Marcel J.M. Pelgrom

Analog-to-Digital Conversion

Second Edition

Springer

Analog-to-Digital Conversion

Marcel J.M. Pelgrom

Analog-to-Digital Conversion

Second Edition

Springer

Marcel J.M.Pelgrom
NXP Semiconductors
Eindhoven
Netherlands

ISBN 978-1-4614-1370-7 ISBN 978-1-4614-1371-4 (eBook)
DOI 10.1007/978-1-4614-1371-4
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012951650

© Springer Science+Business Media, LLC 2010, 2013
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.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 many of the other chapters. Chapter 2 is intended not 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: equations help to understand the way how 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, 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 not only 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 and 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.

Stiphout, the Netherlands, Christmas 2009

Second Edition

In the second edition a number of errors have been corrected and a few topics have been updated. The most important change is the addition of many examples and exercises to assist students in understanding the material.

Stiphout, the Netherlands, Summer 2012

Acknowledgements

Archimedes said: “Give me one fixed point and I will move the Earth.” Home has always served for me as the fixed point from which I could move forward in my work. I owe my wife Elisabeth a debt of gratitude for creating a wonderful home. She herself was once part of this semiconductor world and understands its crazy habits. Yet, the encouragement and support she gave me is invaluable.

This book reflects parts of my 33 years of work in the Philips Natuurkundig Laboratorium and its successor. If there is anything I would call “luck” in my life, it was the opportunity to work in this place. The creativity, energy, opportunities, and people in this laboratory are unique. It is not trivial to create such research freedom in a financially driven industry. My 7 years as a mixed-signal department head have taught me that. Therefore I am truly grateful to those who served in the management of Philips Research and backed me when working on things outside the project scope or looking in unusual directions—just naming here: Theo van Kessel, Kees Wouda, Gerard Beenker, Hans Rijns, and Leo Warmerdam.

A laboratory is just as good as the people who work in it. In my career I met a lot of extraordinary people. They formed and shaped my way of thinking and analyzing problems. They challenged my ideas, took the time to listen to my reasoning, and pointed me in promising directions. I am grateful for being able to use the insights and results of the mixed-signal circuits and systems group. Without the useful discussions and critical comments of the members of this group this book would not exist. However, there are many more colleagues who have contributed in some form.

Without the illusion of being complete, I want to express my gratitude for a pleasant collaboration with: Carel Dijkmans, Rudy van der Plassche, Eduard Stikvoort, Rob van der Grift, Arthur van Roermund, Erik van der Zwan, Peter Nuijten, Ed van Tuijl, Maarten Vertregt, Pieter Vorenkamp, Johan Verdaasdonk, Anton Welbers, Aad Duinmaijer, Jeannet van Rens, Klaas Bult, Govert Geelen, Stephane Barbu, Laurent Giry, Robert Meyer, Othmar Pfarkircher, Ray Speer, John Jennings, Bill Redman-White, Joost Briaire, Pieter van Beek, Raf Roovers, Lucien Breems, Robert van Veldhoven, Kathleen Philips, Bram Nauta, Hendrik van der Ploeg, Kostas Doris, Erwin Janssen, Robert Rutten, Violeta Petrescu, Harry Veendrick, Hans Tuinhout, Jan van der Linde, Peter van Leeuwen, and many others.

This book is based on the lectures in the Philips Center for Technical Training, at universities, and in the MEAD/EPFL courses. I want to thank Prof. Bram Nauta and Prof. Kofi Makinwa for giving me the opportunity to teach at the Universities of Twente and Delft, Prof. Bruce Wooley and Prof. Boris Murmann of Stanford University for their collaboration, and 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

Symbol	Description	Unit symbol
A	Area	cm^2
C	Capacitance	F
C_{ox}	Oxide capacitance	$\mathrm{F/cm}^2$
D_n	Diffusion coefficient of electrons	$\mathrm{cm}^2/\mathrm{sec}$
d_{ox}	Oxide thickness	$\mathrm{cm} = 10^8 \text{Å}$
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	$\mathrm{A/cm}^2$
J_n	Electron current density	$\mathrm{A/cm}^2$
J_p	Hole current density	$\mathrm{A/cm}^2$
K	Substrate voltage influence on the threshold voltage	$\sqrt{\mathrm{V}}$
k	Boltzmann's constant	1.38×10^{-23} J/K
L	Length of transistor gate	cm/µm
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 (300 K)	1.4×10^{10} cm^{-3}
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
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
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/µm
X, Y	General input and output variable	1
x, y, z	Dimension variables	cm/µm
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
ΔP	Difference between two parameters $P_1 - P_2$	
ε	Permittivity in vacuum	8.854×10^{-14} F/cm
$\varepsilon_{\mathrm{ox}}\varepsilon$	Permittivity in silicon dioxide	3.45×10^{-13} F/cm
$\varepsilon_s\varepsilon$	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

Abstract An analog-to-digital converter is an essential building block in modern signal processing. Both the analog-to-digital conversion as well as the digital-to-analog conversion are key functions for optimally exploiting the capabilities of the digital signal processing core. Three basic processes are distinguished in analog-to-digital conversion: the transition between the time-continuous domain and time-discrete domain, the quantization of the signal amplitude, and the relation between physical quantities and numerical quantities.

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 30 years, from mobile phone to internet from medical imaging machines to handheld 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 analog 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

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_1,

© Springer Science+Business Media, LLC 2013

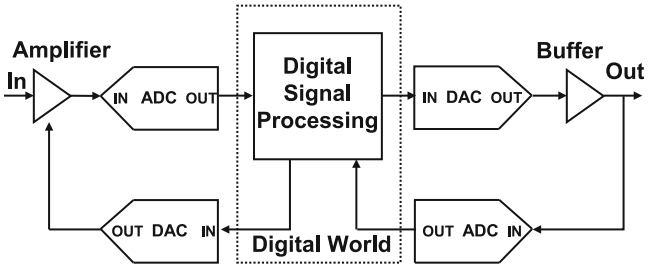


Fig. 1.1 The analog-to-digital and digital-to-analog converters are the ears and eyes of a digital system

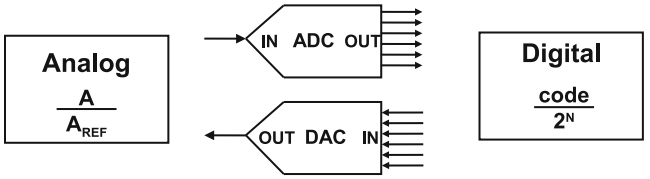


Fig. 1.2 In analog-to-digital conversion a connection is made between the analog world of physical quantities and the digital world of numbers and bits

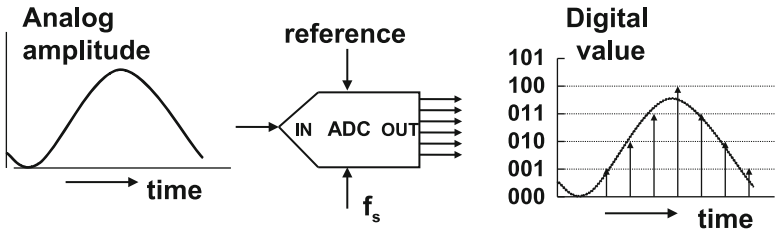


Fig. 1.3 Functions of the analog-to-digital converter: sampling, quantizing, and linking to a reference

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. 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

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 a conversion unit	Reference from a conversion unit

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 setup of this book.

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. 3–6 the three basic functions for conversion are analyzed. Chapter 3 describes the sampling process and gives 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 are nonlinear operations, 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 elaborated in Chap. 8. Oversampling and sigma delta conversion are a special class of conversion techniques and are discussed in Chap. 9.

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 figure-of-merits 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 textbooks review the essential aspects and focus on general principles, circuit realizations, and their merits.

Chapter 2
Components and Definitions

Abstract Mathematics and physics form the basis for electronic design and also for analog-to-digital conversion. This chapter reviews the necessary mathematics and physics. In the mathematics section, the essential transformations of signals, decomposition in harmonic components, and statistical methods are summarized. Some examples are given that are applicable to problems in the following chapters.

The physics sections deal with the behavior of elementary components: resistors, capacitors, and inductors. These components are reviewed with application in integrated circuits in mind. Also the active elements, bipolar and MOS devices, are described and the main model features highlighted. Small-signal and large-signal behavior and stability are discussed. The statistical behavior of large numbers of these devices is reviewed. This chapter concludes with a number of basic circuits, from the differential pair to an oscillator.

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]

[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 the UK.

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_2,

© Springer Science+Business Media, LLC 2013

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]

$n! = 1 \times 2 \times 3 \cdots \times n$	$0! = 1, \quad 1! = 1$
$\binom{n}{m} = \frac{n!}{m!(n-m)!}$	$\binom{n}{n} = \binom{n}{0} = 1$
$(a+b)^n = a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 \ldots$	$(a+b)^2 = a^2 + 2ab + b^2$
$+\binom{n}{n-1}ab^{n-1} + b^n$	$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$
$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1})$	$a^2 - b^2 = (a-b)(a+b)$
$ax^2 + bx + c = 0, \quad x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$	
$\sum_{i=0}^{n} i = \frac{n(n+1)}{2}$	
$\sum_{i=0}^{n} r^i = \frac{1 - r^{n+1}}{1-r}$	$\sum_{i=0}^{\infty} r^i = \frac{1}{1-r} \quad \lvert r \rvert < 1$
$a^b = \mathrm{e}^{b \ln a}$	$a^0 = \mathrm{e}^0 = 1 \quad e = 2.71828$
$\ln(b) = \ln(a) \times {}^a\log(b)$	$\ln(10) = 2.303, \quad {}^{10}\log(10^n) = n$
	${}^{10}\log(2) = 0.301, \quad {}^{10}\log(3) = 0.477$

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

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	$\frac{4}{3}\pi r^3$

Table 2.4 Complex notation [6–8]

$j^2 = -1$	$\lvert a + jb\rvert^2 = a^2 + b^2$
$z = a + jb$	Conjugate(z) : $z^* = a - jb$
Real part$(z) = \mathrm{Re}(z) = \mathrm{Re}(z^*) = a$	Imaginary part$(z) = \mathrm{Im}(z) = -\mathrm{Im}(z^*) = b$,
$\lvert z\rvert^2 = z \times z^* = a^2 + b^2$	$\mathrm{Phase}(z) = \arg(z) = \arctan\left(\frac{\mathrm{Im}(z)}{\mathrm{Re}(z)}\right)$
$\mathrm{e}^{\pm j\alpha} = \cos(\alpha) \pm j\sin(\alpha)$	$\sin(\alpha) = \frac{\mathrm{e}^{j\alpha} - \mathrm{e}^{-j\alpha}}{2j}$, $\cos(\alpha) = \frac{\mathrm{e}^{j\alpha} + \mathrm{e}^{-j\alpha}}{2}$

Table 2.5 Goniometrical relations used in this book

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

The argument of goniometric formulas is expressed in radians $(0 \ldots 2\pi)$, not in degrees [6–8]

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 waveforms describe these events, and Table 2.5 gives a number of regular goniometrical expressions.

Derivatives of functions represent the way a function decreases or increases at every value of its variables. In electronics the derivative helps to understand the behavior of complex functions that are used in a small range of the main variables (Table 2.6).

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

$\frac{\mathrm{d}u(v(t))}{\mathrm{d}t} = \frac{\mathrm{d}u(v)}{\mathrm{d}v}\frac{\mathrm{d}v(t)}{\mathrm{d}t}$	$\frac{\mathrm{d}u(at)}{\mathrm{d}t} = a\frac{\mathrm{d}u(t)}{\mathrm{d}t}$
$\frac{\mathrm{d}}{\mathrm{d}t}(u(t)v(t)) = u(t)\frac{\mathrm{d}v(t)}{\mathrm{d}t} + v(t)\frac{\mathrm{d}u(t)}{\mathrm{d}t}$	$\frac{\mathrm{d}}{\mathrm{d}t}u^n(t) = u^{n-1}(t)\frac{u(t)}{\mathrm{d}t}$
$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{u(t)}{v(t)}\right) = \frac{v(t)\frac{\mathrm{d}u(t)}{\mathrm{d}t} - u(t)\frac{\mathrm{d}v(t)}{\mathrm{d}t}}{v^2(t)}$	$\frac{\mathrm{d}}{\mathrm{d}t}\ln(at) = \frac{a}{t}$
$\frac{\mathrm{d}}{\mathrm{d}t}at^n = ant^{n-1} \quad (n \neq 0)$	$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{at} = a\mathrm{e}^{at}$
$\frac{\mathrm{d}}{\mathrm{d}t}\sin(at) = a\cos(at)$	$\frac{\mathrm{d}}{\mathrm{d}t}\cos(at) = -a\sin(at)$
$\frac{\mathrm{d}}{\mathrm{d}t}\arctan(at) = \frac{a}{1+a^2t^2}$	

Table 2.7 Standard manipulations for integrals of functions [6–8]

$\int x^n\mathrm{d}x = \frac{x^{n+1}}{n+1} + \text{constant} \quad (n \neq -1)$	$\int \frac{1}{x}\mathrm{d}x = \ln(x) + \text{constant}$
$\int \frac{1}{a^2+x^2}\mathrm{d}x = \frac{1}{a}\arctan(x/a) + \text{constant}$	$\int_{x=0}^{\infty} \frac{1}{a^2+x^2}\mathrm{d}x = \frac{\pi}{2a}$
$\int \sin(ax)\mathrm{d}x = \frac{-1}{a}\cos(ax) + \text{constant}$	$\int \cos(ax)\mathrm{d}x = \frac{1}{a}\sin(ax) + \text{constant}$
$\int_{x=-\infty}^{\infty} f(x)\delta(x-x_0)\mathrm{d}x = f(x_0)$	
$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_{t=0}^{x} \mathrm{e}^{t^{-2}}\mathrm{d}t$	$\int_{x=-\infty}^{\infty} \mathrm{e}^{t^{-2}}\mathrm{d}t = \sqrt{\pi}$
$\int_{x=0}^{\infty} \frac{\sin(ax)}{x}\mathrm{d}x = \frac{\pi}{2} \quad (a > 0)$	$\int_{x=0}^{\infty} \frac{\sin(ax)}{x}\mathrm{d}x = \frac{-\pi}{2} \quad (a < 0)0$
$\int_{x=0}^{\infty} \frac{\sin^2(ax)}{x^2}\mathrm{d}x = \frac{a\pi}{2}$	$\int_{x=-\infty}^{\infty} \frac{\sin^2(\pi x)}{(\pi x)^2}\mathrm{d}x = 1$
$\int_{x=0}^{x=L} \sin(2\pi nx/L)\sin(2\pi mx/L)\mathrm{d}x = 0, \quad \text{if } m \neq n$	$\int_{x=0}^{x=L} \sin^2(2\pi nx/L)\mathrm{d}x = 0.5 \quad n,m: \text{ integer}$

Integration of functions is the main mathematical method to form summations over time, area, etc. Elementary manipulation of integrals of functions is given in Table 2.7.

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 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

Table 2.8 Taylor series expansions [6–8]

$(1+a)^n \approx 1+na+\frac{n(n-1)}{2!}a^2+\ldots, \quad \lvert a\rvert \ll 1$	$\frac{1}{1+a} \approx 1-a-a^2\ldots \quad \lvert a\rvert \ll 1$
	$\sqrt{1+a} \approx 1+0.5a, \quad \lvert a\rvert \ll 1$
$e^a \approx 1+a+\frac{a^2}{2!}+\frac{a^3}{3!}, \quad \lvert a\rvert \ll 1$	$\ln(1+a) \approx a-\frac{a^2}{2}+\frac{a^3}{3}, \quad \lvert a\rvert \ll 1$
$\sin(\alpha) \approx \alpha-\frac{\alpha^3}{3!}+\frac{\alpha^5}{5!}$ α in radians, $\lvert\alpha\rvert \ll 1$	$\cos(\alpha) \approx 1-\frac{\alpha^2}{2!}+\frac{\alpha^4}{4!} \quad \lvert\alpha\rvert \ll 1$

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

Table 2.8 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 nonlinear analysis. The Volterra series is a method for including time-dependent effects:

$$\begin{aligned} f(x,t) &= \sum_{i=1}^{\infty}\frac{1}{i!}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} g_i(\tau_1,\ldots,\tau_i)x(t-\tau_1)\times\cdots\times x(t-\tau_i)\mathrm{d}\tau_1\ldots\mathrm{d}\tau_i \\ &= g_0+\frac{1}{1!}\int_{-\infty}^{\infty} g_1(\tau_1)x(t-\tau_1)\mathrm{d}\tau_1+\frac{1}{2!}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(\tau_1)g_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2 \\ &\quad+\frac{1}{3!}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(\tau_1)g_2(\tau_1,\tau_2)g_3(\tau_1,\tau_2,\tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)\mathrm{d}\tau_1\mathrm{d}\tau_2\tau_3\ldots \end{aligned}$$

The first term in the Volterra series is the convolution function as in Eq. 2.149. Laplace and Fourier analogies of the Volterra series are useful to evaluate time-dependent distortion. Various techniques exist to estimate the coefficients g_i.

Example 2.1. Show that $\sin(0.1\pi)*\cos(0.2\pi)=\sin(18^\circ)*\cos(36^\circ)=0.25$

Solution. First multiply numerator and denominator with $\sin(18^\circ)$, then use twice $\sin(\alpha)\cos(\alpha)=\sin(2\alpha)/2$ and on the result $\sin(\alpha)=\cos\left(\frac{\pi}{2}-\alpha\right)$ to get

$$\sin(18^\circ)\cos(36^\circ)=\frac{\sin(18^\circ)\cos(18^\circ)\cos(36^\circ)}{\cos(18^\circ)}=\frac{\sin(36^\circ)\cos(36^\circ)}{2\cos(90^\circ-72^\circ)}=\frac{\sin(72^\circ)}{4\sin(72^\circ)}=0.25.$$

2.1.1 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

$$H(\omega)=\int_{t=-\infty}^{\infty} h(t)\mathrm{e}^{-j\omega t}\mathrm{d}t=\int_{t=-\infty}^{\infty} h(t)\mathrm{e}^{-j2\pi ft}\mathrm{d}t$$

$$h(t) = \frac{1}{2\pi}\int_{\omega=-\infty}^{\infty} H(\omega)\mathrm{e}^{j\omega t}\mathrm{d}\omega = \int_{f=-\infty}^{\infty} H(f)\mathrm{e}^{j2\pi ft}\mathrm{d}f. \tag{2.2}$$

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

$$\begin{aligned} v(\omega) &= j\omega L i(\omega) = j2\pi f L i(\omega) \\ v(\omega) &= R i(\omega) = i(\omega)/g \\ v(\omega) &= \frac{i(\omega)}{j\omega C} = \frac{i(\omega)}{j2\pi f C}, \end{aligned}$$

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 $A(\omega)$ 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:

$$\begin{aligned} A(\omega) &= A^*(-\omega) \qquad \text{or equivalently} \\ \mathrm{Re}(A(\omega)) &= \mathrm{Re}(A(-\omega)) \qquad \text{and} \\ \mathrm{Im}(A(\omega)) &= -\mathrm{Im}(A(-\omega)), \end{aligned}$$

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

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

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

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

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

[2]In modern communication theory, single sideband 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 straightforward Fourier analysis of a physical signal.

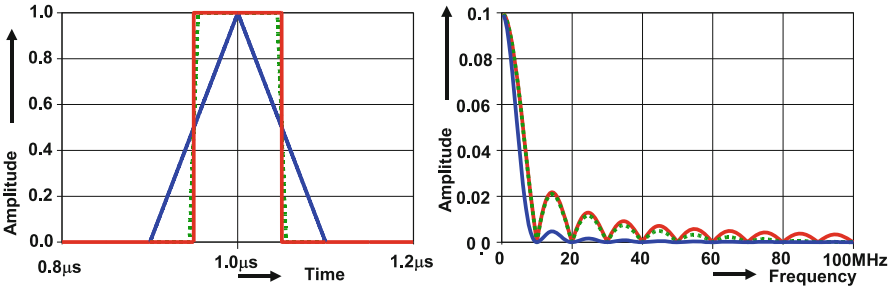


Fig. 2.1 The transform from a box function in the time domain leads to a $\sin(x)/x$ shape in the frequency domain. Three different rise and fall times show the transition from a *block* to a *triangle-shaped* pulse

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, 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

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

with

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

$$a_n = \frac{2}{L}\int_{x=-L/2}^{L/2} f(x)\cos(2\pi nx/L)\mathrm{d}x$$

$$b_n = \frac{2}{L}\int_{x=-L/2}^{L/2} f(x)\sin(2\pi nx/L)\mathrm{d}x.$$

Some more Fourier expansions are given in Table 2.9 [8, p. 405].

Table 2.9 Fourier series expansions of signals repeating at a period T [6–8]

Square wave (transition at $t = 0$)	
$f(t) = \begin{cases} -1 & -T/2 < t < 0 \\ +1 & 0 < t < T/2 \end{cases}$	$f(t) = \sum_{n=1,3,5,\ldots}^{\infty} \frac{4}{n\pi} \sin(2\pi nt/T)$
Square wave symmetrical around $t = 0$	
$f(t) = \begin{cases} -1 & -T/2 < t < -T/4 \\ +1 & -T/4 < t < T/4 \\ -1 & T/4 < t < T/2 \end{cases}$	$f(t) = \sum_{n=1,3,5,\ldots}^{\infty} \frac{4\sin(n\pi/2)}{n\pi} \cos(2\pi nt/T)$
Square wave with T_c high period	
$f(t) = \begin{cases} -1 & -T/2 < t < -T_c/2 \\ +1 & -T_c/2 < t < T_c/2 \\ -1 & T_c/2 < t < T/2 \end{cases}$	$f(t) = \sum_{n=1}^{\infty} \frac{4(-1)^{n+1} \sin(\pi nT_c/T)}{n\pi} \cos(2\pi nt/T)$
Triangle	
$f(t) = \begin{cases} -\frac{2t}{T} & -T/2 < t < 0 \\ +\frac{2t}{T} & 0 < t < T/2 \end{cases}$	$f(t) = \frac{1}{2} - \sum_{n=1,3,5,\ldots}^{\infty} \frac{4}{n^2\pi^2} \cos(2\pi nt/T)$
Saw tooth	
$f(t) = \frac{t}{T} \quad 0 < t < T$	$f(t) = \frac{1}{2} - \sum_{n=1}^{\infty} \frac{1}{n\pi} \sin(2\pi nt/T)$

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

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

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 nonzero 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:

$$10\,^{10}\log\left[1\Big/\left(\frac{\pi^2}{8}-1\right)\right]=6.31\,\mathrm{dB}.$$

Example 2.2. Find the Fourier transform of a square wave in time $f(t)$, defined as

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

that repeats over a period T.

Solution. Evaluation of the integral above shows that $a_0=0$, and $a_n=0$ for all n. Another way of finding this result is to note that this antisymmetrical waveform ($f(t_1)=-f(-t_1)$) must be composed of sine terms with the same antisymmetrical property. The solution cannot contain cosine terms for which symmetry around $t=0$ holds. Also the even sine terms are zero because a sine wave for n is even is identical for $-T/2<t<0$ and $0<t<T/2$. Only the odd b_n coefficients are unequal to zero:

$$b_n=\frac{2}{T}\left[\int_{t=-T/2}^{0}(-1)\sin(2\pi nt/T)\mathrm{d}t+\int_{t=0}^{T/2}(+1)\sin(2\pi nt/T)\mathrm{d}t\right]$$
$$b_n=\frac{-1}{n\pi}\left[(-1)(\cos(0)-\cos(n\pi))+(+1)(\cos(n\pi)-\cos(0))\right]$$
$$=\frac{2(1-\cos(n\pi))}{n\pi}.$$

As expected the resulting terms b_n equal zero for even n and equal $b_n=4/n\pi$ 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):

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

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 $f(t)$ is defined symmetrically around $t=0$:

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

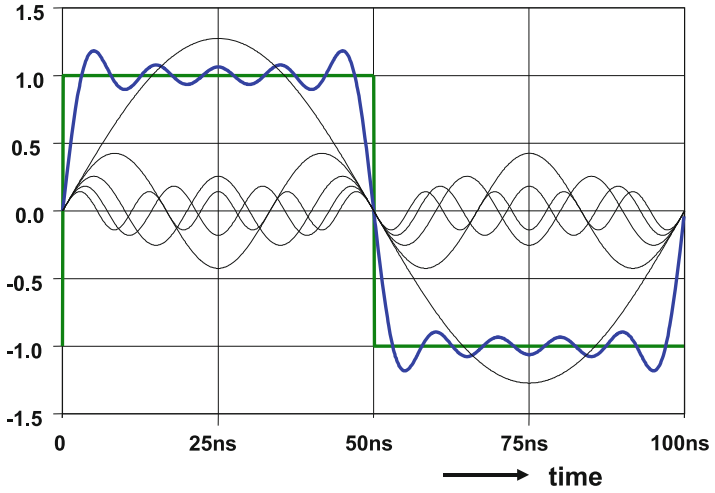


Fig. 2.2 A square wave signal is decomposed in sine waves with frequencies that are odd integers of the fundamental frequency. The first five components are added into an approximation of a square wave

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

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

The even coefficients are zero. The same phase shift that applies to the square wave also makes that the resulting Fourier series experiences a 90° phase shift with respect to the result of first analysis.

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) [9]:

$$\mathrm{THD} = \frac{P_{\mathrm{distortion}}}{P_{\mathrm{fundamental}}}.$$

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]:

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

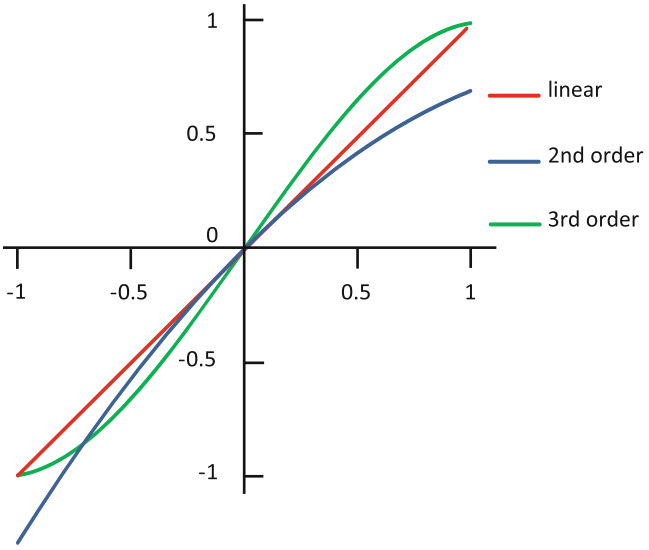


Fig. 2.3 Three transfer curves: $y = x, y = x + ax^2, y = x + bx^3\ a, b < 0$

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

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:

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

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

A transfer function with a small quadratic term $|a| \ll 1$ (see Fig. 2.3):

$$y = x + ax^2$$

will generate 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.

[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.

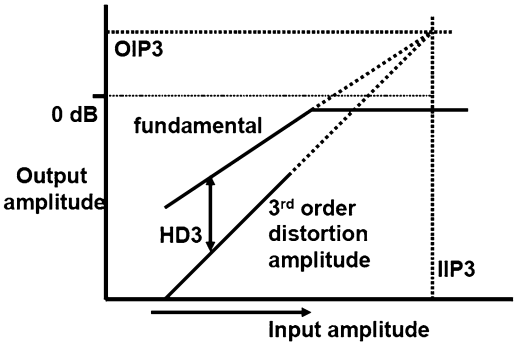


Fig. 2.4 The third-order intercept point is found by extrapolating the fundamental component and the third-order distortion component

The distortion of an input signal $x(t) = V\sin(\omega t)$ is calculated using some goniometric equivalences from Table 2.5[5]:

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

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. In a similar way the third-order distortion is derived:

$$y = x + bx^3$$

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

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. 2.4). 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

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

third-order intercept point can be related to the input axis (IIP3) or the output level (OIP3). Values of IIP3 exceeding 1 volt 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 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

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

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:

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

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. 2.5) is roughly $\frac{3}{4}$ (-2.5 dB) of the magnitude of the third-order distortion HD3.

Example 2.3. If a signal transfer function equals $y = 2x + 0.01x^2$ and $x(t) = 3\sin(\omega t)$, what is the THD in dB?

Solution. The first-order component is $y_1(t) = 6\sin(\omega t)$ and the second-order term $y_2(t) = 0.045\sin(2\omega t)$. So HD2 $= 0.045/6 = 7.510^{-3}$.

The same result follows by using HD2 $= aV/2$ on $y/2 = x + 0.005x^2$.

The THD is found from THD $= 20^{10}\log(\text{HD2}) = -42.5$ dB.

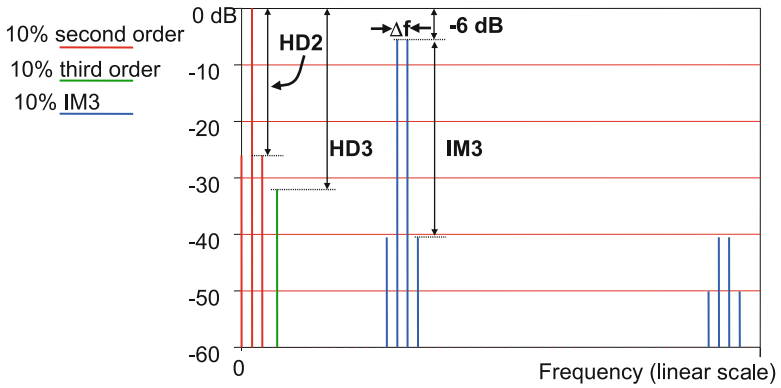


Fig. 2.5 A signal is distorted with $a = b = 0.1$ and gives second- and third-order distortion. A similar distortion is applied to two sine waves at a Δf frequency spacing but at half of the amplitude. This will result in IM3 products at Δf from each carrier

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 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

$$f(s) = \int_{t=0}^{\infty} \mathrm{e}^{-st} f(t)\mathrm{d}t = \int_{t=0}^{\infty} \mathrm{e}^{-\alpha t}\mathrm{e}^{-j\omega t} f(t)\mathrm{d}t. \tag{2.8}$$

The Laplace transform exists if the function $f(t)$ is (piecewise) continuous in the positive time domain and if constants A, B, T exist such that $|f(t)| < A\mathrm{e}^{Bt}$ for $t > T$. The functions $f(t)$ and $f(s)$ have different mathematical formulations, as can be easily seen from Table 2.10. This is reflected in formal textbooks in different typesettings of the function names, e.g., $f(t)$ and $F(s)$ or $\mathbf{f}(s)$. Even more typesettings 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 where the reader might get confused a difference in typesetting 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

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

Table 2.10 Laplace transforms

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

The notation $u(t)$ is the step function at $t = 0$ [6–8]

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 2.10).

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. 2.62).

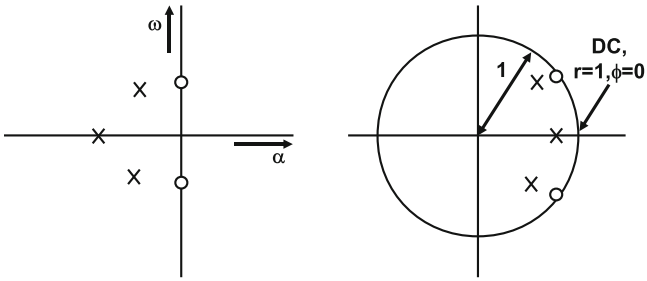


Fig. 2.6 The complex plane for the Laplace transform (s-plane) and the time-discrete plane (z-plane). A real pole, a pair of imaginary poles, and a pair imaginary zeros are depicted

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.,

$$\begin{aligned}\frac{N(s)}{D(s)} &= \frac{s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0} \\ &= \frac{(s+s_{zm})(s+s_{z(m-1)})((s+s_{z2})^2 + s_{z1}^2)}{(s+s_{pn})(s+s_{p(n-1)})((s+s_{p2})^2 + s_{p1}^2)} \\ &= \frac{N_n(s)}{(s+s_{pn})} + \frac{N_{n-1}(s)}{(s+s_{p(n-1)})} + \frac{N_2(s)}{((s+s_{p2})^2 + s_{p1}^2)}.\end{aligned}$$

The roots of the numerator polynomial $(s = -s_{zm}, s = -s_{z2} \pm j s_{z1})$ are called "zeros," and the roots of the denominator polynomial $(s = -s_{pn}, s = -s_{p2} \pm j s_{p1})$ 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 $\mathrm{e}^{-s_{pn}t}$.
- 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 $s = -s_{p2} \pm j s_{p1}$. 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. 2.6 shows 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 feedback.

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.

2.1.5 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/s, $\omega_z \leftrightarrow \omega T_s$ 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 the 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. 2.6 (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:

$$\begin{aligned} z &\leftrightarrow \mathrm{e}^{sT_s} \\ z &\leftrightarrow \mathrm{e}^{j\omega T_s}. \end{aligned} \tag{2.9}$$

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 $f \ll 1/T_s$:

$$\begin{aligned} z &\leftrightarrow 1 + sT_s & s &\leftrightarrow \frac{1}{T_s}(z-1) \\ z &\leftrightarrow \frac{2+sT_s}{2-sT_s} & s &\leftrightarrow \frac{2}{T_s}\frac{(z-1)}{(z+1)}. \end{aligned} \tag{2.10}$$

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 2.7 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:

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

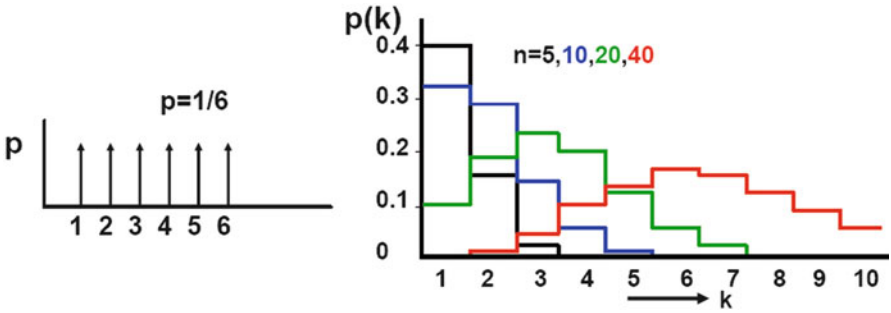


Fig. 2.7 The distribution function of one die, and the binomial distribution for multiple dice with $n = 5, 10, 20, 40$

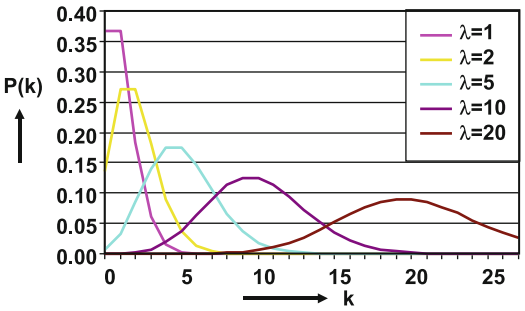


Fig. 2.8 The Poisson distribution with $\lambda = 1, 2, 5, 10$, and 20. The distribution shifts from Poisson to Gaussian for increasing values of the parameter λ

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. 2.7 (right) the probability 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 period. If the average number of expected events is λ then the probability that exactly $k = 0, 1, 2, \ldots$ events will occur is a Poisson distribution:

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

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. 2.8. The normal probability density function or Gaussian distribution has a probability density function (see Fig. 2.8):

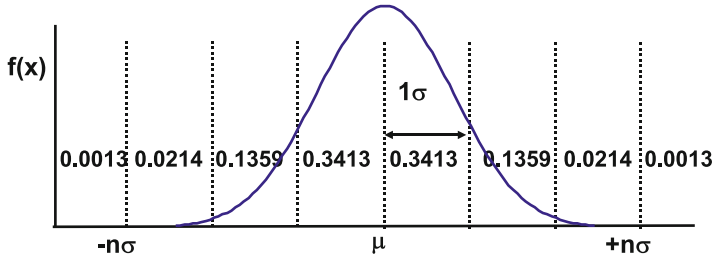


Fig. 2.9 The normal probability distribution with the probability values per 1-σ interval

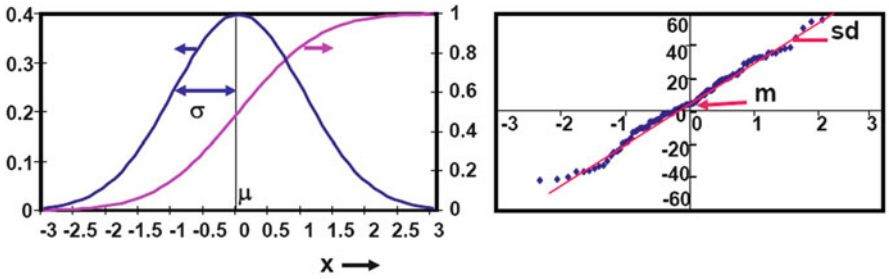


Fig. 2.10 The normal probability distribution and the cumulative probability distribution (*left*) and the normal-scaled cumulative probability distribution (*right*)

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

The discrete variable k has moved into the continuous variable x. A normal distribution 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 $\sigma = np(1-p)$ (Fig. 2.9).

If x is a continuous stochastic variable, the question $p(x = 3)$ has no meaning or equals zero. Nonzero 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:

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

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. 2.10).

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

Table 2.11 Probability that an experimental value exceeds the $n\sigma$ limit on one side in a normal distribution $P((x-\mu) > n\sigma)$[6–8]. The dual-sided rejection is easily found by doubling: $P(|x-\mu| > n\sigma) = 2 \times P((x-\mu) > n\sigma)$

n	P	n	P	n	P	n	P	n	P	n	P
0	0.5000	1.0	0.1587	2.0	0.02275	3.0	$1,350 \times 10^{-6}$	4.0	$31,671 \times 10^{-9}$	5.0	287×10^{-9}
0.1	0.4602	1.1	0.1357	2.1	0.01786	3.1	968×10^{-6}	4.1	$20,657 \times 10^{-9}$	5.1	170×10^{-9}
0.2	0.4207	1.2	0.1151	2.2	0.01390	3.2	687×10^{-6}	4.2	$13,346 \times 10^{-9}$	5.2	100×10^{-9}
0.3	0.3821	1.3	0.0968	2.3	0.01072	3.3	483×10^{-6}	4.3	$8,540 \times 10^{-9}$	5.3	57.9×10^{-9}
0.4	0.3446	1.4	0.0808	2.4	0.00820	3.4	337×10^{-6}	4.4	$5,413 \times 10^{-9}$	5.4	33.3×10^{-9}
0.5	0.3085	1.5	0.0668	2.5	0.00621	3.5	233×10^{-6}	4.5	$3,398 \times 10^{-9}$	5.5	19.0×10^{-9}
0.6	0.2743	1.6	0.0548	2.6	0.00466	3.6	159×10^{-6}	4.6	$2,112 \times 10^{-9}$	5.6	10.7×10^{-9}
0.7	0.2420	1.7	0.0446	2.7	0.00347	3.7	108×10^{-6}	4.7	$1,301 \times 10^{-9}$	5.7	5.99×10^{-9}
0.8	0.2119	1.8	0.0359	2.8	0.00256	3.8	72.3×10^{-6}	4.8	793×10^{-9}	5.8	3.32×10^{-9}
0.9	0.1841	1.9	0.0287	2.9	0.00187	3.9	48.1×10^{-6}	4.9	479×10^{-9}	5.9	1.82×10^{-9}
1.0	0.1587	2.0	0.0228	3.0	0.00135	4.0	31.7×10^{-6}	5.0	287×10^{-9}	6.0	0.987×10^{-9}

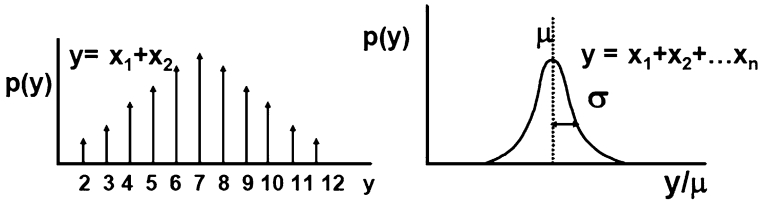


Fig. 2.11 The distribution function, the sum of two dice, and the sum of a large number of dice, illustrating the central limit theorem

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

$$E(x^n) = \int_{x=-\infty}^{\infty} x^n p(x)\mathrm{d}x, \tag{2.15}$$

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. 2.11). If a random variable y is formed by summing k instances of a random variable x: $y = x_1 + x_2 \ldots x_k$ 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

nth-order moment to estimate the variance and mean μ of a distribution:

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

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

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 [10].

The cumulative normal probability distribution in Fig. 2.10 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 into 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. 2.10) 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

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

and the normalized value called the correlation coefficient

$$\text{Correlation coefficient} = \frac{\text{Cov}(x_1, x_2)}{\sigma_{x_1}\sigma_{x_2}}.$$

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

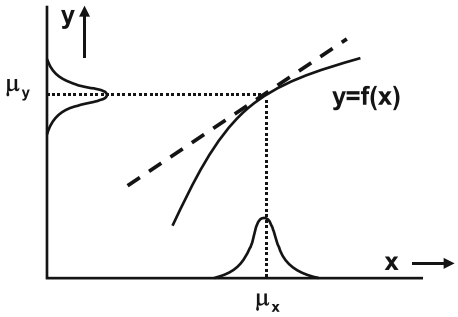


Fig. 2.12 The transformation of one stochastic variable in another via a function $y = f(x)$ uses the Taylor expansion

construct fully dependent stochastic parameters with probability density functions that result in $\mathrm{Cov}(x_1, x_2) = 0$. The mere observation that the covariance between two stochastic variables equals zero is called "uncorrelated."

Example 2.4. Players A and B toss two dice. Player A wins if a "1" or "6" appears or any combination of these; otherwise player B wins. Who will win after many trials?

Solution. This is a classical exam question, and the fast way forward is to invert the problem. In that case the question is what is the probability that both dice show a number in the range "2,3,4,5"? So the probability that player B wins throwing two dice is $4/6 \times 4/6 = 4/9$. Four out of nine times B will win, and A five out of nine.

Example 2.5. What the probability is of having exactly one die showing "6" if six dice are tossed.

Solution. This is a binomial distribution. Use in Eq. 2.11 $n = 6, p = 1/6$, and find $p(k = 1) = 6(1/6)(5/6)^5$. So the probability is 0.4.

Without the formula, the same result is reached by considering that this result requires 5 dice to display one of the numbers "1,2,3,4,5" (probability $= 5/6$) and one to display "6." Actually there are 6 possibilities to reach this result, one for each individual die. The overall probability is therefore $6 \times (1/6) \times (5/6)^5$.

Example 2.6. If a typist makes on average two errors per 10 min what is the probability that he/she makes one error in that time period.

Solution. Now $k = 1, \lambda = 2$ so $p(1,2) = 0.27$. Note that $p(2,2) = 0.27$, $p(3,2) = 0.18$, $p(4,2) = 0.09$, etc.

2.1.7 Functions of Statistical Variables

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. 2.12), the mean and variance of y can be approximated

by the partial derivative using the Taylor series expansion [11]:

$$\mu_y = E(f(x)) \approx f(\mu_x) + \left(\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2}\right)\frac{\mathrm{Var}(x)}{2} = f(\mu_x) + \left(\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2}\right)\frac{\sigma_x^2}{2}$$

$$\mathrm{Var}(y) = \sigma_y^2 = E(f^2(x)) - (E(f(x)))^2 \approx \left(\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right)^2 \mathrm{Var}(x) = \left(\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right)^2 \sigma_x^2.$$

Equivalently the relation $g(x_1, x_2)$, which is a function of two random variables x_1 and x_2, can be calculated:

$$\mathrm{Var}(g) \approx \left(\frac{\mathrm{d}g(x_1)}{\mathrm{d}x_1}\right)^2 \mathrm{Var}(x_1) + \left(\frac{\mathrm{d}g(x_2)}{\mathrm{d}x_2}\right)^2 \mathrm{Var}(x_2) + \left(\frac{\mathrm{d}g(x_1)}{\mathrm{d}x_1}\frac{\mathrm{d}g(x_2)}{\mathrm{d}x_2}\right)\mathrm{Cov}(x_1, x_2).$$

For uncorrelated or independent variables this result reduces to

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

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:

$$\begin{aligned} g(x_1, x_2, x_3, \ldots) &= a_1 x_1 \pm a_2 x_2 \pm a_3 x_3 \ldots \\ E(g) &= a_1 E(x_1) \pm a_2 E(x_2) \pm a_3 E(x_3) \ldots \\ \sigma_g^2 &= a_1^2 \sigma_{x_1}^2 + a_2^2 \sigma_{x_2}^2 + a_3^2 \sigma_{x3}^2 \ldots \end{aligned} \tag{2.19}$$

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.

Example 2.7. A device has a transfer function of the form

$$I = I_0 \mathrm{e}^{\frac{V_g}{V_D}},$$

where V_g is the input variable, I is the output variable, and the other terms are constants. If V_g varies as a normal distribution with $\mu = V_{g0}$ and $\sigma = \sigma_g$ calculate the mean and variance of I.

Solution. This transfer function describes the voltage-to-current transfer of diodes, bipolar transistors, and MOS transistors in weak-inversion regime. Note that the nth derivative of $I(V_g)$ equals IV_D^{-n}. The distribution resulting from a normal distribution in the power term of an exponential function is called a log-normal distribution (see Fig. 2.13). The distribution guarantees positive results but is asymmetrical. Using

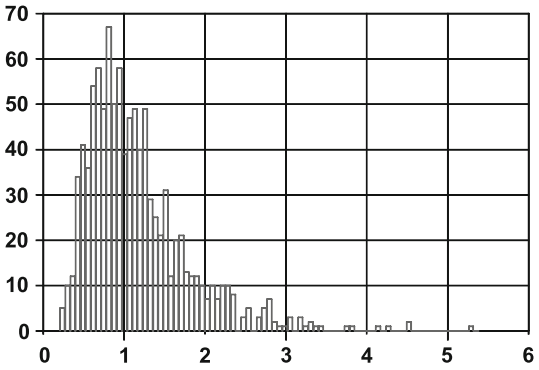


Fig. 2.13 A log-normal distribution generated with $I_0 = 1$ and $\sigma_g = V_D/2$

the equations above,

$$\mu_I = I_0 \mathrm{e}^{\frac{V_{g0}}{V_D}} + \frac{\sigma_g^2}{2V_D^2}$$

$$\sigma_I^2 = \frac{\sigma_g^2}{V_D^2} I_0 \mathrm{e}^{\frac{2V_{g0}}{V_D}}.$$

The mean value shifts. The modulating term is being "rectified" in electronic terms.

Example 2.8. Resistor R_1 with a standard deviation of σ_{R_1} is connected in series with a resistor R_2 with a standard deviation of σ_{R_2}. What is the standard deviation of the series combination if the standard deviations are independent?

Next these two resistors are connected in parallel; what is now the standard deviation?

Solution. For a series connection the following relation holds: $R_{\mathrm{tot}} = R_1 + R_2$. Using Eq. 2.18 yields

$$\sigma_{\mathrm{tot}}^2 = \left(\frac{\mathrm{d}R_{\mathrm{tot}}}{\mathrm{d}R_1}\right)^2 \sigma_{R_1}^2 + \left(\frac{\mathrm{d}R_{\mathrm{tot}}}{\mathrm{d}R_2}\right)^2 \sigma_{R_2}^2$$

with both partial derivatives equal to 1; the expected result is $\sigma_{\mathrm{tot}}^2 = \sigma_{R_1}^2 + \sigma_{R_2}^2$.

A parallel connection is calculated in the same manner:

$$R_{\mathrm{tot}} = \frac{R_1 R_2}{R_1 + R_2} \qquad \sigma_{\mathrm{tot}}^2 = \left(\frac{R_2}{R_1 + R_2}\right)^4 \sigma_{R_1}^2 + \left(\frac{R_1}{R_1 + R_2}\right)^4 \sigma_{R_2}^2 \tag{2.20}$$

Rearranging terms shows

$$\frac{\sigma_{R_{\mathrm{tot}}}^2}{R_{\mathrm{tot}}^2} \frac{1}{R_{\mathrm{tot}}^2} = \frac{\sigma_{R_1}^2}{R_1^2} \frac{1}{R_1^2} + \frac{\sigma_{R_2}^2}{R_2^2} \frac{1}{R_2^2}.$$

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. 2.14) then Ohm's law expresses the current flowing through this piece of material as

$$V = I \times R$$
$$I = G \times V. \tag{2.21}$$

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. 2.15) and in integrated form. The resistor is then constructed in a sheet of material.

Each material is characterized by the intrinsic property "resistivity" or as a symbol: ρ in Ωm. Figure 2.14 shows a rectangular piece of material. Based on the general material property of resistivity, the equivalent resistor value is determined as

$$R = \frac{L\rho}{W \times d} \quad [\Omega] \tag{2.22}$$

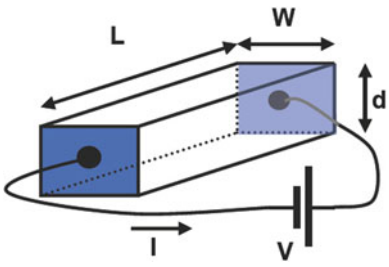
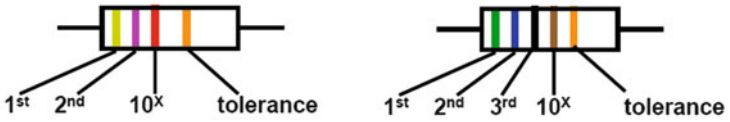


Fig. 2.14 Current flowing through a piece of material



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. 2.15 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 4,700 Ω; the *right-hand* is 5,600 Ω

Table 2.12 Resistivity of (semi) conductors at room temperature; the range indicates values from different sources, e.g., CRC handbook [12, p. e-78] and Eq. 2.68

Material	Resistivity ρ Ωm	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
Polycrystalline silicon	1 Ω–1 k$\Omega/\square$	0.001 K^{-1}

See also Table 2.16

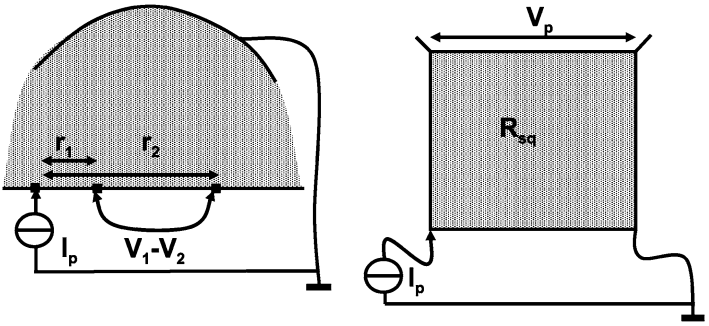


Fig. 2.16 vd Pauw theory and vd Pauw structure for measuring the sheet resistance

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 2.12. 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:

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

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

For general cases van der Pauw [13] 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. 2.16 (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 $2\pi r$:

$$\mathrm{d}V = \frac{-I_p \times R_\square \times \mathrm{d}r}{2\pi r}.$$

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

$$V_1 - V_2 = \int_{V_2}^{V_1} \mathrm{d}V = \int_{r_2}^{r_1} \frac{-I_p R_\square}{2\pi r}\mathrm{d}r = \frac{-I_p R_\square}{2\pi} \ln\left(\frac{r_2}{r_1}\right). \tag{2.23}$$

Based on this theory in IC design the vd Pauw structure is used for measuring sheet resistances. After some mathematical manipulations, the voltage-to-current relation in Fig. 2.16 (right) is found as

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

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

$$V_1 - V_2 = \int_{V_2}^{V_1} \mathrm{d}V = \int_{r_2}^{r_1} \frac{-I_p \rho}{4\pi r^2}\mathrm{d}r = \frac{-I_p \rho}{4\pi}\left(\frac{1}{r_2} - \frac{1}{r_1}\right).$$

In a 10 Ωcm 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

Temperature T is measured in degrees Kelvin °K equal to (T-273.15) °C.[7] 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

$$\text{Energy} = \int_{t=-\infty}^{\infty} P(t)\mathrm{d}t \qquad P(t) = V(t)I(t) = I^2(t)R = \frac{V^2(t)}{R}. \tag{2.24}$$

[7] Better forget about Fahrenheit, Raumur, and Rankine conversion.

Table 2.13 Thermal conductivity, the range indicates values from multiple sources, e.g. [14]

Material	Thermal conductivity κ_{ox} W/Km (Watt per meter per Kelvin)
Aluminum	240
Copper	400
Silicon	130–150
Silicon dioxide	1.1–1.4

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

$$T_R - T_{\text{ambient}} = P \times (K_{m1} + K_{m2} + \cdots) \tag{2.25}$$

where K_{mi} represent the thermal resistances of the ith structure between the power consuming element and the ambient in degrees Kelvin per Watt.

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 (Table 2.13).

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

$$K_m = \frac{t_{ox}}{\kappa_{ox} W L}, \tag{2.26}$$

where t_{ox} is the layer thickness and W and L the lateral dimensions. A heat source of $1 \times 1\,\mu\text{m}$ on top of a $1\,\mu\text{m}$ silicon dioxide layer sees a thermal resistance of 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.

Example 2.9. A chip consuming 0.8 W is mounted in a package with a thermal resistance of $K_{pa} = 40$ K/W between junction and the PCB. The PCB shows a thermal resistance of $K_{pcb} = 10$ K/W towards the equipment casing. What will be the junction temperature if the casing is at 30 °C?

Solution. The chip junction temperature will rise $0.8\,\mathrm{W} \times 10\,\mathrm{K/W} + 0.8\,\mathrm{W} \times 40\,\mathrm{K} = 40\,\mathrm{K}$ over the casing temperature and reach 70 °C.

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:

$$\begin{aligned} R(V) &= R(V_0)(1 + \mathrm{VC}(V - V_0)) \\ \mathrm{VC} &= \frac{1}{R}\frac{\mathrm{d}R}{\mathrm{d}V}. \end{aligned} \tag{2.27}$$

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

$$\begin{aligned} R(T) &= R(T_0)(1 + \mathrm{TC}(T - T_0)) \\ \mathrm{TC} &= \frac{1}{R}\frac{\mathrm{d}R}{\mathrm{d}T}. \end{aligned} \tag{2.28}$$

T_0 is the reference temperature, and T the actual temperature. The temperature coefficient for some materials is given in Table 2.12.

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 measuring voltage creates a current. Next to that all kinds of parasitic resistances are present in the measurement loop. The measurement technique shown in Fig. 2.17 (left) supplies the current to the resistor via a path separate from the voltage measurement. The four-point technique or Kelvin measurement uses the 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

$$V_R = e\left(\frac{R_2}{R_1 + R_2} - \frac{R_4}{R_3 + R_4}\right). \tag{2.29}$$

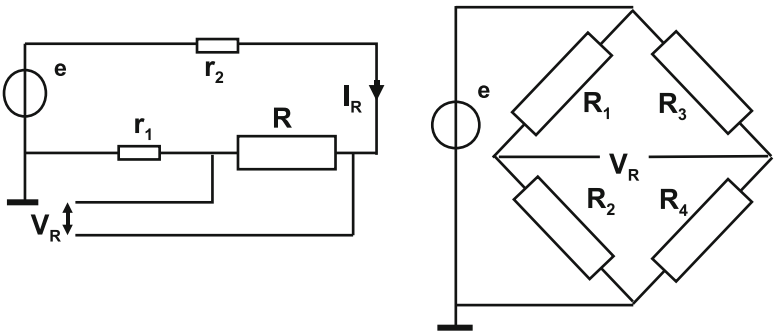


Fig. 2.17 Resistance measurement via a four-point method (*left*) and a Wheatstone bridge (*right*)

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: $R_1R_4 = R_2R_4$. 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 [15] :

$$t_{50} = K_{\text{em}} J^{-n} \mathrm{e}^{(E_a/kT)}, \tag{2.30}$$

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: $1.38 \times 10^{-23}\,\text{m}^2\text{kgs}^{-2}\,\text{K}^{-1}$
E_a = 0.6–0.7 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 [16].

2.2.5 Noise

In electronic devices the flow of current is associated with noise [17]. 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}:

$$\begin{aligned}
\text{thermal noise density in 1 Hz:} S_{vv}(f) &= 4kTR \\
\text{or:} S_v &= \sqrt{4kTR}\,[\mathrm{V}\sqrt{\mathrm{Hz}}] \\
\text{rms noise voltage} v_{\text{noise,rms}} &= \sqrt{\int_{f=f_{\text{low}}}^{f=f_{\text{high}}} 4kTR\,\mathrm{d}f} \\
&= \sqrt{4kTR(f_{\text{high}} - f_{\text{low}})} = \sqrt{4kTR\mathrm{BW}} \quad [\mathrm{V}],
\end{aligned}$$

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$\sqrt{\mathrm{Hz}}$ at room temperature and 1 MΩ results in 126 nV$\sqrt{\mathrm{Hz}}$. 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. 2.18). These two models are equivalent following the Norton theorem:

$$\begin{aligned}
v_{\text{noise,rms}} &= \sqrt{4kTR\mathrm{BW}} \quad [\mathrm{V}] \\
i_{\text{noise,rms}} &= \sqrt{4kT\mathrm{BW}/R} \quad [\mathrm{A}].
\end{aligned} \tag{2.31}$$

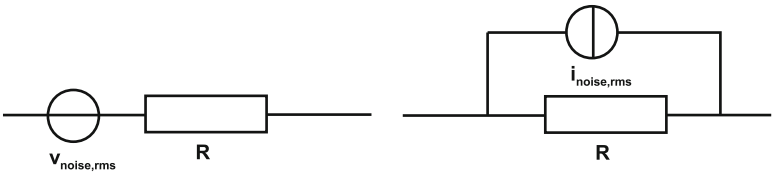


Fig. 2.18 Equivalence of noise voltage and noise current

"$1/f$" 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 autocorrelation time constant. Combining the effect of a multitude of these traps with a uniform distribution of autocorrelation 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

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

$$S_v(f) = \sqrt{\frac{K_{1/f}}{\text{area} \times f}} \quad [\text{V}\sqrt{\text{Hz}}], \tag{2.32}$$

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

$$i_{\text{noise}} = \sqrt{2qI_{\text{bias}}} \quad [\text{A}\sqrt{\text{Hz}}]. \tag{2.33}$$

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 (Sect. 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

$$\text{SNR} = 10\,{}^{10}\log\left(\frac{\text{Signal power}}{\text{noise power}}\right) = 20\,{}^{10}\log\left(\frac{V_{\text{signal,rms}}}{V_{\text{noise,rms}}}\right). \tag{2.34}$$

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.

Example 2.10. A resistor of $100\,\Omega$ is connected in parallel to a resistor of $500\,\Omega$. What is the equivalent resistance? What is the noise spectral density in $\text{V}/\sqrt{\text{Hz}}$? The $100\,\Omega$ resistor is heated to 127 °C; the $500\,\Omega$ resistor is cooled to −173 °C. What is the noise spectral density in $\text{V}/\sqrt{\text{Hz}}$?

Solution.

$$R_{\text{tot}} = 100\,\Omega//500\,\Omega = \frac{100\times 500}{100+500} = 83.3\,\Omega$$

$$v_{\text{noise}} = \sqrt{4kTR_{\text{tot}}} = 1.17\,\text{nV}/\sqrt{\text{Hz}}.$$

To answer the last question the resistors must be treated individually, and there are two options: use a noise voltage source per resistor or a noise current source. The current source solution is easier; it allows a simple root-mean-square addition of the two parallel current sources loaded with the parallel resistance:

$$v_{\text{noise}} = R_{\text{tot}}\sqrt{\frac{4kT_1}{R_1}+\frac{4kT_2}{R_2}} = 1.27\,\text{nV}/\sqrt{\text{Hz}}.$$

2.3 Maxwell Equations

Electrostatic and electromagnetic events are described by Maxwell equations. These reflect the state of the art in electromagnetic field theory in 1864. The Maxwell equations are the fundament for describing electronic elements and semiconductor devices (see, e.g., [18]).

Electromagnetic fields are represented by vectors.[8] $\mathbf{D}(x,y,z)$ is the electric displacement vector in Coulomb/m^2 at any point in space, which is related to the electric field vector $\mathbf{E}(x,y,z)$ as $\mathbf{D} = \varepsilon_r \varepsilon_0 \mathbf{E}$ in linear materials with components in the x,y,z directions $\mathbf{E} = [E_x, E_y, E_z]$. The surface A is represented by its normal vector $\mathbf{A}$.

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

$$\nabla \cdot \mathbf{E}(x,y,z) = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}. \tag{2.35}$$

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 $\mathbf{E}$ as

$$\nabla \times \mathbf{E}(x,y,z) = \left[\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}, \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}, \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \right]. \tag{2.36}$$

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

The first Maxwell equation is also known as Gauss law:

$$\nabla \cdot \mathbf{D} = \varepsilon_r \varepsilon_0 \nabla \cdot \mathbf{E} = \rho, \qquad \oint \varepsilon_r \varepsilon_0 \mathbf{E} \cdot \mathrm{d}\mathbf{A} = \iiint_{\text{volume}} \rho(x,y,z)\, \mathrm{d}x\mathrm{d}y\mathrm{d}z. \tag{2.37}$$

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.[9]

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 = $8.8542 \times$

[8] 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.

[9] 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.

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

$$\chi_e = \varepsilon_r - 1.$$

If a point charge Q_p is surrounded by a sphere with a radius r, the electric field $\mathbf{E}(x,y,z)$ is a vector formed by the line between the point charge and the location where the field 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.[10] The field strength E_p at distance r is found with Eq. 2.37 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

$$F_{12} = \frac{Q_p Q_s}{4\pi r_{12}^2 \varepsilon_r \varepsilon_0}. \tag{2.38}$$

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

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

$$\nabla^2 \cdot V = -\nabla \cdot \mathbf{E} = \frac{-\rho}{\varepsilon_r \varepsilon_0}. \tag{2.39}$$

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

$$V(r_2) - V(r_1) = \int_{r=r_1}^{r_2} -E(r)\mathrm{d}r = \frac{Q}{4\pi\varepsilon_r\varepsilon_0}\left(\frac{1}{r_2} - \frac{1}{r_1}\right). \tag{2.40}$$

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

$$\nabla \cdot \mathbf{B} = 0, \quad \oint \mathbf{B} \cdot \mathrm{d}\mathbf{A} = 0. \tag{2.41}$$

B is the vector of the magnetic flux density in Tesla or $\mathrm{Vs/m^2} = \mathrm{N/Am}$ (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 **B** is related to the magnetic field strength

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

$\mathbf{H}$ as $\mathbf{B} = \mu_r \mu_0 \mathbf{H}$. 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. Another 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 μ_r =600...10,000.

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

$$\nabla \times \mathbf{E} = -\frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t}, \quad \oint \mathbf{E} \cdot \mathrm{d}\mathbf{s} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}. \tag{2.42}$$

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

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

$$\nabla \times \mathbf{B} = \mu_r \mu_0 \mathbf{J} + \varepsilon_r \varepsilon_0 \mu_r \mu_0 \frac{\mathrm{d}\mathbf{E}}{\mathrm{d}t}, \qquad \oint \mathbf{B} \cdot \mathrm{d}\mathbf{s} = \mu_r \mu_0 \mathbf{I} + \varepsilon_r \varepsilon_0 \mu_r \mu_0 \frac{\mathrm{d}}{\mathrm{d}t} \oint \mathbf{E} \cdot \mathrm{d}\mathbf{A}. \tag{2.43}$$

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:

$$\mathrm{d}\mathbf{B} = \frac{\mu_0}{4\pi r^2} \mathbf{I} \times \mathbf{r}_u \mathrm{d}y. \tag{2.44}$$

$\mathbf{B}(x,y,z)$ is a vector of the magnetic field due to a current I in a small portion of wire dy. $\mathbf{r}_u$ is the unit-length vector pointing from the wire fraction dy to the point where the field $\mathbf{B}$ is measured. $\mathbf{I} \times \mathbf{r}_u$ is the outer product between the vector $\mathbf{r}_u$ 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 $\mathbf{I}, \mathbf{r}_u$. Intuitively $I\mathrm{d}y$ can be interpreted as a magnetic charge and dB is the resulting magnetic field on a sphere with radius r.

The four Maxwell laws reduce in vacuum to the Heaviside formulation:

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 0 \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t} \\ \nabla \times \mathbf{B} &= \varepsilon_0 \mu_0 \frac{\mathrm{d}\mathbf{E}}{\mathrm{d}t}, \end{aligned} \tag{2.45}$$

where **E** and **B** are functions of time t and space x, y, z. Substitution leads to

$$\varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} - \nabla^2 \mathbf{E} = 0,$$
$$\varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} - \nabla^2 \mathbf{B} = 0.$$

The combination of the magnetic $\mathbf{B}(t,x,y,z)$ field and the electric $\mathbf{E}(t,x,y,z)$ field defines the electromagnetic wave propagation. The solutions for **B** and **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 $f\lambda = c$.

Electromagnetic waves travel with the velocity of light c, which equals 2.99792×10^8 m/s for vacuum. The velocity of electromagnetic 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 equations couple the magnetic and electric field also in another way. The magnetic field and the electrical field define a characteristic impedance:

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

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.

2.3.1 Inductors

The magnetic field in a coil in Fig. 2.19 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 then 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:

$$\oint \mathbf{B} \cdot \mathrm{d}\mathbf{s} = BL_c = \mu_0 N_w I.$$

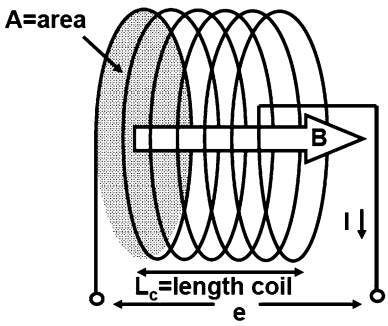


Fig. 2.19 Definitions in a coil

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:

$$L = \frac{N_w \Phi}{I} = \frac{N_w B A}{I} = \frac{\mu_0 N_w^2 A}{L_c} \tag{2.46}$$

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

Faraday's law defines the electromagnetic induction or electromagnetic force[11] "emf" e in a conductor that encloses a magnetic flux (Φ) as

$$e(t) = -\frac{\mathrm{d}\Phi(t)}{\mathrm{d}t}. \tag{2.47}$$

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. 2.19 the number of windings N_w multiplies this potential to yield

$$|e(t)| = N_w \frac{\mathrm{d}\Phi(t)}{\mathrm{d}t} = A N_w \frac{\mathrm{d}B(t)}{\mathrm{d}t} = \frac{A \mu_0 N_w^2}{L_c} \frac{\mathrm{d}I(t)}{\mathrm{d}t} = L \frac{\mathrm{d}I(t)}{\mathrm{d}t}, \tag{2.48}$$

$$I(t) = \frac{1}{L} \int_{\tau=0}^{\tau=t} e(\tau) \mathrm{d}\tau. \tag{2.49}$$

[11] 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.

A sinusoidal current $i(t) = i_0 \sin(2\pi ft) = i_0 \sin(\omega t)$ results in a voltage over the coil as $v(t) = Li_0 \cos(2\pi ft)$. 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

$$E_L = \int_{t=0}^{\infty} e(t)I(t)\mathrm{d}t = \int_{t=0}^{\infty} I(t)\frac{\mathrm{d}LI(t)}{\mathrm{d}t}\mathrm{d}t = \int_{I=0}^{I_L} LI\mathrm{d}I = 0.5LI_L^2. \tag{2.50}$$

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 [19].[12] 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 2.20 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:

$$\mathrm{d}B(x,y) = \frac{\mu_0}{4\pi r^2} I \times r_u \mathrm{d}y = \frac{\mu_0}{4\pi(x^2 + (y - y_B)^2)} I \sin(\theta)\mathrm{d}y = \frac{\mu_0 I}{4\pi} \frac{x\mathrm{d}y}{(x^2 + (y - y_B)^2)^{3/2}}.$$

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], [12, Sect. A]:

[12] 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.

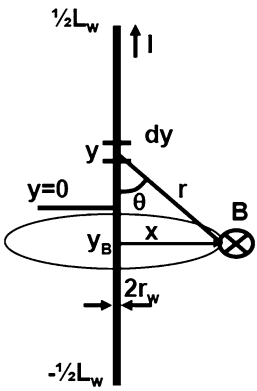


Fig. 2.20 Definitions of magnetic field along a wire

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

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

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:

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

$$= \frac{\mu_0 I}{4\pi} \int_{x=r_w}^{x=\infty} \frac{2(\sqrt{x^2+L_w^2}-x)}{x} \mathrm{d}x = \frac{\mu_0 I}{2\pi} \left[\sqrt{x^2+L_w^2} - L_w \ln\left(\frac{L_w + \sqrt{L_w^2+x^2}}{x} \right) - x \right]_{x=r_w}^{x=\infty}$$

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

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:

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

with L_w and r_w as the length and radius of the wire in meter. Assuming[13] 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:

$$\delta = \frac{1}{\sqrt{\pi \sigma f \mu}}, \tag{2.51}$$

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

[13] Implicitly the return path is estimated at a distance of a few hundred times the wire radius.

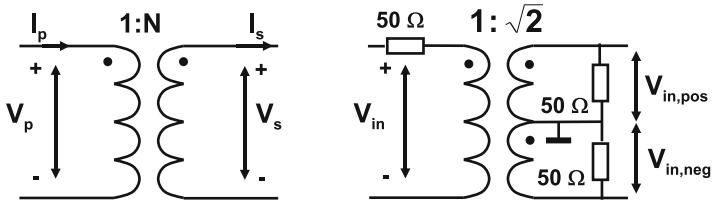


Fig. 2.21 Notation conventions in a transformer and a transformer used to create differential signals. The *dots* indicate the side of the coils with mutually in-phase signals

$$\frac{V_p}{V_s} = \frac{I_s}{I_p} = \frac{\text{number of primary windings}}{\text{number of secondary windings}} = \frac{1}{N}. \tag{2.52}$$

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 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 2.21 (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[14] 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 $0 \leq k \leq 1$ determines the contribution of the mutual inductance:

$$M = k\sqrt{L_p L_s}. \tag{2.53}$$

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

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

[14] Many manufacturers offer RF transformers, e.g., M/A-COM, Mini-Circuits, Coilcraft, and Pulse.

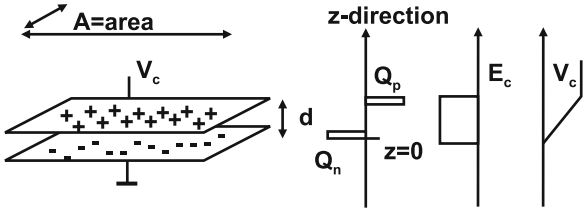


Fig. 2.22 Parallel-plate capacitor

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.

2.3.6 *Capacitors*

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

$$C = \frac{\mathrm{d}Q}{\mathrm{d}V} \quad [\mathrm{F}]$$

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

In Fig. 2.22 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 Eq. 2.37 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

$$E = \iiint\limits_{\text{volume}} \frac{\rho}{A\varepsilon_r\varepsilon_0}\mathrm{d}x\mathrm{d}y\mathrm{d}z = \begin{cases} 0, & z < 0 \\ \dfrac{Q}{\varepsilon_r\varepsilon_0 A}, & 0 < z < d \\ 0, & z > d. \end{cases}$$

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:

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

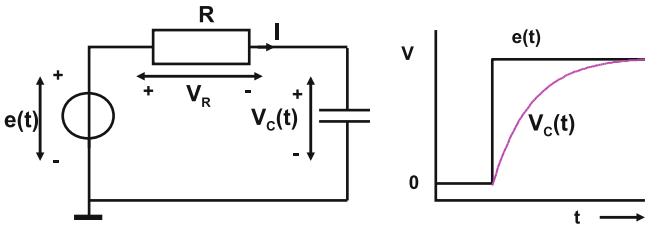


Fig. 2.23 Charging a capacitor

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

$$C = \frac{Q}{V} = \frac{A\varepsilon_r\varepsilon_0}{d}. \tag{2.54}$$

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:

$$I(t) = \frac{\mathrm{d}Q(t)}{\mathrm{d}t} = \frac{\mathrm{d}CV(t)}{\mathrm{d}t} = C\frac{\mathrm{d}V(t)}{\mathrm{d}t} + V\frac{\mathrm{d}C(t)}{\mathrm{d}t}, \tag{2.55}$$

$$V(t) = \frac{1}{C}\int_{\tau=0}^{\tau=t} I(\tau)\mathrm{d}\tau. \tag{2.56}$$

The mathematical equivalence with Eq. 2.49 is obvious.

For time-invariant capacitors only the first term remains.

With a sinusoidal voltage $v(t) = v_0\sin(2\pi ft) = v_0\sin(\omega t)$ the current through is given as $i(t) = Cv_0\cos(2\pi ft)$. 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. 2.23, the voltage over the capacitor can be calculated as follows:

$$I(t) = C\frac{\mathrm{d}V_C(t)}{\mathrm{d}t} \quad \text{and} \quad I(t) = \frac{V - V_C(t)}{R} \quad \text{result in}$$

$$V_C(t) = V(1 - \mathrm{e}^{-t/RC})$$

$$I(t) = \frac{V}{R}\mathrm{e}^{-t/RC}. \tag{2.57}$$

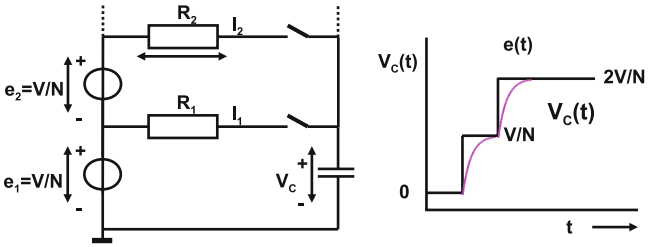


Fig. 2.24 Charging a capacitor from multiple sources

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

$$\begin{aligned} E_e &= \int_{t=0}^{\infty} e(t)I(t)\mathrm{d}t = CV^2 \\ E_R &= \int_{t=0}^{\infty} V_R(t)I(t)\mathrm{d}t = 0.5CV^2 \\ E_C &= \int_{t=0}^{\infty} V_C(t)I(t)\mathrm{d}t = 0.5CV^2. \end{aligned} \tag{2.58}$$

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.

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. 2.24. 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:

$$\begin{aligned} E_{e1} &= CV^2/N^2 \\ E_{R_1} &= 0.5CV^2/N^2 \\ E_C &= 0.5CV^2/N^2. \end{aligned}$$

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

$$
\begin{aligned}
E_{e1} &= CV^2/N^2 + CV^2/N^2 \\
E_{e2} &= CV^2/N^2 \\
E_{R_1} &= 0.5CV^2/N^2 \\
E_{R_2} &= 0.5CV^2/N^2 \\
E_C &= 0.5C(2V/N)^2.
\end{aligned}
$$

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

$$
\begin{aligned}
E_{e1\ldots eN} &= \frac{(N+1)}{2N}CV^2 \\
E_{R_1\ldots RN} &= \frac{1}{2N}CV^2 \\
E_C &= \frac{1}{2}CV^2.
\end{aligned}
$$

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 voltages and control the timing. In audio processing a similar method for transferring efficiently energy to a load is called "class G." Also some charge-based analog-to-digital converters use this basic idea.

Example 2.11. A capacitor with value C is discharged (0 V). A second capacitor of the same size is charged to a voltage V_E. 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?

Solution. The energy stored in a capacitor is $E = 0.5CV^2$. In this case the total energy is $E_{\text{tot}} = 0 + 0.5CV_E^2$. Capacitors in parallel share their charge. So the total charge $(0 + CV_E)$ is distributed over 2 capacitors. The resulting voltage is therefore $V_{E,2} = CV_E/2C = V_E/2$, and the total resulting energy is $E_{\text{tot},2} = 0.5 \times 2C \times (V_E/2)^2 = 0.25CV_E^2$. Is there energy lost?

Equivalent is the thermodynamic problem of two equal-sized containers, each filled with the same number of gas molecules. One container is at 0 °C while the second is at 100 °C. At the moment the containers are connected the temperature of the total gas becomes 50 °C and the lost energy is the increased entropy of the gas.

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,

Table 2.14 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

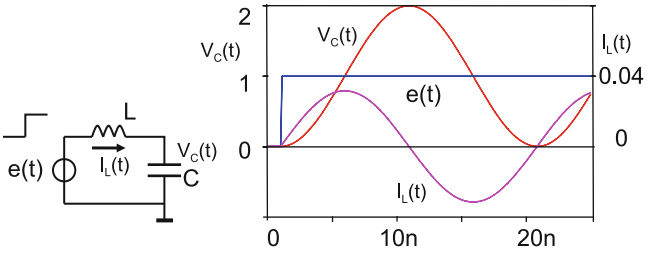


Fig. 2.25 A step response on an *LC* circuit. *Right*: the excitation $e(t)$ is a step function; the voltage on the capacitor $V_C(t)$ swings from 0 to twice the step, while the current through the coil $I_L(t)$ leads over the voltage swing

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

$$P_d = f_d N_d C_d V^2. \tag{2.59}$$

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 2.14). 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. 2.45 and in Table 11.1.

Example 2.12. An *LC* circuit is connected to a voltage source. Explain the voltage behavior on the capacitor on a stepwise voltage change.

Solution. Figure 2.25 shows the circuit. At $t = 0$ the capacitor and the coil are at zero voltage and zero current. The coil will oppose changing the current and the capacitor will tend to keep the voltage constant. So the voltage of the step excursion will appear over the coil while $V_C(t = 0+) = 0$. As shown in Eq. 2.49 the voltage over the coil will cause the magnetic field to build up and the current through the coil

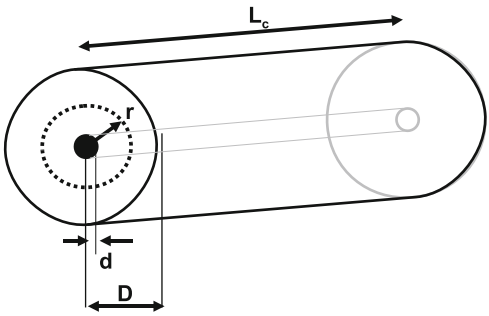


Fig. 2.26 Definitions for a coaxial cable

to increase linear with time. This current will be integrated according to Eq. 2.56 in the capacitance and this double integration results in a parabolic shape of V_C starting from $t = 0$. As the voltage over the capacitor increases, the voltage over the coil $e(t) - V_C(t)$ decreases and consequently the buildup of the current in the coil will stop at the point where $e(t) = V_C(t)$. The current in the coil has reached its maximum and continues charging the capacitor. Now the inverse process starts: a negative voltage over the coil will reduce the current. From symmetry with the first cycle, it is clear that the current will reach $I_L(t) = 0$ at the point where $V_C(t)$ has reached the double amplitude of the step. Now the negative voltage over the coil will start building up a current in the opposite direction that discharged the capacitor. An oscillation is started.

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. 2.26 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 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 Eq. 2.37,

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

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

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

The capacitance of a unit length of coaxial cable is now

$$C_{\text{unit}} = \frac{Q}{V} = \frac{2\pi\varepsilon_r\varepsilon_0}{\ln\frac{D}{d}} \quad [\text{F/m}]. \tag{2.60}$$

Referring to the coaxial cable in Fig. 2.26 the inductance of a coaxial cable per unit length carrying a current I can be derived:

$$\oint B \times \mathrm{d}s = \mu_r\mu_0 I.$$

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

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

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

$$L_{\text{unit}} = \frac{\Phi}{I} = \frac{\mu_r\mu_0 \ln\frac{D}{d}}{2\pi} \quad [\text{H/m}]. \tag{2.61}$$

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

$$Z_0 = \sqrt{\frac{L}{C}} = \frac{\ln\frac{D}{d}}{2\pi}\sqrt{\frac{\mu_r\mu_0}{\varepsilon_r\varepsilon_0}} = 59.988 \times \ln\frac{D}{d} = 138 \times {}^{10}\log\frac{D}{d} \quad [\Omega]. \tag{2.62}$$

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

$$c_{\text{coax}} = \frac{1}{\sqrt{L_{\text{unit}}C_{\text{unit}}}} = \frac{c}{\sqrt{\varepsilon_r\mu_r}}. \tag{2.63}$$

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

$$n \times p = n_i^2 = N_C N_V \mathrm{e}^{-E_G/kT} \propto T^3 \mathrm{e}^{-E_G/kT}. \tag{2.64}$$

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 bandgap prevents electrons to go to the conduction band. In silicon E_G equals 1.12 electronVolt (eV) at 0 K and in germanium the bandgap 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" bandgap 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 are 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

$$p_n = \frac{n_i^2}{N_D} \qquad n_p = \frac{n_i^2}{N_A}. \tag{2.65}$$

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

$$n = n_i \mathrm{e}^{(E_F - E_i)/kT} \qquad p = n_i \mathrm{e}^{(E_i - E_F)/kT}, \tag{2.66}$$

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 constant and equals 1.38×10^{-23} J/K. T is the absolute temperature with T °K.

The above mathematical description is derived from Maxwell–Boltzmann statistics. These statistics were originally formulated for gases, where the Boltzmann constant is linked to the ideal gas R constant via $R = kN_{\text{Avogadro}}$. The Boltzmann constant is the energy increase per unit temperature increase per particle. Nearly all

occurrences of exponential relations in semiconductor descriptions are derived from these statistics. It is important to realize how this concept differs from Newton'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. An energy 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 US literature the term "mil" is used. 1 mil equals 1/1000 in. 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 $n_i = p_i = 1.4 \times 10^{10}\,\mathrm{cm}^{-3}$. 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 is 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:

$$\rho_{R,n} = \frac{1}{qn\mu_n} \qquad \rho_{R,p} = \frac{1}{qp\mu_p} \tag{2.67}$$

q is the charge of an electron $q = 1.6 \times 10^{-19}$ Coulomb. The proportionality constants $\mu_n \approx 1{,}500\,\mathrm{cm}^2/\mathrm{Vs}$ and $\mu_p \approx 450\,\mathrm{cm}^2/\mathrm{Vs}$ are the mobility for electrons and holes, respectively (see Table 2.15). 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. 2.27 [14, p. 32].

Table 2.15 Some electrostatic properties of semiconductors [14]

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_3N_4	7.5	5	1,000	
Germanium Ge	16.0	0.67	8	3,900/1,900
Gallium arsenide GaAs	13.1	1.42	35	8,500/400
Gallium nitride GaN	9.5	3.36	330	380/

The permittivity in vacuum $\varepsilon_0 = 8.85 \times 10^{-12}$ F/m

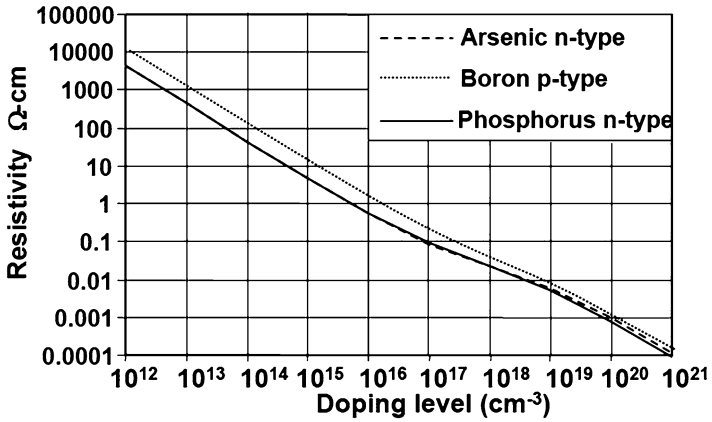


Fig. 2.27 Resistivity of doped silicon: *upper line* is boron (*p*-type); *lower two lines* represent phosphorus and arsenic (*n*-type)

2.4.2 Voltage and Temperature Coefficient

In doped silicon the temperature coefficient (see Eq. 2.28) 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 $T^{1.5}$ 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

$$\mathrm{TC} = \frac{1}{\rho_R}\frac{\mathrm{d}\rho_R}{\mathrm{d}T} = \frac{-\alpha}{T} \approx 0.007\,^{\circ}\mathrm{C}^{-1}. \tag{2.68}$$

A minor dependence with doping is present. In polycrystalline silicon the effects of grain boundaries play a role, strongly influencing the temperature coefficients in poly silicon [21].

Table 2.16 An indication of resistor characteristics in a semiconductor process (0.18 µm to 65 nm generations) from ITRS [20] and various publications

Material	Square resistance $\Omega/\square$	Voltage coeff V^{-1}	Temp coeff K^{-1}	Matching A_R $\%\mu m$
n/p Diffusion	75...125	1×10^{-3}	$1..2 \times 10^{-3}$	0.5
n-Well diffusion	1,000	80×10^{-3}	4×10^{-3}	
n-Polysilicon	50...150		$-1 \ldots +1 \times 10^{-3}$	2
p-Polysilicon	50...150		0.8×10^{-3}	2
Polysilicon (silicide)	3...5		3×10^{-3}	
Aluminum	0.03...0.1		3×10^{-3}	

It is obvious that resistors with voltage dependency affect the distortion behavior of an electronic circuit. These effects are modeled with Eq. 2.27. Table 2.16 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 layout 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

$$\frac{\sigma_{\Delta R}}{R} = \frac{A_R}{\sqrt{\text{area}}}, \tag{2.69}$$

where A_R is expressed in %µm and area in μm^2.

2.4.4 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

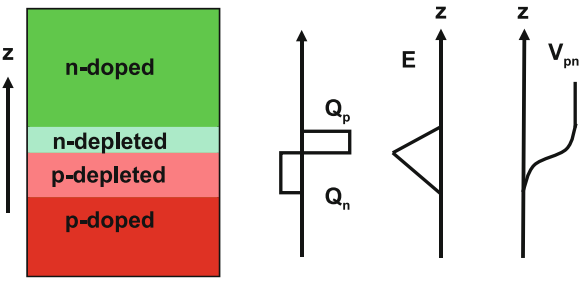


Fig. 2.28 Cross section of a pn-junction in equilibrium

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. 2.28). This potential is called the "built-in" voltage:

$$V_{\text{bi}} = \frac{kT}{q} \ln\left(\frac{N_A N_D}{n_i^2}\right) = 2\phi_F \tag{2.70}$$

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:

$$\begin{aligned} J_n &= q\mu_n nE + qD_n \frac{\mathrm{d}n}{\mathrm{d}z} = 0 \\ J_p &= q\mu_p pE - qD_p \frac{\mathrm{d}p}{\mathrm{d}z} = 0. \end{aligned} \tag{2.71}$$

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

$$D_n = \frac{kT}{q}\mu_n, \quad D_p = \frac{kT}{q}\mu_p. \tag{2.72}$$

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:

$$\begin{aligned} p_n &= \frac{n_i^2}{N_D} \mathrm{e}^{-qV_{\text{pn}}/kT} \\ n_p &= \frac{n_i^2}{N_A} \mathrm{e}^{qV_{\text{pn}}kT}. \end{aligned} \tag{2.73}$$

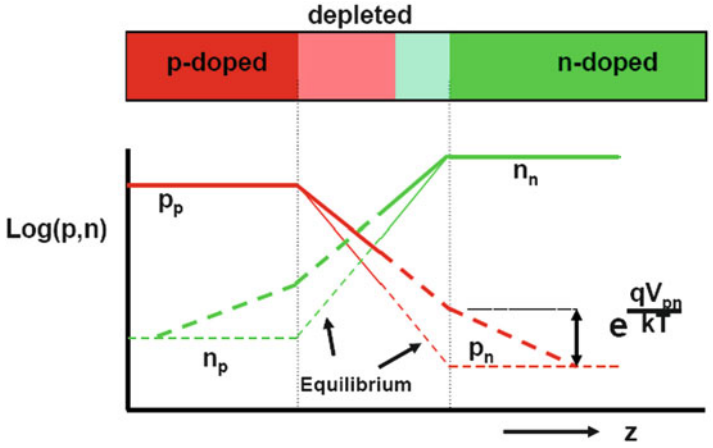


Fig. 2.29 The electron and hole concentrations in a pn-junction in equilibrium and during forward bias

The levels of the minority carriers at the edges of the depletion region are given by these equations (see Fig. 2.29).

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:

$$L_n = \sqrt{D_n \tau_n}. \tag{2.74}$$

τ_n is the generation-recombination time constant. In modern processes the diffusion length is in the order of 1 mm. In that case most recombination takes place 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:

$$\frac{\partial n(z,t)}{\partial t} = D_n \frac{\partial n(z,t)}{\partial z} \tag{2.75}$$

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

$$I = qA\frac{\mathrm{d}n}{\mathrm{d}t} = qAD_n\frac{\mathrm{d}n}{\mathrm{d}z}.$$

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

$$I_{\mathrm{pn}} = qA\left(D_n\frac{\mathrm{d}n}{\mathrm{d}z} + D_p\frac{\mathrm{d}p}{\mathrm{d}z}\right) = qA\left(D_n\frac{n_p}{L_n} + D_p\frac{p_n}{L_p}\right).$$

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

$$I_{\mathrm{pn}} = qA\left(D_n\frac{n_i^2}{N_A L_n} + D_p\frac{n_i^2}{N_D L_p}\right)(\mathrm{e}^{qV_{\mathrm{pn}}/kT} - 1). \tag{2.76}$$

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,

$$I_{\mathrm{pn}} = \frac{qAD_n n_i^2}{N_A L_n}(\mathrm{e}^{qV_{\mathrm{pn}}/kT} - 1) = I_s(\mathrm{e}^{qV_{\mathrm{pn}}/kT} - 1). \tag{2.77}$$

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

Equation 2.77 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:

$$\frac{1}{r_{\mathrm{pn}}} = \frac{\mathrm{d}I_{\mathrm{pn}}}{\mathrm{d}V_{\mathrm{pn}}} = \frac{qI_{\mathrm{pn}}}{kT}. \tag{2.78}$$

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

$$v_{\mathrm{noise,rms}} = \sqrt{4kT\mathrm{BW}r_{\mathrm{pn}}}.$$

Substituting Eq. 2.64 for n_i^2, the forward voltage at current I_{pn} is given:

$$V_{\mathrm{pn}} \approx \frac{E_G}{q} - \frac{kT}{q}\ln\frac{I_{\mathrm{pn}}N_a L_n}{qAD_n N_C N_V} = \frac{E_G}{q} - \frac{kT}{q}\ln\frac{I_{\mathrm{pn}}n_i^2}{I_s N_C N_V}. \tag{2.79}$$

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:

$$C_{\text{pn}} = \frac{\varepsilon_r \varepsilon_0}{z_d} = \sqrt{\frac{qN\varepsilon_r\varepsilon_0}{2(V_{\text{bi}} \pm |V_{\text{pn}}|)}}. \tag{2.80}$$

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

Many parasitic phenomena are associated with diodes [14]:

- 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 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 [14, Chap. 9].

2.4.5 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. 2.30). 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}:

$$I_C = h_{fe} \times I_B. \tag{2.81}$$

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

$$I_C = I_0 \left(\mathrm{e}^{\frac{qV_{\text{BE}}}{kT}} - 1 \right). \tag{2.82}$$

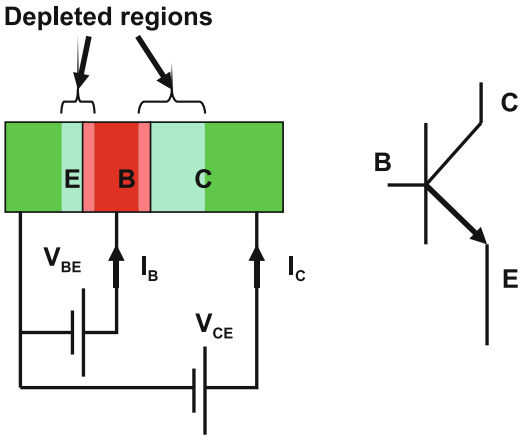


Fig. 2.30 The bipolar transistor: schematic cross section and the symbol for an npn transistor

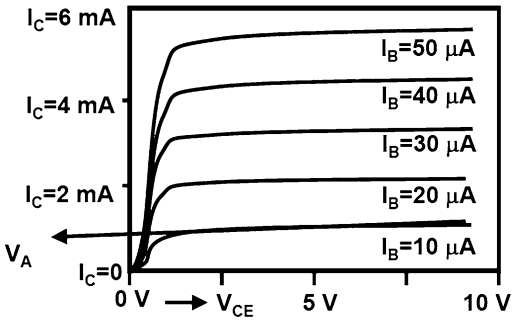


Fig. 2.31 The bipolar transistor: a typical set of characteristics

Consequently the transconductance is found:

$$g_m = \frac{\mathrm{d}I_C}{\mathrm{d}V_{\mathrm{BE}}} = \frac{qI_C}{kT}. \tag{2.83}$$

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 $I_C - V_{\mathrm{CE}}$ characteristics in Fig. 2.31, 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 $I_C - V_{\mathrm{CE}}$ 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 :

$$r_o = \frac{\mathrm{d}v_{ce}}{\mathrm{d}i_c} \approx \left| \frac{V_A}{I_C} \right|. \tag{2.84}$$

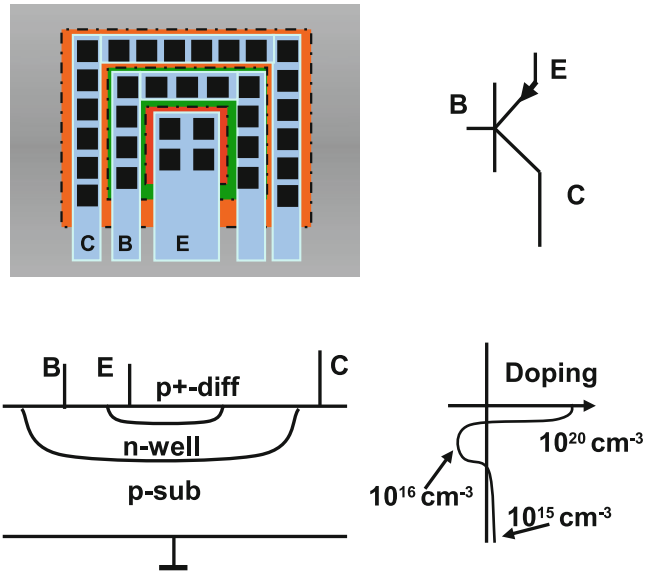


Fig. 2.32 The bipolar pnp transistor can be formed in a CMOS process from the *n*-well and the *p*-diffusion: lay-out, symbol, cross section, and doping profile

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. 2.32). 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 [14, Chap. 3], [22]. 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.

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:

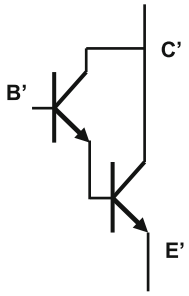
$$\sigma_{\Delta V_{\mathrm{be}}} = \frac{A_{V_{\mathrm{be}}}}{\sqrt{\text{area of emitter}}}. \tag{2.85}$$

For a high-speed bipolar process with thin base regions there is a relatively large impact of base dope fluctuations and $A_{V_{\mathrm{be}}} = 0.3\text{–}0.5\,\mathrm{mV\mu m}$. The base formed by an *n*-well of a parasitic vertical pnp in a CMOS process exhibits much better matching properties: $A_{V_{\mathrm{be}}} = 0.1\text{–}0.2\,\mathrm{mV\mu m}$. These numbers are much better than for thresholds of MOS devices. If, however, the relative current mismatch of bipolar

Table 2.17 Data for vertical pnp transistors

h_{fe}	5–10
V_A	150 V^{-1}
V_{be} at $J_E = 1\,\mu A/\mu m^2$	0.8 V
$\sigma_{V_{be}}$ emitter = 4 μm^2	0.1 mV

Fig. 2.33 The Darlington circuit



and MOS devices differs less, this device with parameters as in Table 2.17 is used in reference circuits (see Sect. 6.2.3).

2.4.6 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 [23] proposed in 1953 a combination of two transistors, being one of the first integrated circuits (Fig. 2.33). 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.

2.4.7 MOS Capacitance

Figure 2.34 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 (assumulation) 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.

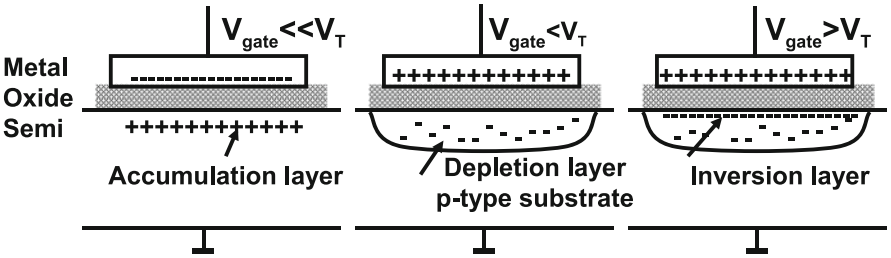


Fig. 2.34 The MOS capacitor: in accumulation, depletion, and inversion

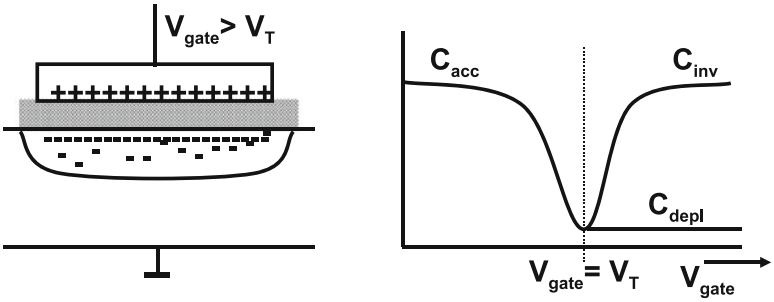


Fig. 2.35 MOS capacitance and C-V curve

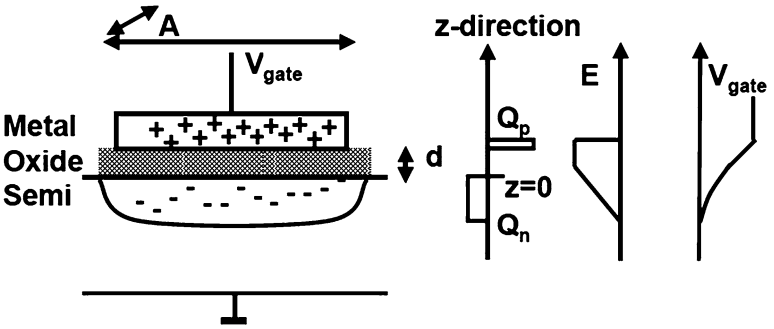


Fig. 2.36 The MOS capacitance consisting of a metal plate, an oxide insulating layer, and a depleted region in the semiconductor

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. 2.35). 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.

The structure in Fig. 2.36 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 2.37 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 buildup 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:

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

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:

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

$$Q_d = qNz_d = \sqrt{2qN\varepsilon_r \varepsilon_0 V_d} \tag{2.86}$$

where Q_d is the depletion charge per unit area. Applying the parallel-plate capacitance formula, the capacitance per unit area is

$$C_d = \frac{\varepsilon_r \varepsilon_0}{z_d} = \sqrt{\frac{qN\varepsilon_r \varepsilon_0}{2V_d}}. \tag{2.87}$$

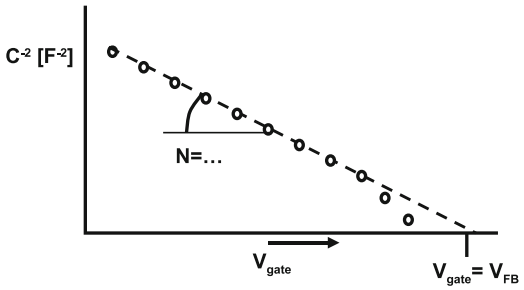


Fig. 2.37 The Mott-Schottky curve of a C-V measurement

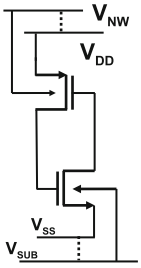


Fig. 2.38 Decoupling capacitors are constructed of gate capacitances. Often used in digital standard-cell blocks

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

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

The Mott–Schottky method in Fig. 2.37 rearranges the terms in the basic capacitance voltage relation to

$$C_d^{-2} = \frac{2(V - V_{\mathrm{FB}})}{qN\varepsilon_r\varepsilon_0}. \tag{2.88}$$

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.

Example 2.13. In Fig. 2.38 a standard schematic of a digital decoupling cell is shown. Explain the function.

Table 2.18 Diffusion capacitances in a 0.25 µm process and a 65 nm CMOS process

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

The numbers are indicative for a node and can vary per foundry [20]

Solution. In a digital standard-cell design the positive power supply is connected both to the source of the PMOS and the n-well (dashed line) for latch-up reasons. For the same reason the source of the NMOS is connected to the substrate. When the power supply is active, both the NMOS and the PMOS inversion layers are present and a decoupling capacitance of $(\text{Area}_{\text{NMOS}}+\text{Area}_{\text{PMOS}}) \times C_{\text{ox}}$ will appear between the V_{DD} and V_{SS} line. In case the power supply is very low or zero, the inversion layers will disappear, and the gates will form depletion capacitors with respectively the substrate and the n-well. The effective capacitance will be in the order of 10–15% of the capacitance with inversion.

Here the gate of the NMOS transistor is connected to V_{DD} via the inversion layer of the PMOS transistor. The series resistor will prevent very fast high-energy spikes on the power supply to damage the gate oxide.

2.4.8 Capacitance Between Layers

In a semiconductor process various capacitors are available through the construction of the layers in the process:

- Diffusion capacitances are based on a depleted semiconductor region (see Sect. 2.4.7 and Eq. 2.87). Semiconductor capacitances suffer from leakage and nonlinearities. These capacitors are asymmetrical. A reverse bias voltage creates a small leakage current, while a forward bias generates an exponentially growing forward current. Table 2.18 compares diffusion capacitors 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. 2.4.7) 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 nanometer processes the specific gate capacitance is around 12 fF/µm^2 (see Table 2.20 on p. 79). This device requires a significant turn-on voltage in excess of the threshold

voltage. Even at a significant turn-on voltage a slight nonlinearity between 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.

2.4.9 Voltage and Temperature Coefficient

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

$$C(V) = C(V_0)(1 + \mathrm{VC}(V - V_0)) \quad \text{with} \quad \mathrm{VC} = \frac{1}{C}\frac{\mathrm{d}C}{\mathrm{d}V}, \tag{2.89}$$

$$C(T) = C(T_0)(1 + \mathrm{TC}(T - T_0)) \quad \text{with} \quad \mathrm{TC} = \frac{1}{C}\frac{\mathrm{d}C}{\mathrm{d}T}. \tag{2.90}$$

The temperature coefficient of plate capacitors is mostly negligible. Capacitances of depleted regions may show significant temperature sensitivity. For large voltage variations the basic Eq. 2.56 must be used.

2.4.10 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 [24–26]. 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

$$\frac{\sigma_{\Delta C}}{C} = \frac{A_{C,\mathrm{area}}}{\sqrt{\mathrm{area}}} = \frac{A_{C,\mathrm{area}}\sqrt{\varepsilon_0\varepsilon_r/d_{\mathrm{ox}}}}{\sqrt{C}} = \frac{A_C}{\sqrt{C}}. \tag{2.91}$$

with area $= Cd_{\mathrm{ox}}/\varepsilon_0\varepsilon_r$. The area is expressed in $\mu\mathrm{m}^2$ and $A_{C,\mathrm{area}}$ in %µm. With $A_C = 0.5\%\sqrt{\mathrm{fF}}$ (see Table 2.19) a 400 fF capacitor has a mismatch of 0.025% or $\sigma_C = 0.1$ fF.

2.4.11 Capacitor Design

The capacitor is the most important passive device for analog electronic designs. A number of options are available to design a capacitor:

Table 2.19 An indication of capacitor characteristics in a semiconductor process ranging from 0.18 µm to 90 nm generations

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	$3–5 \times 10^{-2}$		
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	$5–10 \times 10^{-4}$	-8×10^{-5}	

sources, e.g., ITRS [20] and various publications

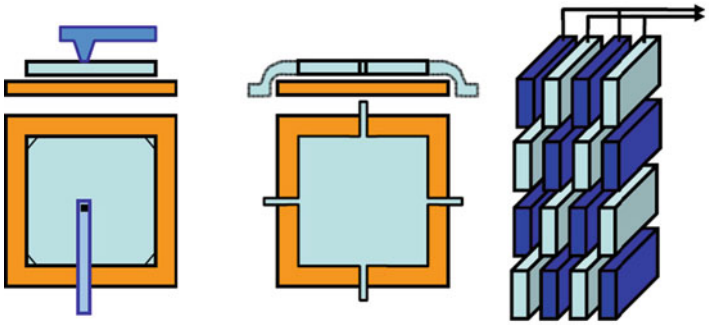


Fig. 2.39 Three capacitor implementations. The *left-side* configuration uses the fact that a top contact is allowed in some processes. The *middle* lay-out connects on the sides whenever top contacts are not permitted. The *right* fringe-capacitor configuration is used in high-metal density configurations

- Interconnects can form plate capacitors and fringe capacitors. The capacitance is defined by Eq. 2.54. 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 nm). 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, is an often applied solution to obtain a high-density capacitor. The combination of flat surfaces due to the chemical-mechanical polishing (CMP) process steps and 8–10 layers of metal lead to a capacitor with excellent properties. 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 and has a good linearity and low parasitics.
- In older processes a double polysilicon capacitor option is often available. In relatively low-doped polysilicon plates partial depletion can occur, which dramatically increases the effective insulator thickness. The depletion of the polysilicon leads to a voltage-dependent capacitance and nonlinearities. Figure 2.39 (left

and middle) shows a few different design styles. The first implementation is used in processes with two polysilicon 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 layer. The process first deposits the two polysilicon layers and after that patterns the second and the first layer. This procedure creates a high-quality capacitor but reduces the layout 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 layout 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."[15] Fringe capacitors use a finger structure and combine lateral and vertical capacitances (see Fig. 2.39). A capacity per unit area of around $1.5\,\mathrm{fF/\mu m^2}$ is obtainable. The low-voltage coefficient and excellent matching properties of fringe and plate capacitors make them often the highest precision components in a process.
- 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 $4\,\mathrm{fF/\mu m^2}$ up to $15\,\mathrm{fF/\mu m^2}$. It remains however questionable whether real gain is achieved over the before-mentioned plate and fringe capacitances.

The reproducibility of the capacitor ratio is of crucial importance in many analog applications and especially in switched-capacitor designs. Special care has to be taken to create capacitor layouts with optimum reproducibility properties. For large arrays of capacitors gradient reduction techniques similar to Fig. 7.17 are applied (see Table 2.19 for some data).

Example 2.14. A $10 \times 10\,\mu\mathrm{m}^2$ conductor is deposited on a 3,700 Å(=370 nm) silicon oxide on top of a conductive substrate. The permittivity in vacuum $\varepsilon_0 = 8.85 \times 10^{-14}\,\mathrm{F/cm}$. What is the resulting capacitor?

Solution. The capacitance is found from

[15]There are many fringe capacitor layouts. For lawyers patent 7.170.178 [27] is a good starting point; engineers might prefer the "examples" section of [28].

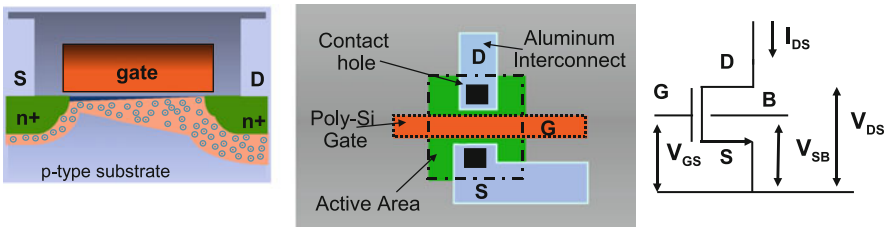


Fig. 2.40 Cross section, layout, and symbol of a MOS transistor

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

This calculation comprises the vertical field under the conductor. In a real experiment also the field at the perimeter must be taken into account.

2.5 MOS Transistor

The metal-oxide-silicon (MOS) transistor in Fig. 2.40 is a charge-controlled device.[16] In the MOS capacitor it was assumed that gate depletes the underlying semiconductor *p*-type substrate from free holes. However, if a source of electrons is in the vicinity of this gate, electrons will flow under the isolating interface and form the inversion layer. This inversion charge under the interface can form an electrical connection between the source and second drain terminal defining 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 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:

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

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

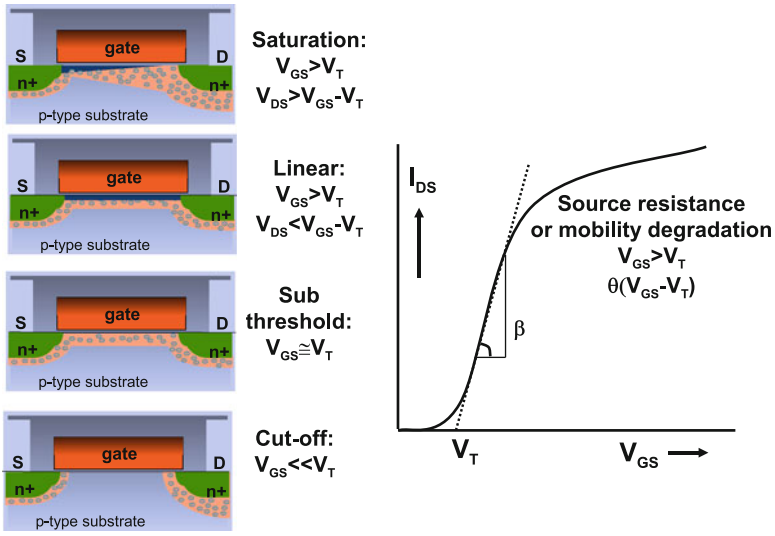


Fig. 2.41 *Left*: various regimes in a NMOS transistor. *Right*: in a I_{DS} versus V_{GS} plot in the linear regime the threshold voltage and current factor can be easily identified

For a 0.18 µm process $d_{ox} = 4\,\text{nm} = 4 \times 10^{-7}\text{cm}$ and $C_{ox} = 8\text{fF}/\mu\text{m}^2$. 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 causes 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. When the electric field of the gate attracts the electrons towards the insulator, the inversion layer or "channel" creates a connection between the two terminals of the MOS transistor. The start of inversion is (arbitrarily) defined as the concentration of the charge in the channel that exceeds the substrate doping level. From an external perspective the gate voltage of an NMOS transistor V_{GS} exceeds a threshold voltage V_T. First the channel will be in "weak inversion." If the excess voltage is more than a few kT/q the "strong-inversion regime" applies. The simplest condition for the definition of (strong) inversion is therefore

$$V_{GS} > V_T. \tag{2.92}$$

Nearly all gate voltage variation is reflected by the inversion charge variation. Just a little change in the surface potential of the inversion charge ϕ_s is needed to increase the charge concentration as demanded by the charge Eq. 2.66. Without sufficient gate voltage, there is no inversion charge, and the structure behaves as a MOS capacitance in depletion. The transistor is in the off state (Fig. 2.41).

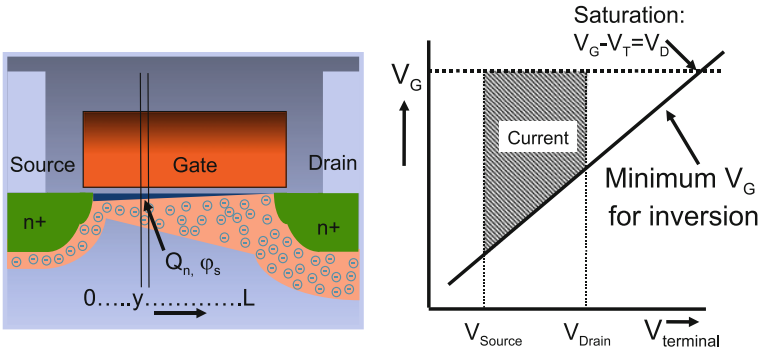


Fig. 2.42 The gradual channel approximation in a transistor allows to analyze the charge balance in the vertical direction (*left*). The Memelink model (*right*) gives a graphical representation of the MOS conductivity. Here all voltages are referred to the substrate potential. The MOS current is proportional to the *shaded area*

The inversion charge layer is characterized at every position y under the gate by a surface potential ϕ_s and an inversion charge density (see Fig. 2.42 (left)): $Q(y) = C_{\mathrm{ox}}(V_G - V_{\mathrm{T}} - \phi_s(y))$. The voltages and charge balance in the vertical direction is considered dominant over the variation in y direction. This is known as the "gradual channel" approximation. From the inversion charge density the conductivity at every position y along the channel is

$$\text{conductivity at position } y = W \times \mu_n \times C_{\mathrm{ox}}(V_G - V_{\mathrm{T}} - \phi_s) = \frac{\mathrm{d}I}{\mathrm{d}\phi_s}.$$

Now the total current can be calculated from the average conductance under the entire gate area:

$$\begin{aligned}\int_{y=0}^{y=L} I(y)\mathrm{d}y = I_{\mathrm{DS}}L &= \int_{y=0}^{y=L} W\mu_n C_{\mathrm{ox}}(V_G - V_{\mathrm{T}} - \phi_s(y))\mathrm{d}y \\ &= \int_{\phi_s=V_S}^{\phi_s=V_D} W\mu_n C_{\mathrm{ox}}(V_G - V_{\mathrm{T}} - \phi_s)\mathrm{d}\phi_s \\ &= W\mu_n C_{\mathrm{ox}}\left((V_G - V_{\mathrm{T}})\phi_s - \frac{1}{2}\phi_s^2\right)\Bigg|_{\phi_s=V_S}^{\phi_s=V_D}. \end{aligned} \tag{2.93}$$

It is implicitly assumed that there is no current flowing in any other direction than from source to drain in the y direction. The current is constant for any cross section at any position y, which means that with reducing inversion charge density the electrons must move at a greater velocity[17]:

[17] There is of course a limitation due to the maximum velocity of carriers.

$$I_{\mathrm{DS}} = \frac{WC_{\mathrm{ox}}\mu}{L}\left(V_{\mathrm{GS}} - V_{\mathrm{T}} - \frac{1}{2}V_{\mathrm{DS}}\right)V_{\mathrm{DS}} = \frac{W\beta_{\square,n}}{L}\left(V_{\mathrm{GS}} - V_{\mathrm{T}} - \frac{1}{2}V_{\mathrm{DS}}\right)V_{\mathrm{DS}} \quad (2.94)$$

In the linear regime the MOS transistor behaves as a gate-voltage-controlled resistor.

In Fig. 2.42 (right) these relations between gate source and drain voltages are visualized. The Memelink model [29] is based on Eq. 2.93 and refers all voltages to the substrate. The horizontal axis corresponds to the voltage of a diffusion next to the gate. The slanted line indicates the minimum gate voltage that is needed to start inversion. For a given gate voltage (on the vertical axis) the distance between the actual gate voltage and the slanted line indicates the local inversion charge density. If the area is multiplied by $C_{\mathrm{ox}}\mu_n W/L$ the total current is found.

In saturation mode the drain voltage rises to a level where at the drain side no longer inversion exists: $V_G - V_D < V_{\mathrm{T}}$. 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 ϕ_s reaches a level where the field reverses. The electrons travel along the substrate-insulator 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

$$(V_{\mathrm{GS}} - V_{\mathrm{DS}}) < V_{\mathrm{T}}, \quad V_{\mathrm{DS}} > (V_{\mathrm{GS}} - V_{\mathrm{T}}), \quad \phi_s(\text{saturation point}) = V_G - V_{\mathrm{T}}. \quad (2.95)$$

Drain–source voltage above this "saturation voltage" will not influence the conductivity in the inversion layer as in the linear regime. This allows the transistor to operate efficiently as a current source. Below this level the transistor operates in the linear regime, and a considerable parallel resistor impairs current source performance. The current is in saturation mode found from Eq. 2.93 by replacing the upper ϕ_s boundary condition by the saturation voltage: $V_{\mathrm{GS}} - V_{\mathrm{T}}$:

$$I_{\mathrm{DS}} = \frac{WC_{\mathrm{ox}}\mu}{2L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2 = \frac{W\beta_{\square,n}}{2L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2. \quad (2.96)$$

The result is the classical "square-law MOS equation": The quantity $WC_{\mathrm{ox}}\mu/L$ 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 2.43 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 $\mu_n \approx 1500\,\mathrm{cm^2/Vs}$ and $\mu_p \approx 450\,\mathrm{cm^2/Vs}$ for the mobility for electrons and holes, the current factors will differ accordingly. Due to many

[18] Some authors prefer K.

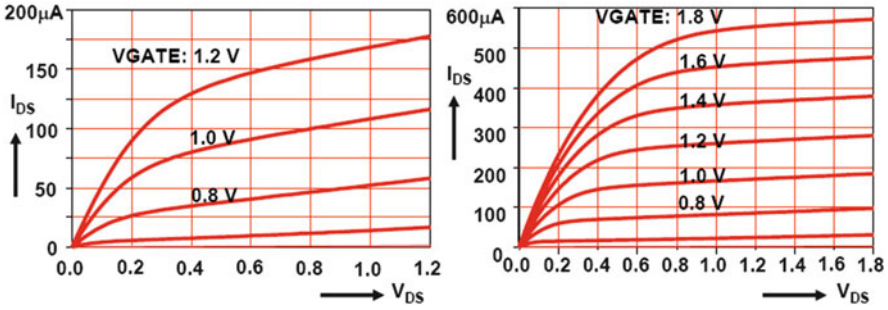


Fig. 2.43 Two drain current I_{DS} versus drain–source voltage V_{DS} plots. The *left* set of curves is a 0.3/0.065 NMOS transistor in a 65-nm process. The *right* set of curves is for a 1/0.18 NMOS transistor in a 0.18 µm process. The gate voltage runs from 1.2 resp. 1.8 V with decrements of 0.2 V

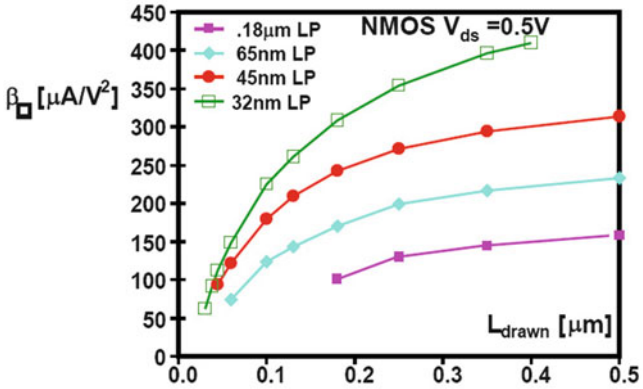


Fig. 2.44 The current factor in a MOS transistor for different processes is much lower than the nominal current factor. A strong dependence on the length can be observed, $V_{GS} - V_T \approx 0.2$–0.3 V [30]

impairments of the current flow under the gate, the real values for the mobilities are normally much lower (see Fig. 2.44). 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.

2.5.1 Threshold Voltage

The threshold voltage is composed of some fixed interface potential contributions and of the voltage that is necessary to build up the depletion charge (Eq. 2.86):

Table 2.20 The values of NMOS and PMOS parameters are indicative for what is used in industry

Process	V_{DD}	d_{ox}	C_{ox}	$V_{T,n}$	$V_{T,p}$	$\beta_{\square,n}$	$\beta_{\square,p}$	$A_{V_{T,n}}$	$A_{V_{T,p}}$	$A_{\beta,np}$
Unit	Volt	nm	fF/μm^2	Volt	Volt	μA/V^2	μA/V^2	mVμm	mVμm	%μm
0.8 μm	3.3	15	2.3	0.6	−0.65	125	55	10.7	18.0	4
0.6 μm	3.3	13	2.7	0.65	−0.8	150	50	11.0	8.5	
0.5 μm	3.3	12	2.9	0.6	−0.6	130	36	9	10	1.8
0.35 μm	3.3	7.7	4.3	0.63	−0.6	190	46	8	7.4	2
0.25 μm	2.5	6	6.9	0.57	−0.53	235	53	6	6	1.5
0.18 μm	1.8	4	8.3	0.48	−0.5	300	80	6	5	1.6
0.13 μm	1.2	2.5	11	0.34	−0.36	590	135	5	5	1.6
90 nm (LP)	1.2	2.3	11.7	0.37	−0.39	550	160	4.5	3.5	
65 nm (LP)	1.2	2.2	12.6	0.32	−0.36	450	200	4.5	3.5	1.2
45 nm (LP)	1.1	1.7	16	0.39	−0.42	300	100	4	4	1.2
28 nm	1.0	1.0		0.35	−0.35			2	2	

Especially the 130 to 28 nm values depend on many technological effects and bias settings. The 28-nm data is based on process with a high-k factor dielectric with metal gate. See for a discussion of the current factor in these processes Fig. 2.44 (source: ITRS [20] and various publications)

$$\begin{aligned} V_T &= V_{FB} + 2\phi_F + \frac{\sqrt{2\varepsilon_r\varepsilon_0 q N_a (2\phi_F + V_{SB})}}{C_{ox}} \\ &= V_{FB} + 2\phi_F + \gamma\sqrt{2\phi_F + V_{SB}} \\ &= V_T(V_{SB} = 0) + \gamma(\sqrt{2\phi_F + V_{SB}} - \sqrt{2\phi_F}), \end{aligned} \tag{2.97}$$

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:

$$\phi_F = \frac{kT}{q} \ln\left(\frac{N_a}{n_i}\right). \tag{2.98}$$

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:

$$\gamma = \frac{\sqrt{2\varepsilon_r\varepsilon_0 q N_a}}{C_{ox}}. \tag{2.99}$$

Example 2.15. Calculate the current for a $W/L = 5/0.2\,\mu$m NMOS transistor in 0.18 μm CMOS, with $V_{GS} = 1.0$ and $V_{DS} = 1.8$ V.

Solution. First the operation regime is determined as $V_{DS} > V_{GS} - V_T$ the transistor operates in the saturation regime and Eq. 2.96 applies. With $V_T = 0.48$ V and $\beta_{\square,n} = 300\,\mu$A/V^2,

$$I_{DS} = \frac{W\beta_{\square,n}}{2L}(V_{GS} - V_T)^2 = \frac{5 \times 300}{2 \times 0.2}(1 - 0.48)^2 = 1{,}014\,\mu A. \tag{2.100}$$

Compare to Fig. 2.43 (right) and check the result against the $W/L = 1/0.18$ device.

2.5.2 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 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 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

$$I_{DS} = \frac{WI_o}{L} e^{q(V_{GS} - V_T)/mkT} \tag{2.101}$$

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. 2.45 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. 2.45 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× and is used for memories. In digital design

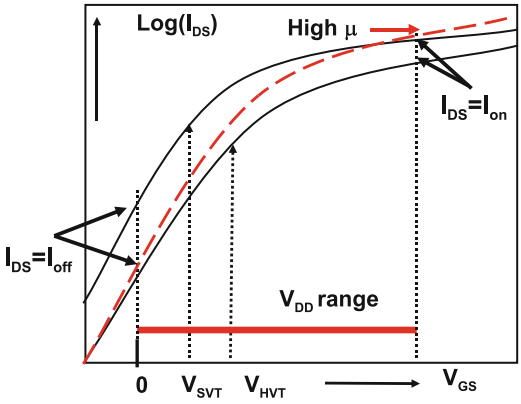


Fig. 2.45 The current in a MOS transistor is determined by weak inversion for gate-source voltages lower than the threshold voltage, while for higher gate-source voltages the strong-inversion regime applies. In digital CMOS processes multiple threshold voltages are available. Here V_{SV_T} and V_{HV_T} 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

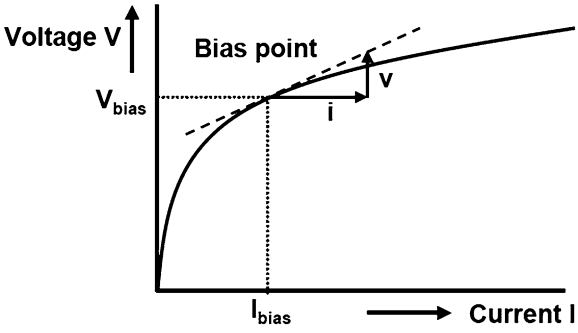


Fig. 2.46 Small-signal linearization in a bias point

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%.

2.5.3 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. 2.46 the relations between voltages and currents are not linear over the entire signal range and bias. In analog design small signals are used

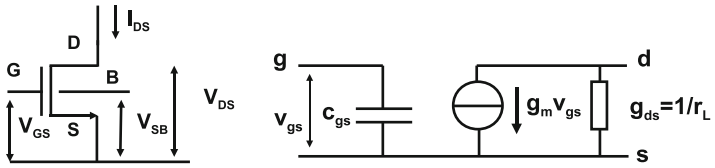


Fig. 2.47 The MOS transistor as a small-signal equivalent circuit

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. 2.47). In small-signal mode all voltages and currents are written in lowercase 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 analyses it is necessary to include the modulation due to the substrate voltage as a controlled current source parallel to $g_m v_{gs}$:

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

Typically the bulk transconductance amounts 10%–20% of the gate transconductance.

2.5.4 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 Eq. 2.96 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 (Table 2.21). This drain-induced-barrier-lowering (DIBL) is the major cause for an increased drain current in a submicron transistor. In older processes an increased drain voltage pushes back the point where inversion still exists 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

Table 2.21 Transistor parameters for 0.18 µm CMOS as specified by ITRS [20] 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 modulation λ	0.05 at $L_{minimum}$	0.05 at $L_{minimum}$
Output modulation 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$^{0.5}$

described by reducing the effective gate length:

$$I_{DS} = \frac{WC_{ox}\mu}{2L(1-\alpha_L V_{DS})}(V_{GS}-V_T)^2 \approx \frac{WC_{ox}\mu}{2L}(V_{GS}-V_T)^2(1+\alpha_L V_{DS}). \quad (2.104)$$

Here α_L 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:

$$I_{DS} = \frac{WC_{ox}\mu}{2L}(V_{GS}-V_T+\lambda V_{DS})^2. \quad (2.105)$$

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 introducing too much series resistance.

2.5.5 *Output Impedance*

In the first model for MOS currents Eq. 2.104, the effect of a drain-voltage variation was described with parameter α_L. The small-signal output impedance is found as

$$g_{ds} = \frac{1}{r_{ds}} = \frac{dI_{DS}}{dV_{DS}} = \frac{\alpha_L}{1-\alpha_L V_{DS}} I_{DS} \approx \alpha_L I_{DS} \quad (2.106)$$

or the small-signal output impedance is $r_{ds} = 1/\alpha_L I_{DS}$.

The output conductance is physically caused by drain-induced barrier lowering and static feedback (Eq. 2.105). 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:

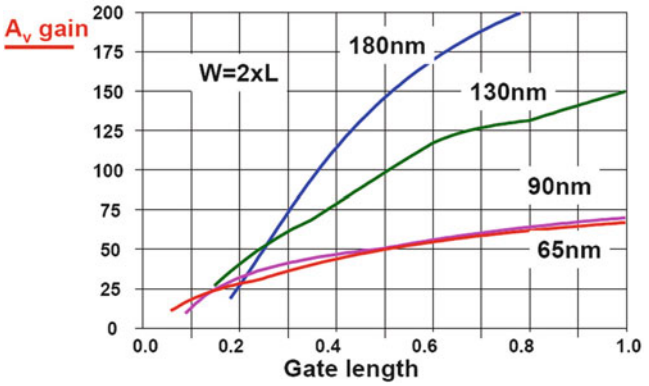


Fig. 2.48 The voltage gain of a transistor with a fixed ratio $W/L = 2$ at constant current as a function of gate length and technology

$$g_{ds} = \frac{1}{r_{ds}} = \frac{dI_{\mathrm{DS}}}{dV_{\mathrm{DS}}} = \frac{\lambda W C_{\mathrm{ox}} \mu}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}} + \lambda V_{\mathrm{DS}}) = \lambda g_m. \tag{2.107}$$

With this definition, the maximum voltage amplification A_V of a transistor is limited by the static feedback:

$$A_V = g_m r_{ds} = g_m / g_{ds} = \frac{1}{\lambda}. \tag{2.108}$$

Rephrased in terms of currents and voltages: the current variation due to an input signal v_{in} is $i_{ds} = g_m v_{\mathrm{in}}$. 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

$$A_V = \frac{v_{\mathrm{out}}}{v_{\mathrm{in}}} = \frac{i_{ds} r_{ds}}{v_{\mathrm{in}}} = \frac{1}{\lambda}. \tag{2.109}$$

Practically A_V is limited to the range 20–50 (see Fig. 2.48).

2.5.6 *Matching*

Circuit design and especially analog-to-digital design heavily rely 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 differs 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:

$$\sigma_{\Delta V_\mathrm{T}} = \frac{A_{V_\mathrm{T}}}{\sqrt{WL}}. \tag{2.110}$$

All technological constants are merged into one parameter A_{V_T} [31].

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

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

$$\frac{\sigma_{\Delta\beta}}{\beta} = \frac{A_\beta}{\sqrt{WL}}. \tag{2.111}$$

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 Eq. 2.18,

$$\left(\frac{\sigma_{\Delta I}}{I}\right)^2 = \left(\frac{\mathrm{d}I}{\mathrm{d}V_\mathrm{T}}\right)^2 \sigma^2_{\Delta V_\mathrm{T}} + \left(\frac{\mathrm{d}I}{\mathrm{d}\beta}\right)^2 \sigma^2_{\Delta\beta}. \tag{2.112}$$

For strong inversion this equation can be written as

$$\left(\frac{\sigma_{\Delta I}}{I}\right)^2 = \left(\frac{2\sigma_{\Delta V_\mathrm{T}}}{V_\mathrm{GS} - V_\mathrm{T}}\right)^2 + \left(\frac{\sigma_{\Delta\beta}}{\beta}\right)^2. \tag{2.113}$$

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:

$$\left(\frac{\sigma_{\Delta I}}{I}\right)^2 = \left(\frac{q\sigma_{\Delta V_\mathrm{T}}}{mkT}\right)^2. \tag{2.114}$$

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

Example 2.16. A pair of NMOS transistors is used for a differential input stage. The size of each transistor is 20 µm/5 µm. The mismatch coefficient is $A_{V_\mathrm{T}} = 10\,\mathrm{mV\mu m}$. What is the 3-$\sigma$ spread between the input gates of this stage. If the transistors are doubled, what is the input-referred mismatch?

Solution. The coefficient A_{V_T} is defined for the difference of the threshold voltage for two equal transistors. So,

$$\sigma_{\Delta V_\mathrm{T}} = \sigma_{V_\mathrm{T1} - V_\mathrm{T2}} = \sqrt{\sigma^2_{V_\mathrm{T1}} + \sigma^2_{V_\mathrm{T1}}} = \frac{A_{V_\mathrm{T}}}{\sqrt{W \times L}}.$$

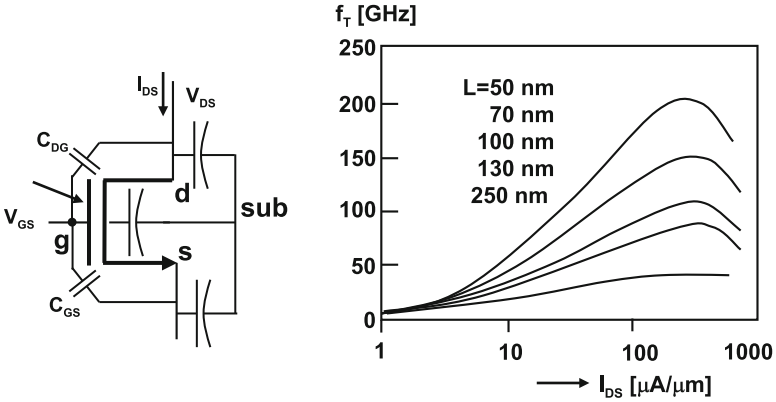


Fig. 2.49 The capacitances around a MOS device in 90 nm technology determine the high-frequency behavior. A characteristic set of performance curves shows a strong dependence for the maximum frequency on gate length and current (after [32]). With technological tricks (under-diffusion) the effective gate length can be made smaller than the lithographical length

So the 1 σ value is 1 mV and $3\sigma = 3$ mV. In case both transistors have equal size it is clear that

$$\sigma_{V_{T1}} = \sigma_{V_{T2}} = \sigma_{\Delta V_T}/\sqrt{2} = \frac{A_{V_T}}{\sqrt{2W \times L}} = 0.7\,\text{mV}.$$

If the transistor width is doubled by increasing the width to $W = 40\,\mu$m, the 1 σ value drops to 0.7 mV. If each transistor of 20 µm/5 µm is replaced by two transistors of the same size the same result is obtained.

2.5.7 High-Frequency Behavior

Figure 2.49 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 [33]. The transit time limits the highest frequencies obtainable with MOS devices to some 1,000 GHz [34].

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 2.49 shows the most relevant capacitors in a MOS device and Table 2.22 shows some values. The detailed behavior of these capacitors is subject to a lot of study [22]. The most common measure of high-frequency behavior is the cut-off frequency:

Table 2.22 Some parasitic transistor capacitors in 90 nm technology

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

source: ITRS [20] and various publications

$$f_T = \frac{g_m}{C_{\text{gate}} + C_{gs} + C_{gd} + C_{\text{parasitcs}}}. \tag{2.115}$$

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 $g_m = W\beta(V_{\text{GS}} - V_{\text{T}})/L$ and only looking at the largest capacitor

$$f_T \approx \frac{\mu_{n,p}(V_{\text{GS}} - V_{\text{T}})}{L^2} \tag{2.116}$$

it is clear that especially a short effective gate length increases the cutoff frequency. In Fig. 2.49 (right) a typical set of curves shows the behavior of the cutoff 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 cutoff frequency a number of other high-frequency indicators are used. The f_{MAX} frequency [32] 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 layout 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.

2.5.8 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 road map expects a gate leakage current density of 10^{-3} A/cm^2, rising at 2 nm to 10^{-1} A/cm^2 and at 1.5 nm to 10 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 corre-

sponding current density is a function of the electric field over the isolator:

$$J_{\mathrm{FN}} = C_1 E_{\mathrm{ox}}^2 \mathrm{e}^{\frac{-C_2}{E_{\mathrm{ox}}}}. \tag{2.117}$$

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.

2.5.9 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 Eq. 2.97. The contributions to the threshold temperature dependence come from the work function ϕ_{MS} and from the thermal potential ϕ_F [35]. Both are related to the Boltzmann equation which describes the carrier concentration as a function of potential difference and temperature. For *n*-channel,

$$\frac{\mathrm{d}V_{\mathrm{T}}}{\mathrm{d}T} = \frac{\mathrm{d}\phi_{\mathrm{MS}}}{\mathrm{d}T} + \frac{2\mathrm{d}\phi_F}{\mathrm{d}T} + \frac{\gamma}{\sqrt{(2\phi_F + V_{\mathrm{SB}})}} \frac{\mathrm{d}\phi_F}{\mathrm{d}T}. \tag{2.118}$$

The temperature dependence of the work function is linked to the doping polarity of the gate and substrate. Equal polarities give 1 mV/K and opposite give −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 Eq. 2.64:

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

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]

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

Evaluating the last formula with $E_G/q = 1.205\,\mathrm{V}$ and $\phi_F \approx 0.35\,\mathrm{V}$ gives a temperature coefficient of the band bending voltage of $-0.9\,\mathrm{mV/°C}$, which brings the

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

contribution of the band-bending in the threshold voltage to a value of $-2\ldots-2.4\,\mathrm{mV/^\circ C}$.

The temperature dependence of the work function depends on the choice of the gate and substrate materials [35]. 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

$$\frac{\mathrm{d}V_\mathrm{T}}{\mathrm{d}T} = \pm 2\ \mathrm{mV/^\circ C}, \tag{2.119}$$

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 in 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:

$$\beta = \left(\frac{T}{T_0}\right)^\alpha \beta(T = T_0). \tag{2.120}$$

At higher temperatures the current factor decreases.

The overall current depends on temperature as

$$\frac{\mathrm{d}I_D}{\mathrm{d}T} = \frac{\mathrm{d}I_D}{\mathrm{d}\beta}\frac{\mathrm{d}\beta}{\mathrm{d}T} + \frac{\mathrm{d}I_D}{\mathrm{d}V_\mathrm{T}}\frac{\mathrm{d}V_\mathrm{T}}{\mathrm{d}T} = \frac{\alpha}{T}I_D - \frac{2I_D}{V_\mathrm{GS}-V_\mathrm{T}}\frac{\mathrm{d}V_\mathrm{T}}{\mathrm{d}T}, \tag{2.121}$$

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 $-2\,\mathrm{mV/^\circ C}$, a temperature cancellation occurs at $V_\mathrm{GS} - V_\mathrm{T} = 0.6\,\mathrm{V}$. In a digital CMOS circuit with a threshold voltage of around $V_\mathrm{T} = 0.4\,\mathrm{V}$, a supply voltage of 1 V will lead to a temperature-insensitive circuit.

2.5.10 Noise

Figure 2.50 shows the dominant noise contributions in a MOS transistor (see [17, 36]). 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

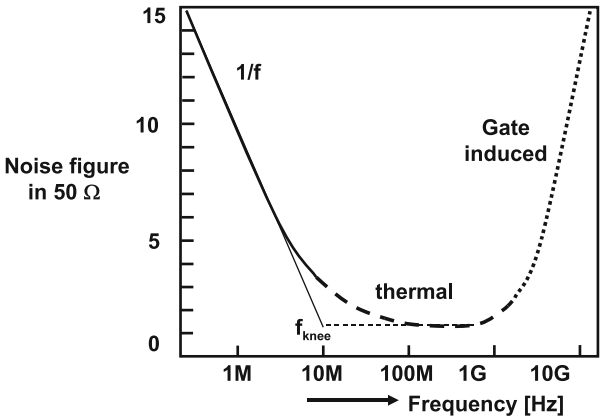


Fig. 2.50 Various noise regimes in a MOS transistor

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

$$S_v(f) = \sqrt{\frac{K_{1/f}}{WLf}} \quad [\mathrm{V}\sqrt{\mathrm{Hz}}], \tag{2.122}$$

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:

$$\begin{aligned} &\text{linear regime} && S_{vv} = 4kTR_{\mathrm{channel}} && S_{ii} = 4kT/R_{\mathrm{channel}} \\ &\text{saturated regime} && S_{vv} = 4\alpha kT/g_m && S_{ii} = 4\alpha kT g_m. \end{aligned}$$

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 $\alpha = 1$ [36].

While the knee frequency for 0.5 µm CMOS processes was in the 100 kHz range, this crossover 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

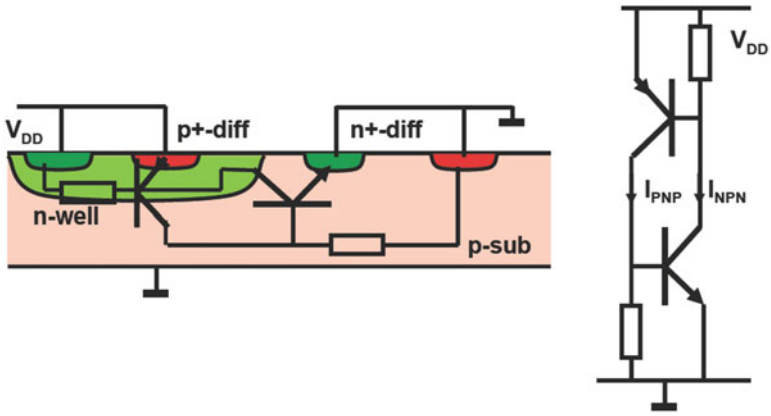


Fig. 2.51 Cross section of a basic latch-up situation in CMOS and a circuit representation

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.

2.5.11 Latch-Up

Figure 2.51 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

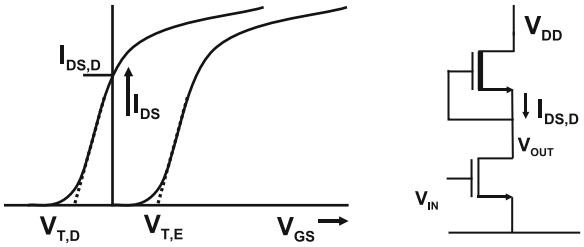


Fig. 2.52 The threshold voltage of an NMOS transistor can be positive or negative. A negative threshold voltage leads to a current when source and gate are connected together. The *right-hand plot* shows a characteristic inverter arrangement

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. 6.2.3 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.

2.5.12 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. 2.52). 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 [14, 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.

2.5.13 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.

In the modeling world 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), are 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 Memelink-Jespers-Wallinga graphical model has been devised in the early 1970s for educational purposes. This model (see Fig. 2.42) refers all voltages to the substrate. It is to a certain extent a predecessor of the EKV model. More features of this model can be found in [29].
- 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 nonexistent asymmetry with the drain. This model originally aimed at digital circuit design and was gradually extended with all kinds of analog-relevant phenomena. The result contains many fitting and numerical parameters. The 300-item parameter list is only fully understood by a handful of specialists. A starting point for reading is offered in [37]. The BSIM model is mainly in use in US companies.
- The PSP model was created in Philips Research. The description of this model aims at a "compact model." Only a limited number of physics-based parameters describe the voltage-charge-current relations. 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, and output conductance [38].

- 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 in applications where a MOS serves as a resistor the model is very accurate. Also the subthreshold behavior is well described. The model is popular in academia as it allows a good insight on the underlying physics (see [39]).
- Next to the above models there are many alternatives. Sakurai describes a model where the quadratic term is replaced by a parameter α [40]. This is a relatively simple approach to a sufficiently accurate model for digital applications. Redman-White and his research group described silicon-on-insulator transistors and the design consequences [41].

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 more 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 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:

$$E = \int_{t=-\infty}^{\infty} V(t) \times I(t) \mathrm{d}t. \tag{2.123}$$

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:

$$P = \lim_{T\to\infty} \frac{1}{T} \int_{t=-T/2}^{T/2} V(t) \times I(t) \mathrm{d}t. \tag{2.124}$$

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 (rms) value:

$$\text{Root-mean-square value:} \quad V_{\text{rms}} = \sqrt{\frac{1}{T} \int_{t=0}^{T} V^2(t) \mathrm{d}t}. \tag{2.125}$$

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

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

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 root-mean-square 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 root-mean-square voltage over the resistor is defined as

$$V_{\text{load,rms}} = \sqrt{e_1^2 + e_2^2}. \tag{2.126}$$

This "root-mean-square" sum is applicable to uncorrelated 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. 2.53 a voltage source is shown with an internal impedance R_{in}. As voltages and potentials exist between two terminals, the ground sign indicates the reference

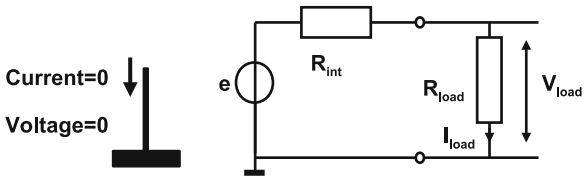


Fig. 2.53 The voltage source has an internal resistance and is loaded with a impedance R_{load}

Table 2.23 Various modes of transfer

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

terminal in the figure. From a fundamental point of view, there is in every circuit exactly one ground symbol that does not conduct any current.[20]

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. 2.53 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 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."

[20]To avoid too many wires, in some figures, multiple ground signs are used. Sticking strictly to a single ground sign is helpful when analyzing various forms of interference.

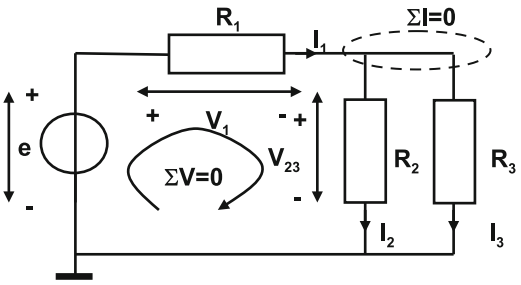


Fig. 2.54 Kirchhoff's law for voltage and current

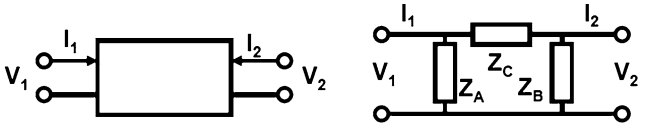


Fig. 2.55 A two-port circuit

2.6.2 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:

$$\Sigma I = 0, \tag{2.127}$$

where the direction of the current (in or out of the summation node) determines the sign. In the example of Fig. 2.54, $+I_1 - I_2 - I_3 = 0$.

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

$$\Sigma V = 0, \tag{2.128}$$

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. 2.54, $-e + V_1 + V_{23} = 0$.

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.

2.6.3 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 Fig. 2.55 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[21]:

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

where each of the z-parameters is defined as

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

$$z_{21} = \left.\frac{v_2}{i_1}\right|_{(i_2=0)} \qquad z_{22} = \left.\frac{v_2}{i_2}\right|_{(i_1=0)}. \tag{2.129}$$

As an example the z-parameters of Fig. 2.55 (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:

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

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

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

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

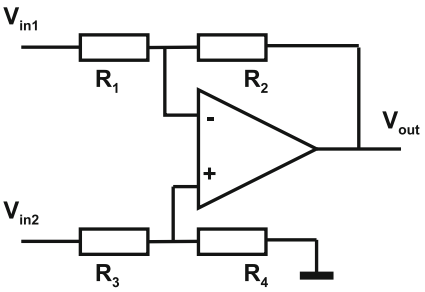


Fig. 2.56 A general operational amplifier configuration

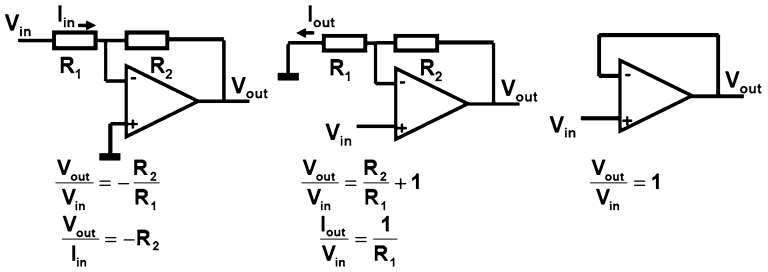


Fig. 2.57 Three classical opamp configurations: inverting, non-inverting, and unity gain

2.6.4 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 2.56 shows a general operational amplifier configuration. Under the above conditions, the analysis gives

$$V_{\text{out}} = \frac{R_4}{R_1}\frac{R_1 + R_2}{R_3 + R_4} V_{\text{in2}} - \frac{R_2}{R_1} V_{\text{in1}}. \tag{2.130}$$

Specific topologies are found by setting resistances to zero or infinite. In Fig. 2.57 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 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

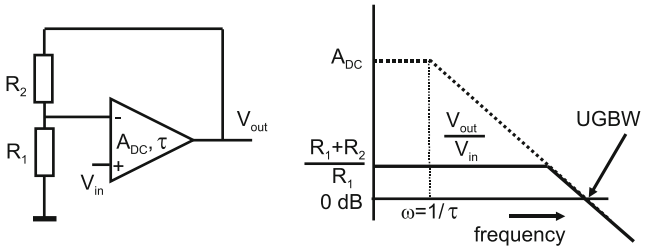


Fig. 2.58 An operational amplifier with limited gain and bandwidth. Open-loop transfer (*dotted line*) and closed-loop transfer (*bold line*)

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

$$\frac{V_{\mathrm{out}}}{V_{\mathrm{in}}} = \frac{R_1+R_2}{R_1}\,\frac{1}{1+\frac{R_1+R_2}{A_{\mathrm{DC}}R_1}+j\omega\tau\frac{R_1+R_2}{A_{\mathrm{DC}}R_1}} \approx \frac{R_1+R_2}{R_1}\,\frac{1}{1+j\omega\tau\frac{R_1+R_2}{A_{\mathrm{DC}}R_1}}. \tag{2.131}$$

The overall transfer equals the inverse of the feedback factor. The brute force amplification of the amplifier is tailored to the needs of the users 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 $\omega_{\mathrm{UGBW}} = A_{\mathrm{DC}}/\tau$ 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 many designs there are more sources that add (unwanted) signals into the amplification chain. In Fig. 2.59 a distortion voltage is added to the output of the opamp. The transfer of this voltage to the output is calculated in a similar manner as in Eq. 2.131:

$$\frac{V_{\mathrm{out}}}{V_{\mathrm{dis}}} = \frac{1+j\omega\tau}{1+\frac{A_{\mathrm{DC}}R_1}{R_1+R_2}+j\omega\tau}. \tag{2.132}$$

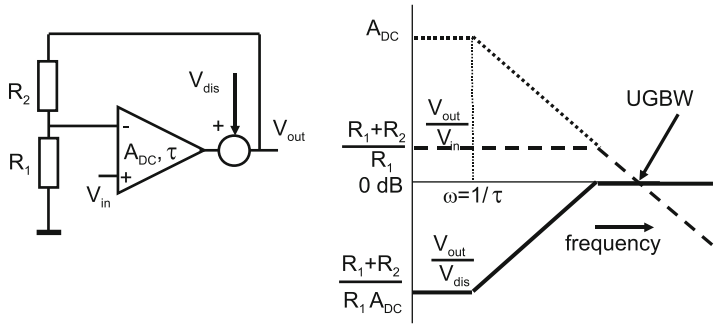


Fig. 2.59 The closed-loop transfer curves of operational amplifier with feed back path (*dotted line*) and an additional distortion source (*bold line*)

This distortion voltage is suppressed by the product of opamp gain and feedback factor. As the opamp gain decreases, the suppression is less, and at very high frequencies the opamp is of no influence, and the distortion voltage is fully present in the output.

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:

$$i_{\text{out}} = g_m v_{\text{in}}. \tag{2.133}$$

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_mC filters, OTAs replace opamps.

2.6.5 *Differential Design*

In CMOS technology signals are mostly represented by voltages. In Fig. 2.60 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. 2.60 shows two amplifiers each processing a single-ended signal, however, in their combination processing the difference between both signal lines.

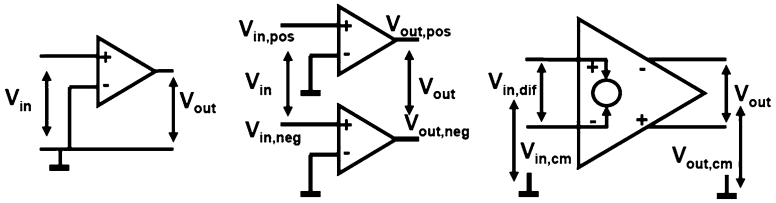


Fig. 2.60 Single-ended amplifier configuration, pseudo-differential, and full-differential design

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 antiphase 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:

$$v_{\text{out,pos}} = 0.5V_a \sin(\omega t) + bV_a^2 \sin(2\omega t) + cV_a^3 \sin(3\omega t). \tag{2.134}$$

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

$$v_{\text{out,neg}} = -0.5V_a \sin(\omega t) + bV_a^2 \sin(2\omega t) - cV_a^3 \sin(3\omega t). \tag{2.135}$$

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

$$v_{\text{out}} = V_a \sin(\omega t) + 2cV_a^3 \sin(3\omega t). \tag{2.136}$$

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. 2.60. 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

$$v_{\text{out}} = A_{\text{com}}(v_{\text{in,pos}} + v_{\text{in,neg}}) + A_{\text{diff}}(v_{\text{in,pos}} - v_{\text{in,neg}}). \tag{2.137}$$

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

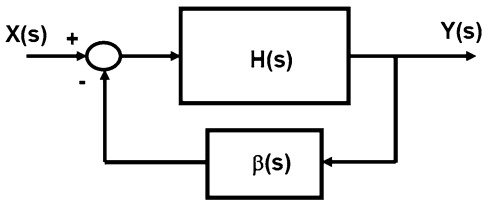


Fig. 2.61 The general scheme of a feedback system

between both is called "the common-mode rejection ratio" or CMRR :

$$\text{CMRR} = 20\,^{10}\log\left(\frac{A_{\text{com}}}{A_{\text{diff}}}\right). \tag{2.138}$$

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:

$$\text{PSRR} = 20\,^{10}\log\left(\frac{\Delta V_{\text{out}}}{\Delta V_{\text{supply}}}\right). \tag{2.139}$$

This ratio defines the transfer from power supply variations to the output of the amplifier. Both the PSRR and the CMRR are frequency dependent.

2.6.6 Feedback

A generally used technique in many systems and of course in analog IC design is feedback. In Fig. 2.61 the transfer function $H(s)$ is connected in a feedback loop via the feedback path multiplier β. Analysis of the resulting transfer results in a Laplace description:

$$\frac{Y(s)}{X(s)} = \frac{H(s)}{1+\beta(s)H(s)}. \tag{2.140}$$

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

The feedback path signal can be added to the input signal or subtracted from the input signal, assuming that there is no inversion present in $\beta H(s)$. Negative feedback or degenerative feedback as drawn in Fig. 2.61 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 signals 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.

The form of the denominator in the transfer function is essential. The main features of a transfer function can often be characterized by the following second-order function[22]:

$$\frac{1}{1+\beta H(s)} = \frac{1}{s^2 + 2\zeta\omega_0 s + \omega_0^2}. \tag{2.141}$$

ζ is the damping factor and ω_0 is the "eigenfrequency" of natural frequency of the system. 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., supply voltage) or it will oscillate. A zero denominator corresponds to the Barkhausen[23] condition for oscillation[24]:

$$|\beta H(s)| = 1 \quad \text{and} \quad \arg(\beta H(s)) = -180^\circ. \tag{2.142}$$

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

$$s_{1,2} = \omega_0(-\zeta \pm \sqrt{\zeta^2 - 1}).$$

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

$\zeta > 1 \rightarrow$ overdamped $\quad f(t) = c_1 e^{\gamma_1 \omega_0 t} + c_2 e^{\gamma_2 \omega_0 t}$

$$\gamma_{1,2} = -\zeta \pm \sqrt{\zeta^2 - 1}$$

$\zeta = 1 \rightarrow$ critically damped $f(t) = c_1 e^{-\omega_0 t} + c_2 t e^{-\omega_0 t}$

[22] 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.

[23] 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.

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

$$\zeta<1 \rightarrow \text{underdamped} \qquad f(t) = c_1 \mathrm{e}^{-\zeta\omega_0 t} \sin(\omega_0 t\sqrt{1-\zeta^2} + \arctan(\sqrt{1-\zeta^2}/\zeta)). \tag{2.143}$$

$f(t)$ shows the basic terms in the time response. Most practical loops are overdamped, 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

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

If $2\zeta^2 > 1$ 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 $2\zeta^2 < 1$ the frequency response will show peaking at $\omega^2 = \omega_0^2(1-2\zeta^2)$. The damping factor ζ 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

$$\sqrt{\frac{(\omega^4 + 2\omega^2\omega_0^2(2\zeta^2-1) + \omega_0^4)\big|_{\omega=0}}{(\omega^4 + 2\omega^2\omega_0^2(2\zeta^2-1) + \omega_0^4) + \omega_0^4)\big|_{\omega^2=\omega_0^2(1-2\zeta^2)}}} = \frac{1}{2\zeta\sqrt{1-\zeta^2}} \approx \frac{1}{2\zeta} = Q. \tag{2.144}$$

For small damping factors ($\zeta \ll 1$) the last term becomes equal to the quality factor Q as defined for filters (see Sect. 2.6.8). With $Q = 1/2\zeta$ the transfer function in the frequency domain for a second-order resonance becomes

$$\frac{1}{\omega^2 + 2\zeta\omega_0\omega + \omega_0^2} = \frac{1}{\omega^2 + \omega_0\omega/Q + \omega_0^2}. \tag{2.145}$$

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

$$f(t) = \frac{1}{2}c_1 \mathrm{e}^{+t/\tau} - \frac{1}{2}c_2 \mathrm{e}^{-t/\tau}, \tag{2.146}$$

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

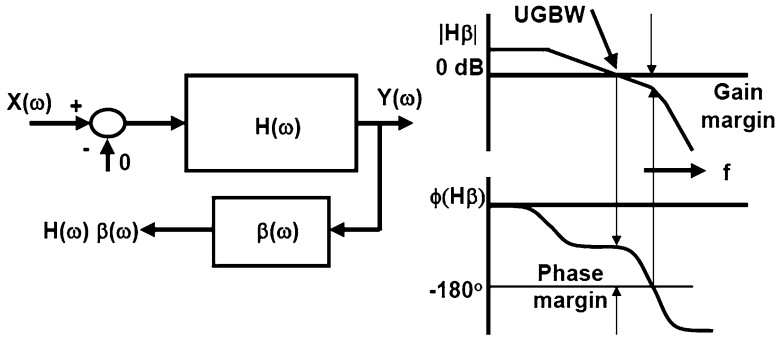


Fig. 2.62 Open-loop analysis of stability using the Bode plots

2.6.7 Bode Plots

Figure 2.62 shows one of the analysis methods for feedback systems: the "Bode plot." For an open-loop analysis, the loop is broken near the summation point. Proper analysis requires to take the loading of the loop at the position where the loop is opened into account.[25] Inspection of the frequency behavior of $\beta H(s)$ requires to replace $s = \alpha + j\omega$ by $s = 0 + j\omega$. The resulting frequency amplitude and phase plots ("Bode plots") 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 the amplitude gain is "$1\times$" (also coined: unity or 0 dB). 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 estimation 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 practical loss of amplification will turn the loop instable. This situation is called "conditionally stable" and not advised.

In this example the open-loop transfer function has 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 communication and electronic applications a logarithmic measure is used for the power gain and amplitude gain:

$$A(\text{in dB}) = 10\,{}^{10}\log\left(\frac{P_Y}{P_X}\right) = 10\,{}^{10}\log\left(\frac{Y^2}{X^2}\right) = 20\,{}^{10}\log\left(\frac{Y}{X}\right) \tag{2.147}$$

[25] An interesting alternative to the stability analysis that does take the mutual loading into account, is proposed by David Middlebrook [42] presents a popular introduction.

assuming equal impedance levels for the signals 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.

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 [43]. 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:

$$h(\tau) = 0, \quad \forall \tau < 0. \tag{2.148}$$

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:

$$y(t) = x(t) * h(t) = \int_{\tau=0}^{\tau=\infty} h(\tau)x(t-\tau)\mathrm{d}\tau = \int_{\tau=0}^{\tau=\infty} h(t-\tau)x(\tau)\mathrm{d}\tau. \tag{2.149}$$

This integral is a convolution function and is most easily evaluated in the Laplace domain:

$$Y(s) = H(s) \times X(s), \tag{2.150}$$

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 LTI filters are analyzed and synthesized in the Laplace domain.

Many filter types exist, depending on their function (low pass, high pass or band pass) or a specific property. However, most properties can be seen on a simple

second-order transfer function, similar to the functions analyzed in the section on feedback systems (Sect. 2.6.6):

$$H(j\omega) = \frac{\omega_0^2}{(j\omega)^2 + 2\zeta\omega_0 j\omega + \omega_0^2}. \tag{2.151}$$

At DC this transfer is unity and at high frequencies the roll-off will be of second order. Modifying ζ gives rise to various shapes of the transition between the unity regime and the roll-off (see, e.g., Fig. 2.64).

An important parameter to characterize the roll-off is the frequency where the signal amplitude is -3 dB lower than unity:

$$|H(j\omega)|^2 = \left|\frac{\omega_0^2}{-\omega^2 + 2j\zeta\omega_0\omega + \omega_0^2}\right|^2 = \frac{\omega_0^4}{\omega^4 + 2\omega^2\omega_0^2(2\zeta^2 - 1) + \omega_0^4}\bigg|_{\omega=\omega_{-3\,\mathrm{dB}}} = \frac{1}{2}$$

resulting in $\omega_{-3\mathrm{dB}} = \omega_0((1-2\zeta^2) \pm \sqrt{(1-2\zeta^2)+1}$. In case $\zeta = 1/\sqrt{2}$ the -3 dB frequency equals the natural frequency ω_0.

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

$$|H(j\omega)|^2 = \frac{1}{1+\omega^{2N}}. \tag{2.152}$$

The definition of maximum flatness means that all high-order derivatives of $|H(j\omega)|^2$ are zero. In case an additional term proportional to ω^{2K} would be present in the denominator the $2K$th derivative would be nonzero.

The conditions for the complex poles of a second-order Butterworth filter $s_{1,2} = -a \pm jb$ can now be calculated:

$$H(s) = \frac{1}{(s+a)^2 + b^2}.$$

The maximum flatness condition for a Butterworth filter is

$$|H(j\omega)|^2 = \frac{1}{(a^2+b^2)^2 + (2a^2 - 2b^2)\omega^2 + \omega^4},$$

where the coefficient of the ω^2 term must be zero, so $a = \pm b$. A -3 dB attenuation occurs at a frequency where $\omega_{-3\,\mathrm{dB}}^2 = a^2 + b^2$ and the poles are found at

Table 2.24 Transfer functions for low-pass filters with a cut-off frequency at 1 rad/s

Order	Butterworth	Bessel	Chebyshev
1	$\frac{1}{s+1}$	$\frac{1}{s+1}$	$\frac{1}{s+1}$
2	$\frac{1}{s^2+1.41s+1}$	$\frac{1}{s^2+3s+3}$	$\frac{1}{1.41s^2+0.911s+1}$
3	$\frac{1}{s^3+2s^2+2s+1}$	$\frac{1}{s^3+6s^2+15s+15}$	$\frac{1}{3.98s^3+2.38s^2+3.7s+1}$
4	$\frac{1}{s^4+2.61s^3+3.41s^2+2.61s+1}$	$\frac{1}{s^4+10s^3+36s^2+95s+105}$	$\frac{1}{5.65s^4+3.29s^3+6.6s^2+2.3s+1}$

$$s_{1,2} = -\omega_{-3\,\mathrm{dB}}\sqrt{0.5} \pm j\omega_{-3\,\mathrm{dB}}\sqrt{0.5}$$

$$H(s) = \frac{1}{(s/\omega_{-3\,\mathrm{dB}})^2 + (s/\omega_{-3\,\mathrm{dB}})\sqrt{2}+1}$$

$$H(s) = \frac{\omega_{-3\,\mathrm{dB}}^2}{s^2 + s\omega_{-3\,\mathrm{dB}}\sqrt{2} + \omega_{-3\,\mathrm{dB}}^2}$$

$$\omega_0 = \omega_{-3\,\mathrm{dB}}, \quad \zeta = 1/\sqrt{2}.$$

In filter textbooks these formulas are often normalized to $\omega_{-3\,\mathrm{dB}} = 1$ rad/s (see Table 2.24):

$$H_{\mathrm{normalized}}(s) = \frac{1}{s^2 + s\sqrt{2}+1}.$$

- Bessel filters optimize the flatness of the group delay. The group delay is defined as:

$$\text{group delay} = \tau_g = -\frac{\partial \arg(H(\omega))}{\partial \omega} = -\frac{\partial \phi(\omega)}{\partial \omega}. \tag{2.153}$$

  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:

$$\frac{\phi(\omega)}{\omega} = \text{constant}. \tag{2.154}$$

  The requirement of linear phase may apply only in a part of the frequency range, depending on the application.
- 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 bands.

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.

Example 2.17. Determine the amplitude, phase, and group delay of a first-order filter.

Solution. A first-order filter is defined as

$$H(\omega)=\frac{1}{1+j\omega\tau}.$$

The -3 dB bandwidth of this filter is $f_{-3\,\mathrm{dB}}=1/2\pi\tau$. Now the amplitude, phase, and group delay are (see Table 2.4)

$$|H(\omega)|^2=H(\omega)\times H^*(\omega)=\frac{1}{1+\omega^2\tau^2}, \tag{2.155}$$

$$\mathrm{Phase}(H(\omega))=\arg(H(\omega))=\arctan\left(\frac{-\omega\tau}{1}\right), \tag{2.156}$$

$$\text{Group delay}(H(\omega))=\tau_g(\omega)=-\frac{\mathrm{d}\arg(H(\omega))}{\mathrm{d}\omega}=\frac{\tau}{1+\omega^2\tau^2}. \tag{2.157}$$

Example 2.18. A first-order filter has an attenuation of $\sqrt{1/2}=0.7071$ or 3 dB at the corner frequency f_c. Higher-order filters can be realized by cascading first-order sections. Calculate the filter order and the corner frequency if the attenuation at a given bandwidth f_BW must remain at 3 dB.

Solution. A cascade of first-order filters gives as amplitude transfer:

$$|H(f)|=\left|\frac{1}{1+\omega^2/\omega_c^2}\right|^{k/2}=\left|\frac{1}{1+f^2/f_c^2}\right|^{k/2},$$

where k is the order and f_c the corner frequency. If $|H(f=f_\mathrm{BW})|=\sqrt{1/2}$ is given then f_c can be calculated for $k=1,\ldots,8$ (see Fig. 2.63).

2.6.9 RLC Filters

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

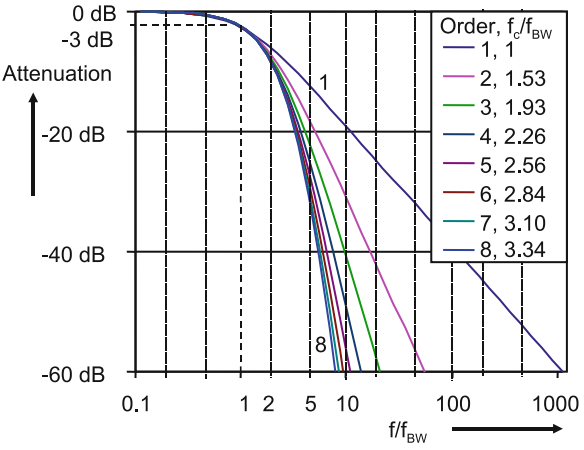


Fig. 2.63 The corner frequency f_c is chosen such that the attenuation of k cascaded first-order filter sections at $f = f_{BW}$ is 3 dB. The frequencies are normalized to the bandwidth f_{BW}

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 filter.

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 Eq. 2.145):

$$\omega_0 = 2\pi f_0 = \frac{1}{\sqrt{LC}}. \tag{2.158}$$

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 breakdown of an insulator. Normally the LC circuit will also contain some resistor, in the example of Fig. 2.64 a resistance in series with the coil. The resistor will dissipate some energy, thereby limiting the total energy available. The ratio

$$\zeta\omega_0 = \frac{R}{2L} \tag{2.159}$$

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.

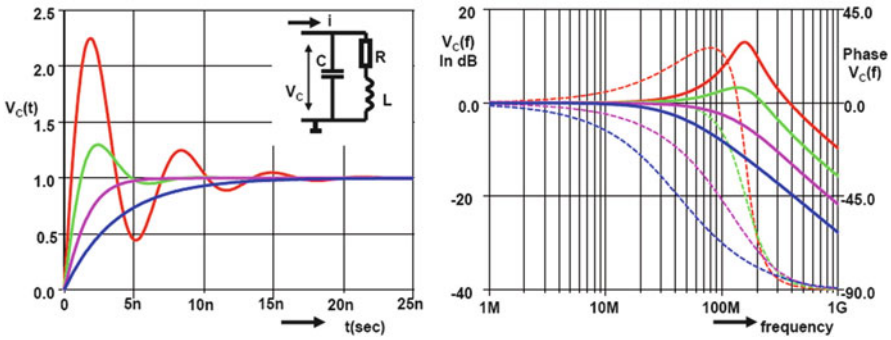


Fig. 2.64 A parallel connection of a capacitor with a coil and resistor with $L = 1\,\mu\text{H}$, $C = 1\,\text{pF}$, and $R = 500, 1000, 2000$, and $4000\,\Omega$. The quality factor varies from 2, 1, 0.5 to 0.25. In the time domain a high-quality factor creates ringing; in the frequency domain, the transfer curve peaks; $Q = 0.5$ corresponds to critical damping

The ratio between the resistor and impedance of the coil or capacitor at the resonance frequency is called the quality factor:

$$Q = \frac{\omega_0 L}{R} = \frac{1}{\omega_0 RC} = \frac{1}{R}\sqrt{\frac{L}{C}}. \tag{2.160}$$

Compare Eq. 2.144 to Eq. 2.160. 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 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. 2.64 is calculated in the Laplace domain:

$$Z(s) = \frac{R + sL}{LCs^2 + sRC + 1} = \frac{1}{LC}\frac{R + sL}{s^2 + 2\zeta\omega_0 s + \omega_0^2}. \tag{2.161}$$

The condition for critical damping is $\zeta = 1$ which reduces the denominator to $(s + \omega_0)^2$ and the time response to an exponential decaying function $e^{-\omega_0 t}$. If $\zeta > 1$, or the condition of "overdamping," the time-domain response of the filter is composed of two exponential decaying terms. If $\zeta < 1$, or the condition of "underdamping," the time-domain response is composed of an exponentially decaying sine wave (see Eq. 2.143).

Combinations of coils and capacitors implement a second-order resonance function as in Eq. 2.145. In integrated circuits realizations exist that do not require

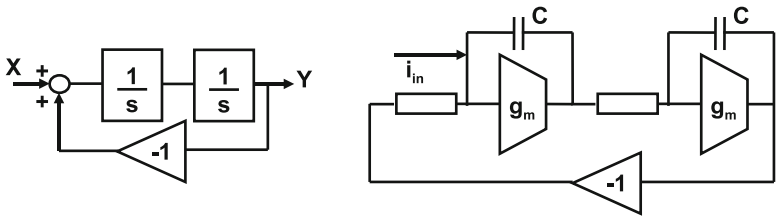


Fig. 2.65 The ideal resonator consists of two integration functions in a unity feedback configuration. *Right*: a practical $g_m - C$ resonator

the use of coils. Figure 2.65 shows a popular example of a resonator constructed with a transconductance stage and a capacitor.

Example 2.19. Design a second-order RLC Bessel filter with a -3 dB frequency of 5 MHz.

Solution. Moving the cutoff frequency from $\omega_{-3\,\mathrm{dB}} = 1$ rad/s to $f_{-3\,\mathrm{dB}} = 5$ MHz means multiplying the poles with $2\pi \times f_{-3\,\mathrm{dB}}$ or substituting the parameter s in Table 2.24 by $s/(2\pi f_{-3\,\mathrm{dB}})$.

Comparison of the transfer function of a simple R-L-C low-pass filter gives

$$H(s) = \frac{1}{s^2 + \sqrt{2}s \times 2\pi \times 5 \times 10^6 + 4\pi^2 \times 25 \times 10^{12}} = \frac{1}{s^2 + 4.42 \times 10^7 s + 9.8 \times 10^{14}}$$

$$H(s) = \frac{1}{s^2 + s\frac{R}{L} + \frac{1}{LC}}$$

$$\frac{R}{L} = 4.42 \times 10^7 \quad LC = 1/(9.8 \times 10^{14}).$$

Choosing $R = 50, \rightarrow L = 1.13\,\mu\mathrm{H}, \quad C = 909\,\mathrm{pF}$

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.

2.6.10 Sallen–Key $g_m - C$ Filters and Gyrators

Next to these transfer functions the choice for the topology must be made. A popular scheme is the Sallen–Key realization (see Fig. 2.66) [44]. The transfer function is

$$H(s) = \frac{1}{s^2 R_1 R_2 C_1 C_2 + sC_2(R_1 + R_2) + 1}$$

$$= \frac{1}{R_1 R_2 C_1 C_2 (s^2 + s(R_1 + R_2)/R_1 R_2 C_1 + 1/R_1 R_2 C_1 C_2)}.$$

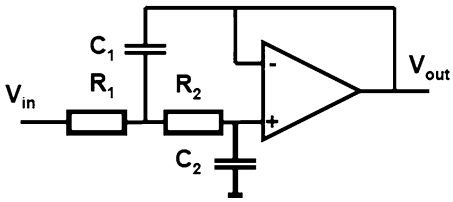


Fig. 2.66 The Sallen–Key active filter topology

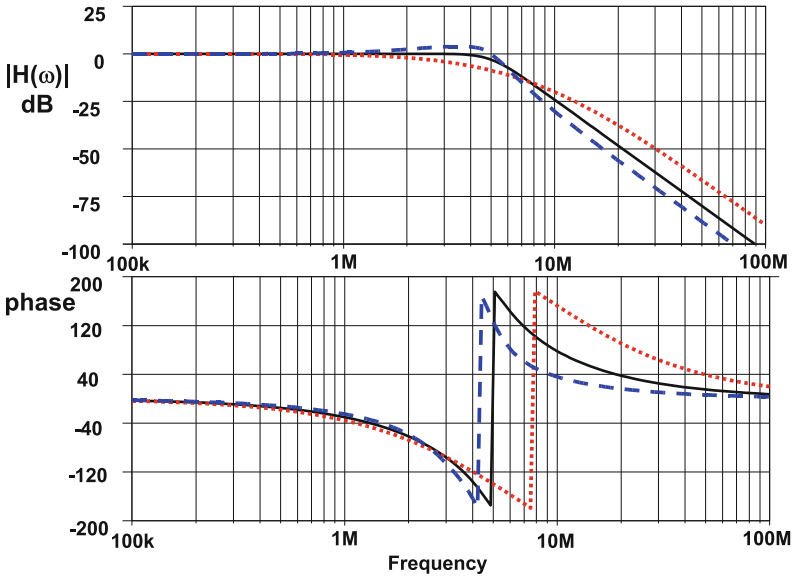


Fig. 2.67 The Sallen–Key transfer function for a fourth-order Butterworth (*solid*), Bessel (*dotted*), and Chebyshev (*dashed*) filter. The filter is composed of two cascaded sections of Fig. 2.66

Or in terms of a general second-order system with its damping and resonance frequency,

$$H(s) = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \quad \text{with} \tag{2.162}$$

$$\omega_0^2 = \frac{1}{R_1 R_2 C_1 C_2}, \quad 2\zeta\omega_0 = \frac{R_1 + R_2}{R_1 R_2 C_1}. \tag{2.163}$$

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 Table 2.24. Sallen–Key filters can easily be cascaded to define higher-order filter functions. A result for a fourth-order filter is given in Fig. 2.67. Changes in the component values 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.

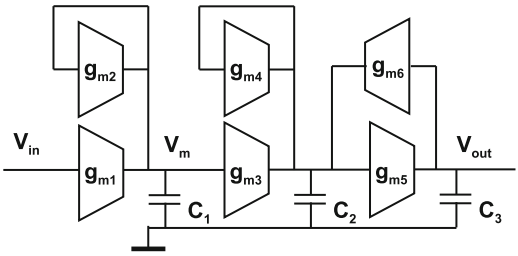


Fig. 2.68 A filter realized with transconductances and capacitors. The first section is a first-order filter, the second section a second-order filter

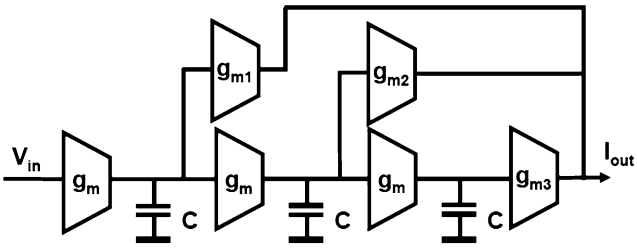


Fig. 2.69 The filter coefficients of this feed-forward $g_m - C$ filter are implemented with weighted transconductances

A popular filter technique in the field of integrated circuits is the $g_m - C$ filter [45]. 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 2.68 shows a third-order filter consisting of a first-order and a second-order section, with a transfer function:

$$H(s) = \frac{V_{\text{out}}(s)}{V_{\text{in}}(s)} = \frac{g_{m1}/C_1}{(s + g_{m2}/C_1)} \frac{g_{m3}g_{m4}/C_2C_3}{(s^2 + sg_{m5}/C_3 + g_{m4}g_{m6}/C_2C_3)}. \tag{2.164}$$

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 certain extent independent of the damping implementation. With equal g_m and C the transfer is

$$H(s) = \frac{V_{\text{out}}(s)}{V_{\text{in}}(s)} = \frac{1}{(1 + sC/g_m)(1 + sC/g_m + (sC/g_m)^2)}. \tag{2.165}$$

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. 2.69. Now the filter transfer function is

$$\begin{aligned} H(s) &= \frac{I_{\text{out}}(s)}{V_{\text{in}}(s)} = g_{m1}\left(\frac{g_m}{sC}\right) + g_{m2}\left(\frac{g_m}{sC}\right)^2 + g_{m3}\left(\frac{g_m}{sC}\right)^3 \\ &= \frac{(g_{m3} + g_{m2}sCg_{m2}/g_m + g_{m1}s^2C^2/g_m^2)}{(sC/g_m)^3}. \end{aligned} \tag{2.166}$$

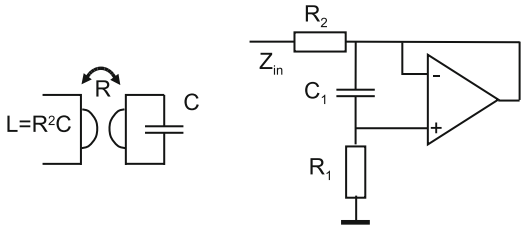


Fig. 2.70 A simple gyrator implementation

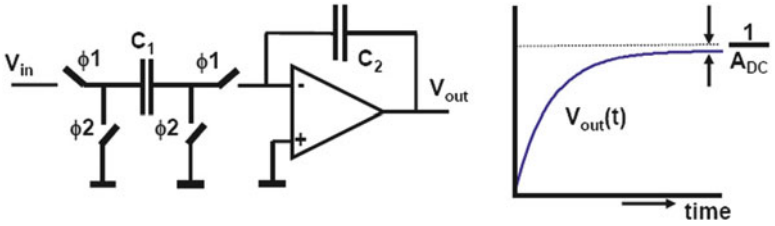


Fig. 2.71 An elementary switched-capacitor circuit

For many years gyrator-filter topologies were extensively studied. A gyrator is a mathematical model of a two-port device. The voltage on port 1 is proportional to the current in port 2 an vice verse via a constant R with the dimension Ω. A capacitor on one side is equivalent to a coil with value R^2C on the other side. In [46] the implementation of Fig. 2.70 is proposed. If the current through R_1 and C_1 is ignored the input impedance is $R_2 + j\omega R_1 R_2 C_1$, which is equivalent to a resistor and coil in series. Also topologies using cross-coupled pairs are used.

Example 2.20. Design a second-order low-pass filter using capacitors, resistors, and ideal opamps for a signal band of 20 kHz.

Solution. A Sallen–Key filter can fulfill these requirements with $\omega_0 = 2 \times 2 \times 10^4$. The filter characteristics and the impedance level are still free to choose. $2\zeta\omega_0 = \omega_0\sqrt{(2)}$ will result in a Butterworth filter. Equation 2.162 indicates that

$$\omega_0^2 = \frac{1}{R_1 R_2 C_1 C_2} = (2\pi 2 \times 10^4)^2,$$

$$2\zeta\omega_0 = \omega_0\sqrt{(2)} = \frac{R_1 + R_2}{R_1 R_2 C_1}.$$

Choosing $R_1 = R_2 = 1{,}000\,\Omega$, the capacitors are found $C_1 = 11.3$ nF and $C_2 = 5.5$ nF.

2.6.11 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 [47–53]. Figure 2.71

shows a basic integrator configuration. After every clock cycle a charge sample $C_1 V_{\text{in}}$ 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 $V_{\text{in}} C_1/(C_1 + C_2)$. 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 nth sample at the output equals

$$V_{\text{out}}(z)z^n = V_{\text{out}}(z)z^{n-1} + \frac{C_1}{C_2} V_{\text{in}}(z)z^n \rightarrow \frac{V_{\text{out}}(z)}{V_{\text{in}}(z)} = \frac{C_1/C_2}{1 - z^{-1}}. \tag{2.167}$$

In case the gain of the amplifier is limited to A_{DC} the output voltage will show a small deviation of the ideal case [47]:

$$\frac{V_{\text{out}}(z)}{V_{\text{in}}(z)} = \frac{C_1/C_2}{1 - z^{-1}} \frac{1}{1 + \frac{C_1 + C_2}{A_{\text{DC}} C_2}} \approx \frac{C_1/C_2}{1 - z^{-1}} \left(1 - \frac{C_1 + C_2}{A_{\text{DC}} C_2}\right). \tag{2.168}$$

The fraction $V_{\text{out}}/A_{\text{DC}}$ 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 $1/A_{\text{DC}}$.
- 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 $V_{\text{in}} - V_{\text{out}}/A_{\text{DC}}$. 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:

$$\left| \frac{C_{1,\text{real}}}{C_{2,\text{real}}} - \frac{C_{1,\text{ideal}}}{C_{2,\text{ideal}}} \right| < \frac{1}{A_{\text{DC}}}. \tag{2.169}$$

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. 3.1.6). One charge transfer cycle with two switch configurations results in an input-referred noise of $v^2_{\text{noise}} = 2kT/C$. 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 shorthand analysis, the contributions of the switched capacitors can be added up in the energy domain. For an extensive analysis see, e.g., [53].

The bandwidth of the switched-capacitor integrator depends on the UGBW 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 UGBW $\omega_{\text{UGBW}} = A_{\text{DC}}/\tau$ and feedback factor β:

$$\tau_{\text{sc}} = \frac{\tau}{\beta A_{\text{DC}}} = \frac{1}{\beta \omega_{\text{UGBW}}}. \tag{2.170}$$

In order to obtain a settling error smaller than $1/A_{\text{DC}}$, A_{DC} should be in the range of $\approx 1{,}000 - 10{,}000$ for a $N = 10 - 14$-bit analog-to-digital converter. A time period T_{sc} is needed of some 7 to 10 time constants:

$$\mathrm{e}^{-T_{\text{sc}}/\tau_{\text{sc}}} < \frac{1}{A_{\text{DC}}} < 2^{-N}. \tag{2.171}$$

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

$$I_{\text{out,slew}} = \frac{V_{\text{out,max}}(C_2 + C_{\text{load}})}{\tau_{\text{sc}}}. \tag{2.172}$$

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

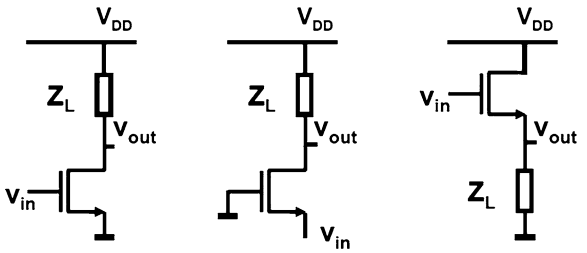


Fig. 2.72 Single-transistor amplifiers: grounded-source, grounded-gate, and grounded-drain configurations

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. 4.21.

2.7 Electronic Circuits

2.7.1 Classification of Amplifiers

Electronic circuit design is extensively described in, e.g., [54, 55]. 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. 2.72).[26] Single-transistor amplification stages are classified along the terminal that is connected to signal 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. 2.73 are drawn with one or two transistors. In reality complex circuits implement the general ideas that are common to these simple amplifier stages.

[26] 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 in an improper use of the ground symbol in this book.

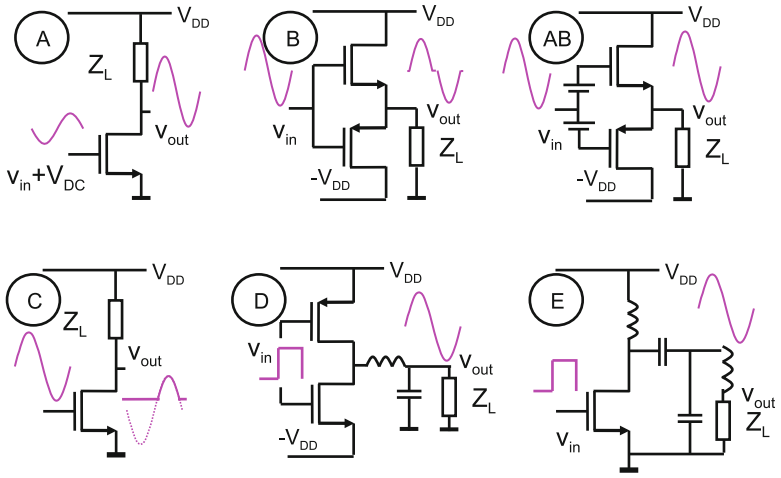


Fig. 2.73 The modes of operation in transistor amplifiers are described by letters: class A, B, AB, C, D, E, etc

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

$$\frac{\int_{t=0}^{t=T} \hat{V}\sin(\omega t) I \sin(\omega t)\mathrm{d}t}{\int_{t=0}^{t=T} 2\hat{V} I_{\mathrm{load}}\mathrm{d}t} = 25\% \tag{2.173}$$

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 even-order distortion is not appreciated by the human ear.[27] 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

[27] In contrast to odd-order distortion as generated by transformers in tube amplifiers.

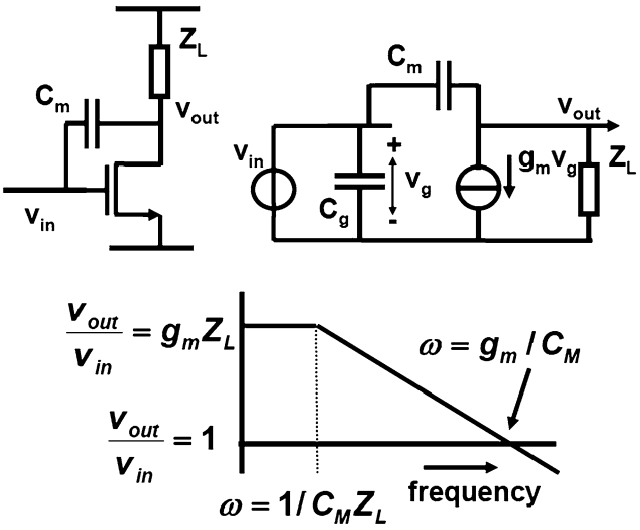


Fig. 2.74 A transistor as amplifier

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. 2.74. The input voltage to the gate is composed of a DC voltage with superimposed a signal voltage: $V_{\mathrm{GS}} + v_{gs}$. On the drain side also a DC component and a signal component will appear: $V_{\mathrm{DS}} + v_{ds}$. Expanding the square-law Eq. 2.105 gives the total output current:

$$\frac{WC_{\mathrm{ox}}\mu}{2L}(V_{\mathrm{GS}} + v_{gs} - V_{\mathrm{T}} + \lambda(V_{\mathrm{DS}} + v_{ds}))^2$$

$$= \frac{WC_{\mathrm{ox}}\mu}{2L}(V_{\mathrm{GS}} - V_{\mathrm{T}} + \lambda V_{\mathrm{DS}})^2 + \frac{WC_{\mathrm{ox}}\mu}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}} + \lambda V_{\mathrm{DS}})v_{gs}$$

$$+ \frac{WC_{\mathrm{ox}}\mu}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}} + \lambda V_{\mathrm{DS}})\lambda v_{ds} + \frac{WC_{\mathrm{ox}}\mu}{2L}(v_{gs}^2 + \lambda^2 v_{ds}^2 + 2\lambda v_{gs}v_{ds})$$

$$= I_{\mathrm{DS}} + g_m v_{gs} + g_{ds} v_{ds} + i_{\text{second order}}.$$

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:

$$i_{ds} = g_m v_{gs} + g_{ds} v_{ds} \tag{2.174}$$

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 of Fig. 2.74 (right) can be evaluated by using the small-signal equivalent diagram. The transfer is found as[28]

$$\frac{v_{\mathrm{out}}}{v_{\mathrm{in}}} = \frac{Z_L(g_m - j\omega C_M)}{1 + j\omega Z_L C_M}. \tag{2.175}$$

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 UGBW:

$$\frac{v_{\mathrm{out}}}{v_{\mathrm{in}}} = \left| \frac{g_m}{j\omega_{\mathrm{UGBW}} C_M} \right| = 1, \qquad f_{\mathrm{UGBW}} = \frac{g_m}{2\pi C_M}. \tag{2.176}$$

[28]The easiest start to the mathematical analysis is to list Kirchhoff's current equations and start eliminating variables.

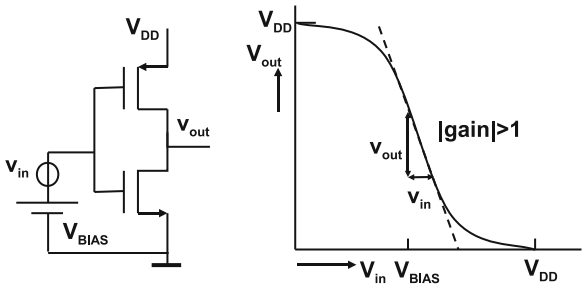


Fig. 2.75 Large- and small-signal transfer for an inverter

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 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 Eq. 2.175 becomes 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 output 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

$$\frac{v_{\text{in}}}{i_{\text{in}}} = \frac{1 + j\omega Z_L C_M}{j\omega C_M (Z_L g_m + 1)}. \tag{2.177}$$

For low frequencies the numerator can be set to "1," and the effective input capacitance equals the value of the Miller capacitor times (1+ DC amplification). This allows to create large load capacitors without the need for spending many picoFarads. 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 $+C_M V_{\text{DD}}$ to $-C_M V_{\text{DD}}$.

2.7.3 Inverter

The large-signal amplification of the inverter in Fig. 2.75 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 midrange 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

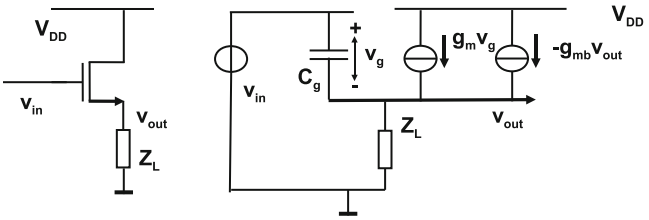


Fig. 2.76 The source follower circuit and its small-signal equivalent diagram

transistors as a load, so the amplification is

$$A_V = \frac{v_{\text{out}}}{v_{\text{in}}} = -\frac{g_{mN} + g_{mP}}{g_{dsN} + g_{dsP}} \approx -6 \cdots -10. \tag{2.178}$$

In large-signal mode the voltage over the Miller capacitor of an inverter (the capacitor between the input and output terminal) is changing its polarity. Effectively 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 its 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. 2.76). 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:

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{Z_L(g_m + j\omega C_g)}{1 + j\omega Z_L C_g + g_m Z_L + g_{mb} Z_L}. \tag{2.179}$$

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

$$\frac{v_{\text{in}}}{i_{\text{in}}} = \frac{1 + j\omega Z_L C_g + g_m Z_L + g_{mb} Z_L}{j\omega C_g}. \tag{2.180}$$

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

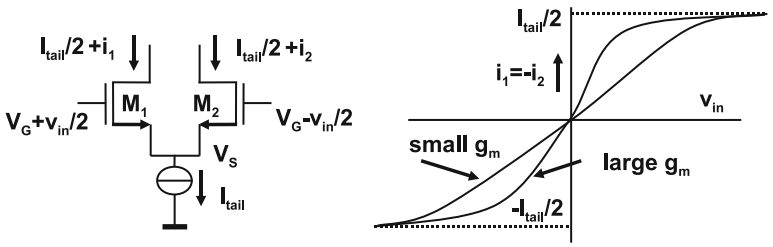


Fig. 2.77 The differential pair and the DC transfer curve for two values of the transconductance

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 on the source node is

$$\frac{v_{\text{out}}}{i_{\text{out}}} = \frac{Z_L}{1 + j\omega Z_L C_g + g_m Z_L} \tag{2.181}$$

which goes to $1/g_m$ in practical design.

2.7.5 *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 2.77 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. 2.78). In linear approximation the small-signal current is related to the small-signal input via the transconductance of the input transistors:

$$i = i_1 - i_2 = \frac{\mathrm{d}I_{\text{DS1}}}{\mathrm{d}V_{G_1}} v_{\text{in}}/2 - \frac{\mathrm{d}I_{\text{DS2}}}{\mathrm{d}V_{G_2}}(-v_{\text{in}}/2) = g_m v_{\text{in}} \tag{2.182}$$

assuming equal transistor sizes and bias conditions. In advanced design this approximation is not sufficient and a nonlinear analysis is used. The transfer is linear

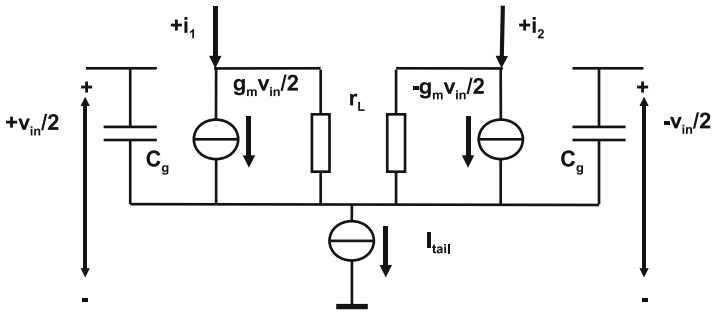


Fig. 2.78 The small-signal equivalent scheme of the differential pair. The ideal tail current source behaves for small signals as a circuit open

up to a certain input voltage excursion, then saturation will occur. Figure 2.77 (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 (Eq. 2.82). Therefore the relation between the differential collector current and the input voltage is also independent of sizes:

$$\frac{i}{I_{\text{tail}}} = \frac{I_{C1} - I_{C2}}{I_{C1} + I_{C2}} = \frac{\mathrm{e}^{q\mathrm{v}_{\text{in}}/2kT} - \mathrm{e}^{-q\mathrm{v}_{\text{in}}/2kT}}{\mathrm{e}^{q\mathrm{v}_{\text{in}}/2kT} + \mathrm{e}^{-q\mathrm{v}_{\text{in}}/2kT}} = \tanh(q/kT \times \mathrm{V}_{\text{in}}/2). \tag{2.183}$$

This equation shows that the transfer of bipolar differential pairs becomes rather nonlinear if the differential voltage approaches the thermal voltage of $kT/q = 26\,\mathrm{mV}$ at room temperature. In MOS design there is an additional degree of freedom: the width over length W/L ratio of the gate. Figure 2.77 (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: $I_{\text{DS}} = f(V_G - V_S)$, [9]. 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 Eq. 2.105 or for bipolar devices of Eq. 2.82.

For the differential pair controlled by a fully symmetrical signal this results in the set of equations:

$$\begin{aligned} I_{\text{DS1}} &= I_{\text{tail}}/2 + i_1 = f(V_{G_1} - V_S) = f(V_G + v_{\text{in}}/2 - V_S) \\ I_{\text{DS2}} &= I_{\text{tail}}/2 + i_2 = f(V_{G_2} - V_S) = f(V_G - v_{\text{in}}/2 - V_S). \end{aligned} \tag{2.184}$$

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

$$v_{\text{in}}/2 = h(I_{\text{tail}}/2 + i_1) - h(I_{\text{tail}}/2 - i_2). \tag{2.185}$$

i_1 is (much) smaller than the tail current, so a Taylor expansion can be used for the function h:

$$h(I_{\text{tail}}/2 + i_1) = h(I_{\text{tail}}/2) + \frac{i_1}{1!}h'(I_{\text{tail}}/2) + \frac{i_1^2}{2!}h''(I_{\text{tail}}/2) + \frac{i_1^3}{3!}h'''(I_{\text{tail}}/2) + \ldots.$$

In an ideal differential structure with $i = i_1 - i_2$ the even terms are canceling each other while the odd components are doubling:

$$v_{\text{in}} = \frac{i}{1!}h'(I_{\text{tail}}/2) + \frac{i^3}{3!}h'''(I_{\text{tail}}/2) + \ldots.$$

This series of terms can be reversed [7, Equation 3.6.25]:

$$i = \frac{1}{h'(I_{\text{tail}}/2)}v_{\text{in}} - \frac{h'''(I_{\text{tail}}/2)}{3!(h'(I_{\text{tail}}/2))^4}v_{\text{in}}^3 + \ldots.$$

When only the third-order term is taken into account and some elementary calculus is applied [7, Equations 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,

$$I_C = f(V_{\text{be}}) = I_o e^{qV_{\text{be}}/kT} \quad \text{with the } n\text{th derivative} \quad f^{(n)} = \left(\frac{q}{kT}\right)^n I_C \tag{2.186}$$

the resulting expression is

$$\frac{i}{I_{\text{tail}}} = \frac{v_{\text{in}}}{2kT/q} - \left(\frac{v_{\text{in}}}{kT/q}\right)^3. \tag{2.187}$$

With a first-order square-law model for the MOS transistors, the MOS current I_{DS} equals

$$\begin{aligned} I_{\mathrm{DS}} &= f((V_G - V_S) = \frac{\beta}{2}(V_G - V_S - V_T)^2 \\ & f'((V_G - V_S) = \beta(V_G - V_S - V_T) = g_m \\ & f''((V_G - V_S) = \beta \\ & f'''((V_G - V_S) = 0 \end{aligned} \tag{2.188}$$

resulting in

$$i = g_m v_{\mathrm{in}} - \frac{\beta^2}{2g_m} v_{\mathrm{in}}^3 \quad \text{or} \quad \frac{i}{I_{\mathrm{tail}}} = \frac{v_{\mathrm{in}}}{(V_G - V_S - V_{\mathrm{T}})} - \left(\frac{v_{\mathrm{in}}}{V_G - V_S - V_{\mathrm{T}}}\right)^3. \tag{2.189}$$

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. 2.77 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 antiphase 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 Eq. 2.184 the voltage V_S can be found as

$$V_S = V_G - (h(I_{\mathrm{tail}}/2 + i_1) + h(I_{\mathrm{tail}}/2 - i_2)). \tag{2.190}$$

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:

$$\frac{v_{s,2}}{(V_G - V_S - V_{\mathrm{T}})} = \frac{v_{\mathrm{in}}^2}{2(V_G - V_S - V_{\mathrm{T}})^2}. \tag{2.191}$$

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 becomes relevant the analysis in this section can serve as a starting point.

Differential amplifiers allows to operate on the difference between two voltages. These voltages are referenced to ground or another fixed voltage. In fully differential

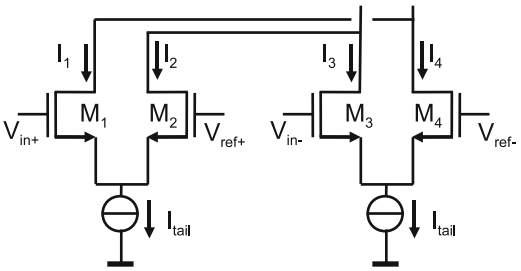


Fig. 2.79 A dual differential pair or differential difference pair allows to subtract two floating voltages. In this case $(V_{\text{in}+} - V_{\text{in}-})$ and $(V_{\text{ref}+} - V_{\text{ref}-})$ could represent the input and reference voltage in analog-to-digital converters

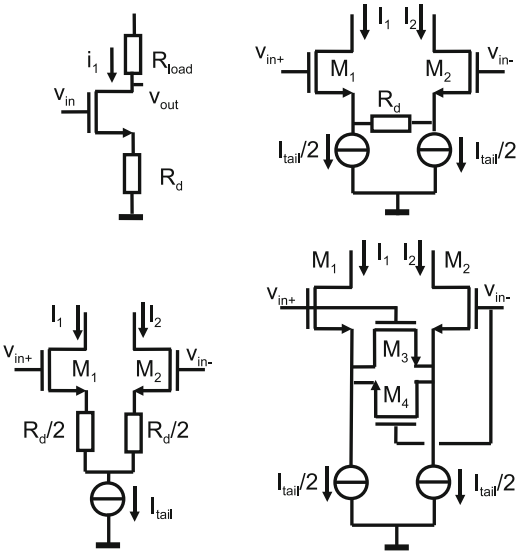


Fig. 2.80 Degeneration in various circuits

designs, subtraction of floating voltages is required. Next to capacitive means, the differential difference pair allows to subtract floating voltages. In Fig. 2.79 two floating voltages are subtracted. Care must be taken to keep the common-mode voltages of both floating voltages closely together. It is possible to connect $(V_{\text{in}+}$ and $V_{\text{in}-})$ to one differential pair and $(V_{\text{ref}+}$ and $V_{\text{ref}-})$ to the other, but this will unbalance both differential pairs considerably.

2.7.6 *Degeneration*

The nonlinearity of the voltage-to-current transfer of a MOS and bipolar transistor can be reduced with degeneration. Figure 2.80 (upper left) shows a single transistor with a degeneration resistor R_d. The signal current is found as

$$i_1 = \frac{g_m}{1 + g_m R_d} v_{\text{in}}.$$

For the condition[29] $g_m R_d \gg 1$ the signal current is simply v_{in}/R_d. If the drain is connected via a resistor R_{load} to the power supply, a simple voltage amplification of $-R_{\text{load}}/R_d$ results.

Degeneration also affects the output impedance of a MOS transistor. In case of a single MOS transistor with a resistor R_d to ground, the output impedance is increased from r_{out} to $r_{\text{out}}(1 + g_m R_d)$.

The differential pair is also linearized by 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. 2.80). The transfer is

$$v_{\text{in}} = (i_1 - i_2)(R_d/2 + 1/g_m). \tag{2.192}$$

The degeneration resistor helps to linearize the voltage to current transfer and allows to process larger input voltages. In the topology in Fig. 2.80 (middle) 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 [56]).

In Fig. 2.80 (right) a very popular degeneration implementation is shown [57]. The resistor is replaced by two transistors M_3 and M_4 which linearizes the input range. Several variants are possible, e.g., with fixed voltages on the gates of M_3 and M_4.

2.7.7 Mixers and Variable Gain Amplifiers

Variable-gain amplifiers are needed to adjust the input signal range to the signal-handling capabilities of the succeeding circuits, such as analog-to-digital converters. Their function is mathematically expressed as

$$V_{\text{out}} = \frac{V_{A,\text{in}} \times V_{B,\text{in}}}{V_{\text{bias}}}.$$

It is clear that this description in voltages can be modified to an expression where one or more terms are replaced by currents. A CMOS variable-gain topology

[29] The transconductance of bipolar transistors equals the DC current divided by the thermal voltage. This condition is therefore met if the bias voltage over the resistor is $4\text{–}5 \times kT/q \approx 125\,\text{mV}$. The same ratio holds for a MOS transistor, with the remark that the thermal voltage is replaced by $V_{\text{GS}} - V_{\text{T}} \approx 200\,\text{mV}$, and consequently the bias voltage over the resistor reaches the 1 V level, clearly illustrating the limitation of this linearization technique in MOS design (Remark by W. Redman-White).

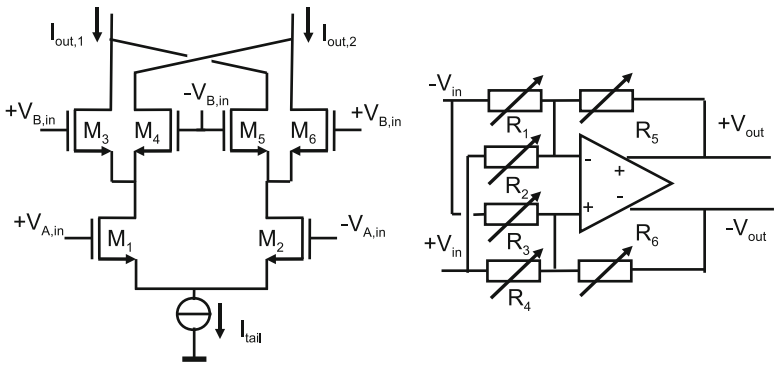


Fig. 2.81 *Left*: the Gilbert cell multiplier. *Right*: general topology for variable-gain amplifiers

is shown in Fig. 2.81 (left). Originally this topology was proposed with bipolar transistors [58]. The exponential relation between current and voltage of Eq. 2.82 allows to design a multiplier circuit that will allow a high-quality and high-speed multiplication over a large-signal range. The bipolar but also the CMOS variant are still very popular as RF mixers. Often one of the two input signals can be replaced by a block wave.

Variable-gain amplifiers, with a controlling signal of much lower bandwidth than the input signal, can be designed with other architectures. The basic operation is often an implementation of Fig. 2.81 (right). Amplification stages with variable degeneration (Fig. 2.80) implement the topology of Fig. 2.81 (right) with variable R_1 and R_4. In "passive mixers" the resistors R_1 to R_4 are MOS switches. Examples of variable-gain amplifiers in advanced CMOS processes are found in [59].

2.7.8 Current Mirror

The current mirror copies current from one branch into another (see Fig. 2.82). 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

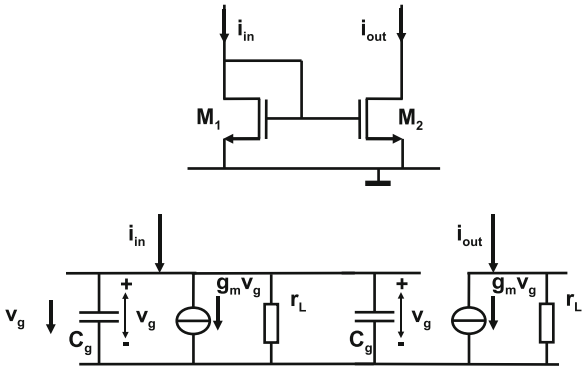


Fig. 2.82 The current mirror and its small-signal equivalent scheme

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 Eq. 2.105).

Based on the small-signal schematic diagram, the input–output transfer is (for equally sized transistors)

$$\frac{i_{\text{out}}}{i_{\text{in}}} = \frac{g_m}{g_m + 1/r_L + 2j\omega C_g} - \frac{v_{\text{out}}}{i_{\text{in}} r_L}. \tag{2.193}$$

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 2.83 shows two alternatives to the regular current mirror. The left schematic is popular in bipolar configurations. The base current creates an offset 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 M_3 also the feedback in the circuit helps to improve the output impedance. To visualize this effect assume that a current is forced into the

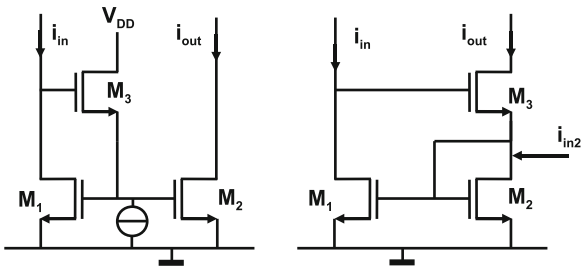


Fig. 2.83 The buffered current mirror and the Wilson mirror

output terminal. This current will see the output impedance of device M_3 but will also be copied by device M_2 into device M_1. As a result this device will pull down the gate of device M_3, which strongly opposes the injected current.

The source node of device M_3 forms a low-ohmic input terminal. A small negative voltage change Δv on this node leads to a current change in device M_1 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 M_3, 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:

$$\frac{\Delta v}{\Delta i_{\text{source3}}} = \frac{1}{g_{m3}g_{m1}R_{\text{drain1}}}. \tag{2.194}$$

There is similarity between the Wilson mirror and the regulated cascode in Fig. 2.85. If the gate of device M_2 of the Wilson mirror is disconnected, the devices M_2 and M_3 form the right-hand branch of the regulated cascode where device M_1 is a simple amplifier with its own threshold voltage featuring as V_{cas} in Fig. 2.85.

Both current mirrors in Fig. 2.83 have as a disadvantage the stack of two threshold voltages in their input terminals, which makes them impractical for low-supply voltage applications.

2.7.9 Cascode and Regulated Cascode

The finite output impedance of a MOS transistor often limits the achievable voltage gain in high-precision design. Some improvement is possible with degeneration (see Sect. 2.7.6). Cascoding[30] transistors, as shown in Fig. 2.84 (left), are a more effective 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

[30]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.

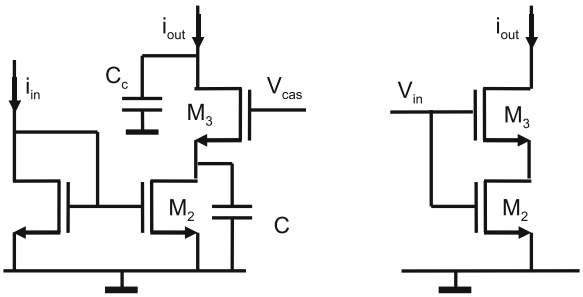


Fig. 2.84 *Left*: The cascode circuit. *Right*: Poor man's cascode

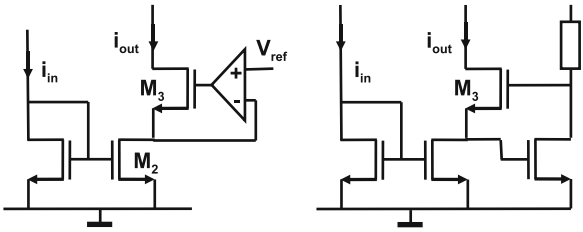


Fig. 2.85 The regulated cascode (gain boost [60, 61]) and a simple realization

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, with λ as the static feedback factor defined in Eq. 2.107. The current variation in this transistor is then $\lambda_3\times$ lower. Adding a cascode transistor increases the output impedance r_2 of the current copier M_2 to roughly $r_{(ds,2)}(1 + 1/\lambda_3)$ for the combination of both transistors. A more rigorous analysis gives the overall output impedance:

$$\frac{v_{\text{out}}}{i_{\text{out}}} = \frac{g_{m,3} r_{ds,3} r_{ds,2} + r_{ds,3} + s r_{ds,3} r_{ds,2} C + r_{ds,2}}{s r_{ds,2} C + 1} \tag{2.195}$$

$$\approx \frac{(g_{m,3} r_{ds,3} + 1) r_{ds,2}}{s r_{ds,2} C + 1} = g_{m,3} r_{ds,3} r_{ds,2}\big|_{s=0} = r_{ds,2}/\lambda_3, \tag{2.196}$$

where $r_{ds,2}/\lambda_3$ is the dominant factor. The output impedance of M_2 is amplified by the gain of M_3. 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.

A poor man's cascode or split-length MOS transistor is shown in Fig. 2.84 (right). The lower transistor is in the linear region and acts as a degeneration resistor, thereby improving the output impedance. A second application of this circuit is the use of

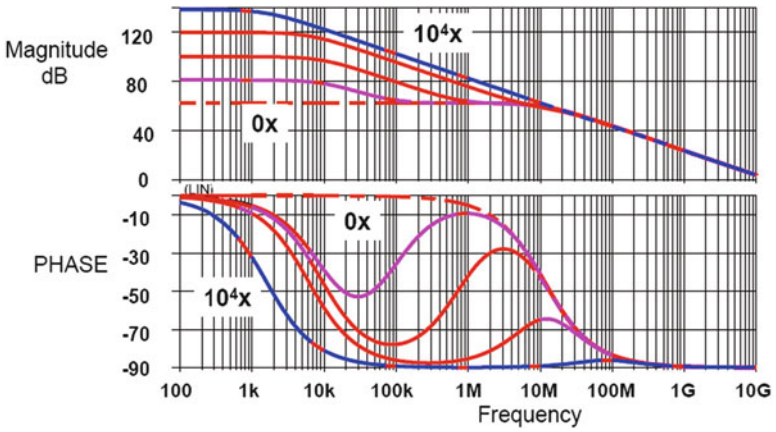


Fig. 2.86 Magnitude and phase of the output impedance of a 90-nm designed cascode stage (*dotted line*) and of a regulated cascode with amplification factors of: 10 ,10^2,10^3,10^4 and an internal pole at 10 kHz. The gain in the feedback path translates in more gain at the output node of the circuit

the intermediate node as a low-ohmic current input. However, as the lower MOS transistor is in its linear region, some signal current will be lost.

Another approach for increasing the output impedance is the regulated-cascode or gain-boost technique [60] (see Fig. 2.85). 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 2.85 (right) shows a simple implementation of the amplifier where the threshold voltage of the amplifier MOS serves as the reference voltage. Figure 2.86 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, 1,000 and 10,000 times improve the overall gain. The lower amplification curves allow to identify the complex output impedance of this circuit. The transfer description is

$$\frac{v_{\text{out}}}{i_{\text{out}}} = \frac{g_{m,3}r_{L,3}r_L(A+1+s\tau_{fb})}{1+s\tau_{fb}+sg_{m,3}r_{L,3}r_LC_3(A+1+s\tau_{fb})}. \tag{2.197}$$

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:

$$\frac{v_{\text{out}}}{i_{\text{out}}} \approx \frac{g_{m,3}r_{L,3}r_L(A+1)}{1+s\tau_{fb}}. \tag{2.198}$$

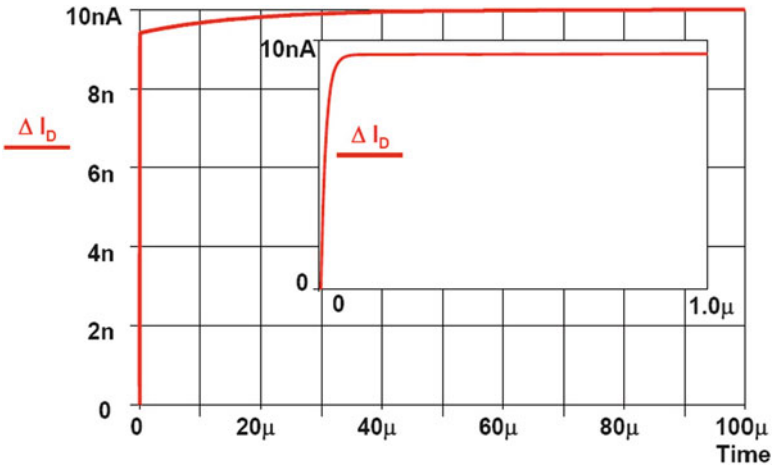


Fig. 2.87 Example of the response of a doublet in a regulated cascode. The plot shows the steep initial reaction to a change (see also the inset), followed by a much slower final settling. This regulated cascode circuit corresponds to Fig. 2.86 with an amplification factor of 100 and a time constant in the feedback loop of 16 µs

At the moment this roll-off has reached its unity-gain frequency ($\omega\tau_{fb}/A = 1$), the zero flattens the curve to the original cascode impedance:

$$\frac{v_{\text{out}}}{i_{\text{out}}} \approx \frac{g_{m,3}r_{L,3}r_L(A+1)(1+s\tau_{fb}/(A+1))}{s\tau_{fb}} \approx g_{m,3}r_{L,3}r_L. \tag{2.199}$$

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:

$$\frac{v_{\text{out}}}{i_{\text{out}}} = \frac{g_{m,3}r_{L,3}r_L}{sg_{m,3}r_{L,3}r_LC_3} = \frac{1}{sC_3}. \tag{2.200}$$

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 Eq. 2.197. 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 Eq. 2.197 can be rewritten in the form of a formal pole-zero doublet:

$$H(s) = \frac{1+s\tau_z}{(1+s\tau)(1+s\tau_p)}, \tag{2.201}$$

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

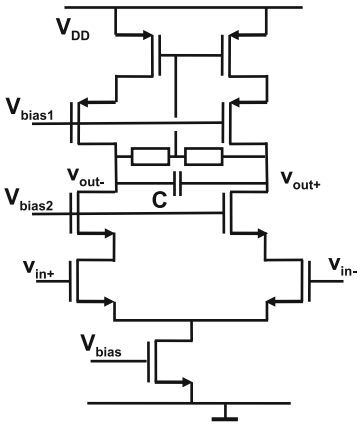


Fig. 2.88 The telescopic fully differential amplifier. The two resistors provide the common-mode feedback path

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 decomposition of the above expression in case of a step wise current excitation leads to two exponential decay terms [62]:

$$i_{\text{out}}(t) = i_{\text{step}} \left(1 - \frac{\tau - \tau_z}{\tau - \tau_p} \mathrm{e}^{-t/\tau} - \frac{\tau_z - \tau_p}{\tau - \tau_p} \mathrm{e}^{-t/\tau_p} \right). \tag{2.202}$$

In Fig. 2.87 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 2.88 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 UGBW is at the frequency where the input voltage v_{in} is equal to the output voltage over the capacitor, resulting in

$$f_{\text{UGBW}} = \frac{g_m}{2\pi C}. \tag{2.203}$$

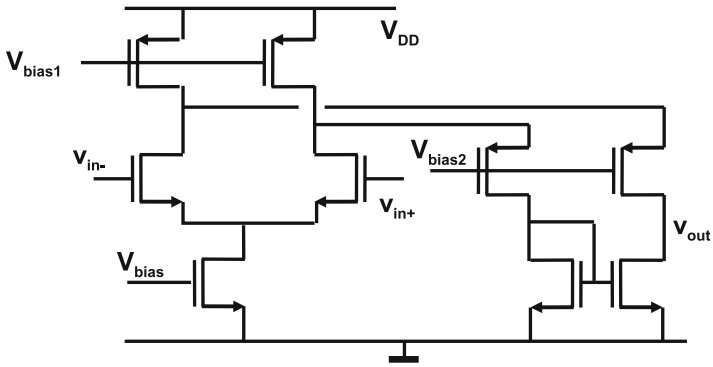


Fig. 2.89 The folded-cascode amplifier

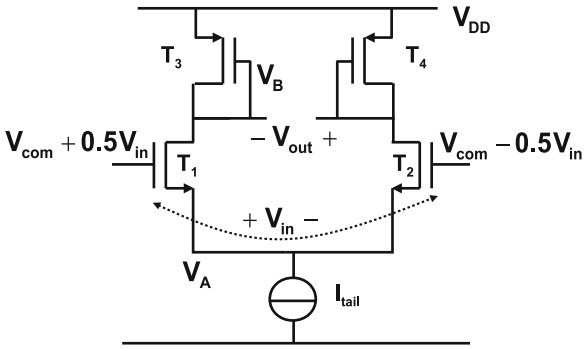


Fig. 2.90 A simple amplifier consisting of a differential stage and a diode-connected load

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 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 2.89 shows the circuit. The topmost 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.

Example 2.21. Figure 2.90 shows a simple amplifier stage. Two NMOS input transistors T_1 and T_2 amplify the input signal to two loading PMOS transistors T_3 and T_4. Define the gain.

Solution. First the large-signal condition is specified: the input signal is considered negligible and the tail current distributes equally over both branches. The DC bias voltages are imposed by the lowest impedances on each node. The transconductance of the input transistors sets the voltage V_A for the node connecting the current source and the input transistors that

$$0.5I_{\text{tail}} = \frac{W_1 C_{\text{ox}} \mu_n}{2L_1}(V_G - V_S - V_{\text{T}})^2 \qquad V_A = V_{\text{com}} - V_{\text{T}} - \sqrt{\frac{I_{\text{tail}} L_1}{W_1 C_{\text{ox}} \mu_n}}.$$

As T_1 equals T_2 the same result is achieved for T_2. A circuit simulator will also incorporate the effect of the current source output impedance. Here that is ignored. The DC bias point V_B for the nodes connecting the NMOS and PMOS transistors is found by looking at the lowest impedance. The PMOS transistor is connected such that its transconductance will dominate over the output impedance of the NMOS transistor, so

$$0.5I_{\text{tail}} = \frac{W_3 C_{\text{ox}} \mu_p}{2L_3}(V_G - V_S - V_{\text{T}})^2 \qquad V_B = V_{\text{DD}} - V_{\text{T}} - \sqrt{\frac{I_{\text{tail}} L_3}{W_3 C_{\text{ox}} \mu_p}}.$$

For each transistor the transconductance can be specified:

$$g_{m1} = g_{m2} = \sqrt{\frac{I_{\text{tail}} W_1 C_{\text{ox}} \mu_n}{L_1}} \qquad g_{m3} = \sqrt{\frac{I_{\text{tail}} W_3 C_{\text{ox}} \mu_p}{L_3}}.$$

In small-signal operation the simplification is used that these values remain constant for every value of the signal. The input differential signal is equally divided over both input transistors, causing small-signal currents in T_1 and T_2:

$$i_1 = g_{m1} \times 0.5 v_{\text{in}} \qquad i_2 = g_{m2} \times -0.5 v_{\text{in}}.$$

Each of these small-signal currents generates a small-signal voltage over the transconductance of T_3 and T_4. The difference between these two nodes is v_{out}:

$$v_{\text{out}} = \frac{i_2}{g_{m4}} - \frac{i_1}{g_{m3}} = \frac{g_{m1} v_{\text{in}}}{g_{m3}} \qquad A_v = \frac{v_{\text{out}}}{v_{\text{in}}} = \frac{g_{m1}}{g_{m3}}.$$

So the gain of this stage equals the ratio of the two transconductances. The same formula applies if the PMOS transistors are replaced by NMOS transistors with their gates connected to the power supply.

2.7.11 *Miller Amplifier*

Figure 2.91 shows a simple two-stage "Miller amplifier" [63–66]. 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).

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 [67].

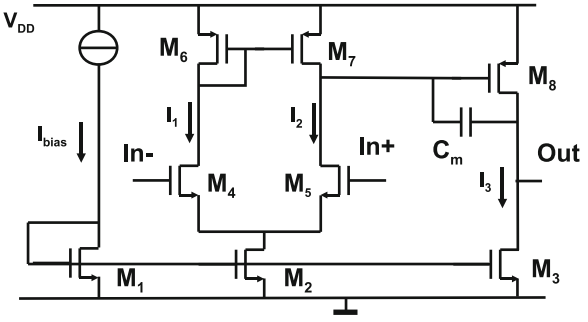


Fig. 2.91 The Miller opamp

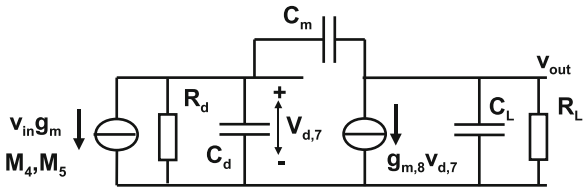


Fig. 2.92 The Miller opamp small-signal equivalent circuit

The current mirror and the output driver carry signals and a near[31]-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. 2.92 a small-signal equivalent circuit is shown. The input pair M_4, M_5 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. 2.92 with the Miller capacitance gives

$$g_m v_{\text{in}} + \frac{v_d}{Z_d} - \frac{v_{\text{out}} - v_d}{Z_m} = 0$$

$$g_{m,8} v_d + \frac{v_{\text{out}} - v_d}{Z_m} + \frac{v_{\text{out}}}{Z_L} = 0.$$

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

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{g_m Z_d Z_L (g_{m,8} Z_m - 1)}{Z_d Z_L g_{m,8} + Z_L + Z_d + Z_m}. \tag{2.204}$$

For frequencies where Z_m is infinitely large the equation reduces to the simple dual pole transfer, and for DC to the amplification factor,

$$\frac{v_{\text{out}}}{v_{\text{in}}} = g_m Z_d Z_L g_{m,8} \rightarrow A_{\text{DC}} = g_m R_d g_{m,8} R_L. \tag{2.205}$$

The frequency behavior defined by the two poles is shown as the upper lines in the Bode plots of Fig. 2.93. 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 $A_{\text{DC}} = 1/\lambda_{M_7}\lambda_{M_8}$ 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 $Z_m = 1/j\omega C_m$:

[31] Compare the current factor versus gate length in Fig. 2.44.

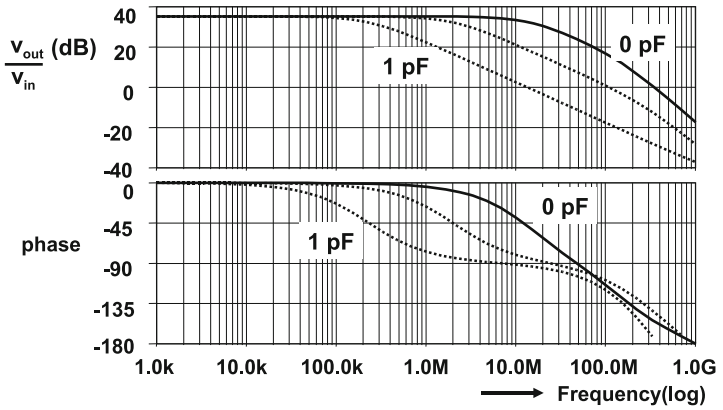


Fig. 2.93 The amplification and phase behavior of the Miller amplifier for Miller capacitors of 0 (*solid line*), 0.1, and 1 pF.

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{g_m Z_d Z_L (g_{m,8} Z_m - 1)}{Z_d Z_L g_{m,8} + Z_L + Z_d + Z_m}$$
$$\approx \frac{g_m Z_d Z_L (g_{m,8} Z_m - 1)}{Z_d Z_L g_{m,8}} = g_m \left(Z_m - \frac{1}{g_{m,8}} \right) \approx \frac{g_m}{j\omega C_m}.$$

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

$$\text{UGBW} = \frac{g_m}{2\pi C_m}. \tag{2.206}$$

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." Although many rules of thumb exist on the choice of the C_m value, it makes sense to derive C_m from the observation that the UGBW from the Miller compensation should be considerably smaller than unity-gain frequency

of the output pole:

$$\text{UGBW} = \frac{g_m}{2\pi C_m} = 0.25 \times \frac{g_{\text{out}}}{2\pi C_{\text{out}}} \tag{2.207}$$

with C_{out} the output load. The factor 0.25 is on the safe side.

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. In a folded-cascode opamp the right-half-plane zero is eliminated by feeding the return current from the Miller capacitor in the cascode stage (see Fig. 7.46) [68].

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. 2.5.3). 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:

$$I_3 T_{\text{slew}} = V_{pp}(C_L + C_M + C_{\text{parasitic}}), \tag{2.208}$$

where T_{slew} is the time allowed to build up a signal of the amplitude V_{pp} over the total capacitive load. $T_{\text{slew}} \approx 1/(2\text{BW})$ 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:

$$\text{UGBW} = \frac{g_m}{2\pi C_M} = \frac{1}{2\pi C_M}\sqrt{\frac{2I_{1,2}W\beta_\square}{L}}. \tag{2.209}$$

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. 2.57). 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 $V_{\text{GS}} - V_{\text{T}}$ 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 dimension-dependent parameter variations. For example, if a transistor must be designed with twice the current of a device with dimensions $W/L = 10/0.4\,\mu\text{m}$ than the choice is between a $W/L = 10/0.2\,\mu\text{m}$ device, a $W/L = 20/0.4\,\mu\text{m}$ device or two $W/L = 10/0.4\,\mu\text{m}$ devices in parallel. If all dimensions turn out 0.1 µm larger that will result in a (10/0.5) versus (10/0.3) or 17% error for the first option, a (20.1/0.5) versus (10.1/0.5) gives 0.5% error for the second option, and (10.1/0.5) versus 2×(10.1/0.5) thus no error for the last option.
  The length is preferably 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 one standard width and multiplying the number of unit transistors. Integer ratios of unit transistors are the optimum design strategy to obtain best matching.
- 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 and mismatch.
- For signal-amplifying devices the drive voltage of transistors $(V_{GS} - V_T)$ is chosen at approximately 200 mV as this gives the largest transconductance. A value of around 100 mV is for high temperatures too close to subthreshold. Current sources are biased at larger drive voltages to reduce their transconductance. The maximum drive voltage is limited as it determines the saturation drain voltage below which the device will operate as a controlled resistor.

Example 2.22. Design a Miller opamp as in Fig. 2.91 in a CMOS 0.18 µm process that will drive a 0.5 V amplitude ($1\,V_{pp}$) in unity-gain feedback on a capacitor of 2 pF to a frequency of 5 MHz with −60 dB THD and minimum current.

Solution. From the above design considerations, $W_4 = W_5, W_6 = W_7$, $L_1 = L_2 = L_3, L_6 = L_7 = L_8$ and $W_3/W_2 = W_8/(W_6 + W_7)$.

In this design the transistors M_1, M_2, and M_3 are current sources that are not in the amplification path. For a 0.18 µm process the choice $L_1 = L_2 = L_3 = 0.3\,\mu m$ is a compromise between obtaining a high output impedance and the area of the transistors. The PMOS transistors do amplify and their gate size acts as a load capacitor, so these gate lengths are chosen close to the technology limit: $L_6 = L_7 = L_8 = 0.2\,\mu m$.

The output specification is the starting point. The slew-rate sets the current in the output stage. A simple triangle approximation of the maximum output sine wave gives (see Eq. 2.208)

$$I_{\text{output}} = \frac{1\,V \times 2\,pF}{10\,ns} = 200\,\mu A.$$

Setting the DC current at 250 µA and the drive voltage at 0.2 V gives $W_3/L_3 = 20, W_8/L_8 = 150$. The opamp of Fig. 2.91 in 0.18 µm CMOS: $I_{\text{bias}} = 50\,\mu A$, $W_1/L_1 = W_2/L_2 = 10/0.3, W_3/L_3 = 50/0.3, W_4/L_4 = W_5/L_5 = 30/0.18, W_6/L_6 = W_7/L_7 = 15/0.2, W_8/L_8 = 150/0.2$ all dimensions in µm.

With $C_m = 0.5$–1 pF, a UGBW of 200 MHz is possible, with $\sigma_{V_{\text{in}}} = 3\,mV$.

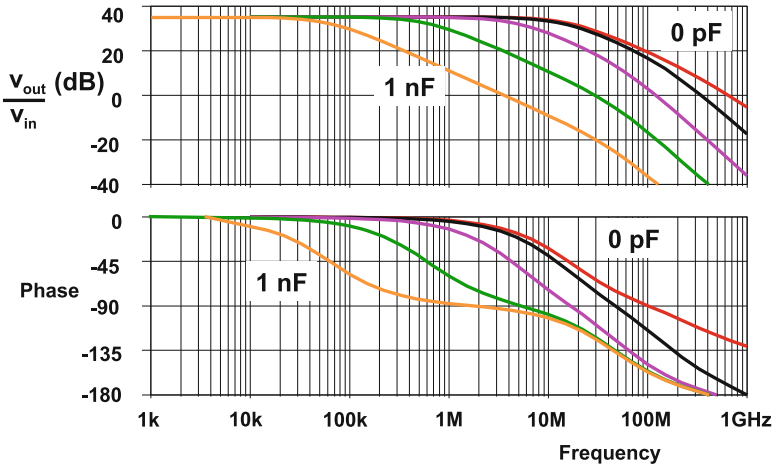


Fig. 2.94 The amplification and phase behavior of the amplifier for load capacitors of 0, 1 pF, 10 pF, 100 pF, and 1 nF. A larger load capacitor creates a low-frequency character

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 2.94 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 simulation increasing from 0 to 1 nF. Without output load the amplifier's 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. 2.6.6 the general theory of feedback systems was summarized. In electronic circuits signals are represented by currents or voltages. Both voltages and currents are used as inputs or outputs of a circuit.

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

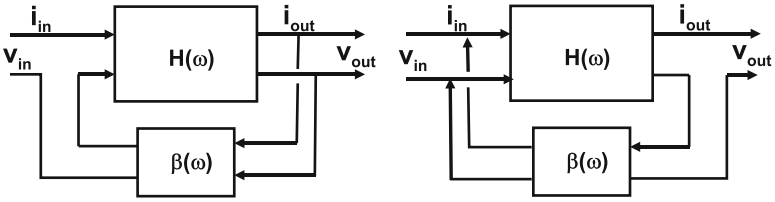


Fig. 2.95 A series–shunt and a shunt series configuration

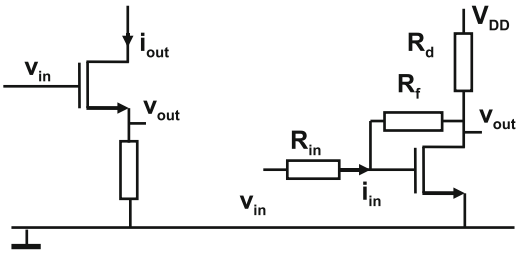


Fig. 2.96 An example of a series–series circuit and a shunt–shunt circuit

Table 2.25 lists the four transfer modes. The feedback classification as used in Fig. 2.95 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 [54] 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. For example, feeding the circuit topology in Fig. 2.95 (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 will be further reduced by a shunt feedback; a low-ohmic output voltage node will show a lower impedance through a shunt feedback.

Figure 2.96 (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 $(v_{\text{in}} - i_{\text{out}}R_s)$ showing the impact of the feedback.

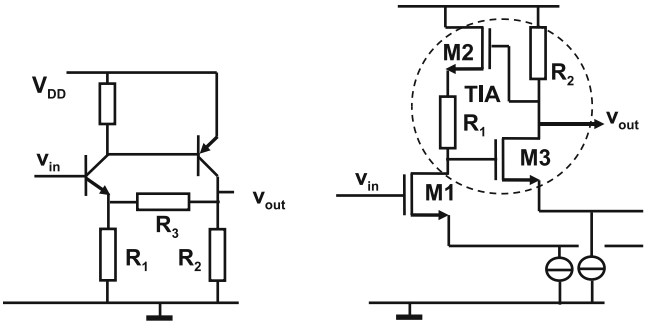


Fig. 2.97 Two circuits using the Cherry-Hooper technique

The circuit in Fig. 2.96 (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 [69] 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. 2.97 (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 [70]. A popular circuit (see Fig. 2.97 (right)) is built with an input transconductance stage and an output formed by a feedback transimpedance amplifier (TIA). Again the similarities with the Wilson mirror (Fig. 2.83) and the gain boost technique (Fig. 2.85 (right)) are instructive.

2.7.15 Bias Circuits

Figure 2.98 shows four schemes to provide bias voltages and currents for circuits. In the upper left circuit schematic a resistor determines the current:

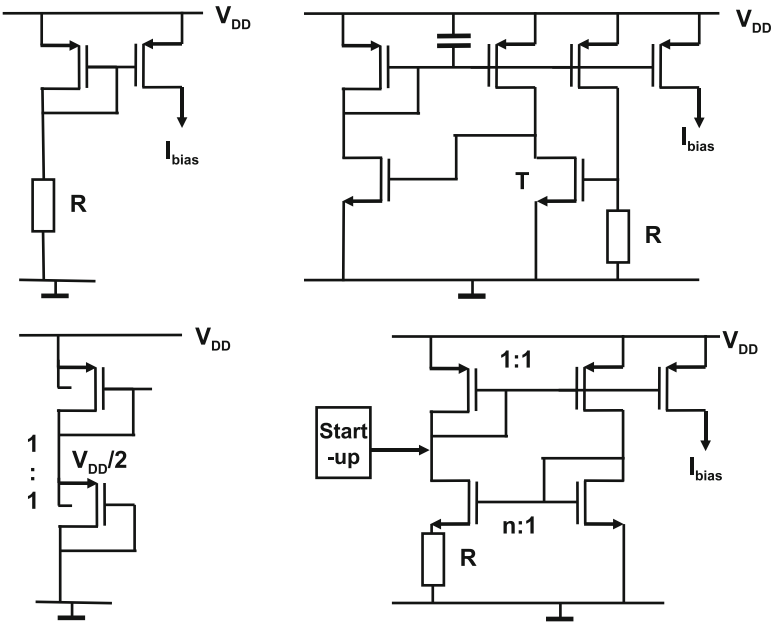


Fig. 2.98 Four bias circuits, *top*: simple resistor-based current and a feedback circuit that will set the current to V_T/R. *Bottom*: a voltage divider and a beta-resistor circuit

$$I = \frac{V_{DD} - V_T}{R}. \tag{2.210}$$

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. 2.98 (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. 2.98 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:

$$I = \frac{V_T}{R}. \tag{2.211}$$

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. 2.98 (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

$$I = \sqrt{\frac{2I}{R^2\beta}}\left(1 - \frac{1}{\sqrt{n}}\right) \tag{2.212}$$

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.

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.

Example 2.23. In a 0.18 µm technology the NMOS transistor current factor is $\beta_\square = 300\,\mu\text{A}/\text{V}^2$. What current is generated with NMOS transistor W/L ratios of 2/0.4 and 8/0.4 ($n = 4$) and $R = 10\,\text{k}\Omega$?

Solution. Assume a start-up circuit prevents $I = 0$ to occur. Now $\sqrt{I}$ is a valid divisor, yielding $I = \frac{2}{R^2\beta}(1 - \frac{1}{\sqrt{n}})^2 = 3\,\mu\text{A}$.

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. These Barkhausen oscillation conditions are described by Eq. 2.142. Lower gain will create extinction of the oscillation; higher gain will create saturation. The design of an oscillator has to meet these two requirements. 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

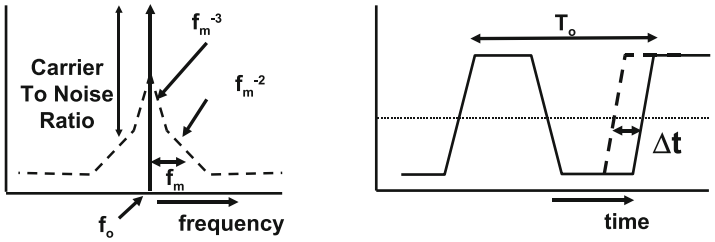


Fig. 2.99 *Left*: the frequency spectrum of an oscillator, *right*: time jitter

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:

$$Q = 2\pi \frac{\text{Energy stored}}{\text{Energy dissipated/cycle}}. \tag{2.213}$$

Consider as an example a sinusoidal current $\hat{i}\sin 2\pi f_0 t$ flowing in an oscillating RLC tank (Fig. 2.64). The maximum energy stored in the inductor is $\hat{i}^2 L/2$; the energy dissipated per cycle is $\hat{i}^2 R/2f_0$. Substitution of these quantities results in Eq. 2.160.

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. 2.99 (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:

$$\theta(t) = \int_{\tau=0}^{\tau=t} \omega_0 + k(\tau)\mathrm{d}\tau \tag{2.214}$$

as the argument of a sine wave description:

$$v_0(t) = V_A \cos(\theta(t)) = V_A \cos\left(\omega_0 t + \int_{\tau=0}^{\tau=t} k(\tau)\mathrm{d}\tau\right). \tag{2.215}$$

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,

$$v_0(t) = V_A \cos(\omega_0 t)\cos(m\sin(\omega_m t)) - V_A \sin(\omega_0 t)\sin(m\sin(\omega_m t)). \tag{2.216}$$

The nested sine and cosine terms can be rewritten with Bessel functions:

$$\begin{aligned} v_0(t) = V_A [&J_0(m)\cos(\omega_0 t) - J_1(m)(\cos(\omega_0 t - \omega_m t) - \cos(\omega_0 t + \omega_m t)) \\ &+ J_2(m)(\cos(2\omega_0 t - 2\omega_m t) - \cos(2\omega_0 t + 2\omega_m t)) - \ldots], \end{aligned}$$

where the Bessel function is defined as

$$J_n(m) = \frac{1}{\pi}\int_{\theta=0}^{\theta=\pi} \cos(m\sin(\theta) - n\theta)\mathrm{d}\theta. \tag{2.217}$$

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:

$$v_0(t) = V_A \left(\cos(\omega_0 t) - \frac{m}{2}\cos(\omega_0 t - \omega_m t) + \frac{m}{2}\cos(\omega_0 t + \omega_m t)\right). \tag{2.218}$$

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:

$$\frac{\text{Power of side frequencies}}{\text{Power of carrier}} = (m/2)^2 + (m/2)^2 = \frac{k_0^2}{2\omega_m^2}. \tag{2.219}$$

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 [71] 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 [71]:

$$S_{ff}(\omega) = \frac{2FkT}{P_0} \frac{\omega_0^2}{(2Q)^2(\omega - \omega_0)^2}. \tag{2.220}$$

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 oscillator allows a low phase noise.

Figure 2.99 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 $(\omega - \omega_0)$ 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,

$$\frac{\Delta t}{T_0} = \frac{\Delta\theta}{2\pi}. \tag{2.221}$$

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:

$$\left(\frac{\sigma_{t,\mathrm{rms}}}{T_0}\right)^2 = \frac{2\int_{f_{\mathrm{low}}}^{f_{\mathrm{high}}} S_{ff}(f)\mathrm{d}f}{(2\pi)^2}. \tag{2.222}$$

The integration cannot start at $\omega = \omega_0\,\mathrm{Hz}$ 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

$$\sigma_{t,\mathrm{rms}} \approx \sqrt{\frac{(\Delta f) S_{ff}(f_0 - \Delta f)}{2\pi^2 f_0^2}}. \tag{2.223}$$

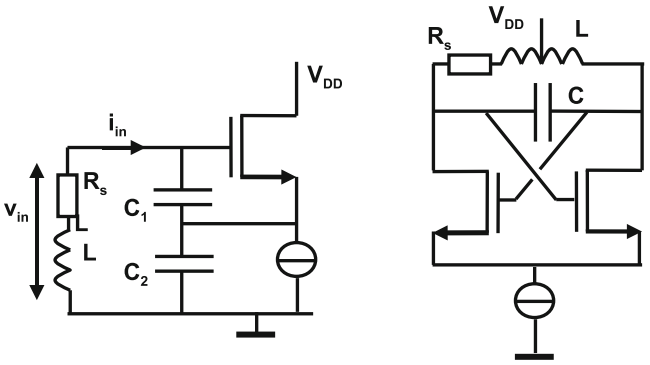


Fig. 2.100 A Colpitts oscillator and a cross-coupled LC oscillator

With $f_0 = 1\,\text{GHz}$, $\Delta f = 1\,\text{MHz}$, and $S_{ff}(999\,\text{MHz}) = 10^{-12}$ equivalent to $\mathcal{L}(999\,\text{MHz}) = -120\,\text{dBc/Hz}$, a jitter of 0.2 ps is found. Taking the bands beyond $\Delta f = 100\,\text{kHz}$ 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 [71–74].

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:

$$\text{Oscillator number} = 10\log(S_{ff}(\omega_m)) + 20\log\left(\frac{\omega_m}{\omega_0}\right) = \text{CNR}(\omega_m) - 10\log\left(\frac{\omega_0^2}{\omega_m^2}\right)$$

$$\text{Figure of Merit} = \text{Oscillator number} + 10\log\left(\frac{\text{Power}}{1\,\text{mW}}\right). \tag{2.224}$$

A typical LC oscillator shows an oscillator number in the range of -180 to $-190\,\text{dB}$.

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 mW. Good figures of merit are between $-160\,\text{dB}$ for relaxation oscillators and $-180\,\text{dB}$ for high-Q LC oscillators [75, pp. 177–179].

Figure 2.100 (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[32]:

$$\omega_0 = \frac{1}{\sqrt{L\frac{C_1 C_2}{C_1+C_2}}}. \tag{2.225}$$

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:

$$(v_{\text{in}} - v_s) j\omega C_1 + g_m(v_{\text{in}} - v_s) = v_s j\omega C_2. \tag{2.226}$$

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

$$\frac{v_{\text{in}}}{i_{\text{in}}} = \frac{1}{j\omega C_1} + \frac{1}{j\omega C_2} + \frac{g_m}{j\omega C_1 j\omega C_2}. \tag{2.227}$$

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 [76].

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 nonlinear phenomena introduce specification loss. For example, 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:

[32] Ignoring the small effect that R_s has on the oscillation frequency.

Table 2.26 Parameters for the equivalent circuit of a quartz crystal

$C_0 = 4\,\text{pF}$	$C_1 = 16\,\text{fF}$
$H_1 = 5\,\text{mH}$	$R_1 = 50\,\Omega$
$f_s = 18\,\text{MHz}$	$Q = 10.000$

- 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 $C_h = C_1C_2/(C_1 + C_2)$, which consumes considerably less chip area than $(C_1 + C_2)$. This fact not only can be an advantage because of less area, but it also increases the sensitivity for parasitics.
- Figure 2.100 (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 100–1,000 range. Higher-quality factors require to use mechanical resonators like a quartz crystal as a resonator. Figure 2.101 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. 2.101 (right)). C_1 and L_1 represent 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 antiresonance frequency) the capacitor C_0 is in series with C_1, forming slightly smaller capacitor than C_1. These two resonance frequencies and the quality factor are specified as

$$f_s = \frac{1}{2\pi}\sqrt{\frac{1}{L_1C_1}} \qquad f_p = \frac{1}{2\pi}\sqrt{\frac{C_0 + C_1}{L_1C_1C_0}} \qquad Q = \frac{1}{2\pi f_s R_1 C_1}. \tag{2.228}$$

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.

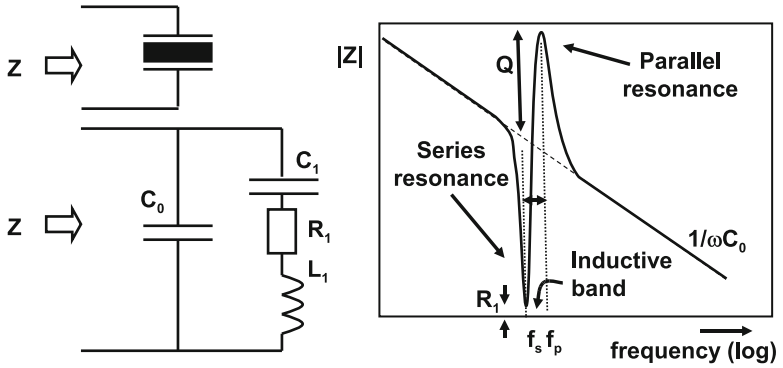


Fig. 2.101 The electrical equivalent circuit for crystal contains a series branch shunted with C_0. *Right*: A schematic view of the impedance of a crystal (not to proportion!)

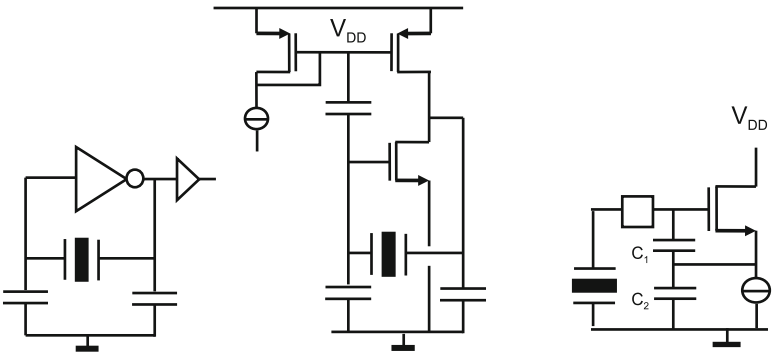


Fig. 2.102 A Pierce crystal oscillator, a push–pull arrangement [78], and a one-pin oscillator [79]

The stable mechanical construction allows to create high-quality oscillators. Comparison of Fig. 2.102 [77–79] 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. 2.102).

The capacitors and crystal of the basic Colpitts/Pierce oscillator can be rearranged. Moving one side of the crystal to the ground node allows to use only one additional connection (pin of a package) and the crystal oscillator of Fig. 2.102 (right) is obtained.

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 layout tricks improve the quality factor somewhat.

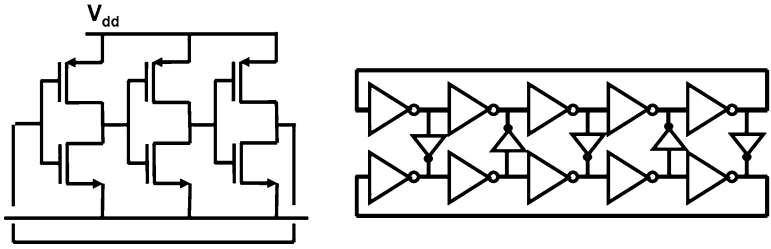


Fig. 2.103 A inverter ring as an oscillator and pseudo-differential ring oscillator

- 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 2.103 (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.[33] This inverter delay is determined by the (dis)charging time of the intermediate node by the transistors:

$$t_{d,\mathrm{inv}} = \alpha_{\mathrm{inv}} \frac{C_{\mathrm{node}} V_{\mathrm{DD}}}{I_{\mathrm{DS,sat}}}, \tag{2.229}$$

where $\alpha_{\mathrm{inv}} \approx 0.5$ depending on, e.g., the NMOS/PMOS ratio. If $I_{d,\mathrm{sat}} \propto (V_{\mathrm{DD}} - V_{\mathrm{T}})^2$ 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 (PLLs). In order to improve the performance of the oscillator a pseudo-differential approach is applied as is seen in Fig. 2.103 (right). The two inverter loops run in antiphase. The antiphase behavior is guaranteed by the small coupling inverters between the taps of the lines. The current consumption is relatively constant.

A multivibrator of Fig. 2.104 (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

[33] People advertising the virtues of technologies reverse this definition: the gate delay is the inverse fraction of the measured oscillation frequency.

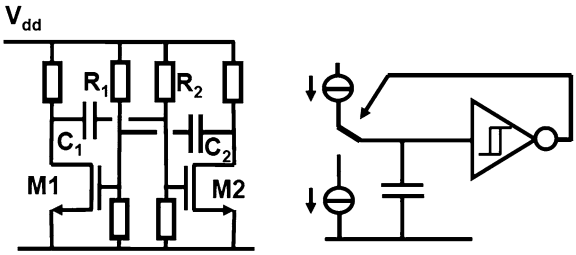


Fig. 2.104 A traditional astable multivibrator and a current source multivibrator

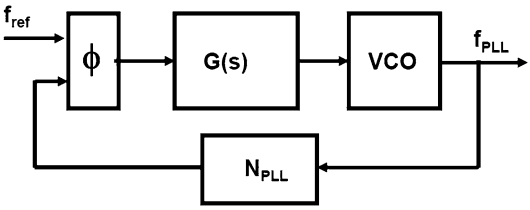


Fig. 2.105 A phase-locked loop circuit

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:

$$f = \alpha_{\text{mul}} \frac{1}{R_1 C_2 + R_2 C_1} \tag{2.230}$$

with $\alpha_{\text{mul}} \approx 0.7$. Recently multivibrators that are more adopted to integrated circuit design as in Fig. 2.104 (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 [80] 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% [81].

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 is needed. A circuit that locks the oscillation frequency to a reference frequency is a "PLL" . Figure 2.105 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 $f_{\text{out}} = N_{\text{PLL}} f_{\text{ref}}$.

The transfer of the phase demodulator can be modeled as $v_d = K_d(\theta_{\text{ref}} - \theta_{\text{out}}/N_{\text{PLL}})$. The transfer of the filter is $H(s)$. The voltage-controlled oscillator will generate an output frequency proportional to the input voltage $f_{\text{out}} = K_v v_{osc}$. The phase is the time integral of the frequency $f_{\text{out}}(s) = s\theta_{\text{out}}(s)$; therefore an additional $1/s$ factor must be added in the transfer for the oscillator. This results in an overall transfer function of

$$\frac{\theta_{\text{out}}(s)}{\theta_{\text{ref}}(s)} = \frac{K_v K_d H(s)/s}{1 + K_v K_d H(s)/N_{\text{PLL}} s}. \tag{2.231}$$

There are some noise sources in a PLL: the noise of the reference signal, the noise contributed by the phase detector and the divider, 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. In [82] noise sources are analyzed and the noise performance is improved by replacing the divider by a subsample mechanism.

Exercises

2.1. If a signal transfer function equals $y = 2x + 0.01x^3$ and $x(t) = 0.5\sin(\omega t)$, what is the THD in dB?

2.2. Show that for a smooth form of distortion (like in the previous exercise) the change of the second-order distortion amplitude is twice (in dB) the change in fundamental amplitude and that the third-order component changes three times the change in the fundamental.

2.3. Change the function in Example 2.7 on p. 29 in $I = K_0(V_g - V_{\text{T}})^2$, and calculate the mean and variance.

2.4. In Example 2.8 on p. 30, change the two parallel resistors in two parallel conductances $G_1 = 1/R_1$; if the relative standard deviation of the conductances equals the relative standard deviation of the resistors ($\sigma_{R_1}/R_1 = \sigma_{G_1}/G_1$), show that the same result is obtained.

2.5. Show in Example 2.10 on p. 39 that using noise voltage sources requires to weigh the contribution of each noise voltage source with the applicable resistor ratio and gives the same result.

2.6. Where is in Example 2.11 on p. 52 the lost energy?

2.7. An RF amplifier is constructed by replacing the resistors in Fig. 2.80 (upper left) by coils. Derive a DC and RF amplification description. Derive the impedance

seen on the gate when the gate capacitance is included. What is a particular additional effect in this circuit when the voltages on the source and drain are observed?

2.8. A small closed box has two electrical connections. The box may contain a voltage source of 1 V in series with a resistor of 1 Ω or a current source of 1 A with a parallel resistor of 1 Ω. According to the Norton theorem both are electrically equivalent. Still you can tell what is in the box without opening it. How?

2.9. Show in Example 2.21 on p. 138 that the same formula for the gain is obtained for the NMOS loaded circuit. What are the differences between the PMOS and NMOS loaded circuits?

2.10. Derive, for Example, 2.21 on p. 138 the voltage gain if the PMOS load transistors are replaced by current sources.

2.11. Compare the different unity-gain bandwidths for the various opamp topologies. Which opamp will give the best UGBW for a fixed amount of current?

2.12. Compare the different DC gains for the various opamp topologies. Which opamp will give the best DC gain for the lowest supply voltage?

Chapter 3
Sampling

Abstract Sampling is one of the main processes in an analog-to-digital converter. The sampling theory is examined and the crucial elements are extensively discussed. The relation with other techniques, such as modulation and sampling of noise, is described.

The second section discusses the design of time-discrete filters. These filters form an important building block in several conversion structures especially in sigma–delta conversion. Down-sample filters transform the bit-stream format into a more usable pulse-code format. The essential properties of finite impulse response (FIR) and infinite impulse response (IIR) filters are reviewed.

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:

$$t = \frac{n}{f_s} = nT_s, \quad n = -\infty, \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots, \infty. \tag{3.1}$$

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., [83, 84]. 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.

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_3,

© Springer Science+Business Media, LLC 2013

Sampling transforms a time-continuous signal in a time-discrete signal and can be applied on various signals. Most common is sampling of analog time-continuous signals into analog time-discrete signals. Also time-continuous digital signals (like pulse-width modulated signals) can be sampled.

The mathematical description of the sampling 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:

$$\int_{t=-\infty}^{\infty} f(t)\delta(t-t_0)\,\mathrm{d}t = f(t_0). \tag{3.2}$$

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":

$$\delta(t) = \begin{cases} 0, & -\infty < t < 0 \\ \frac{1}{\varepsilon}, & 0 < t < \varepsilon \\ 0, & \varepsilon < t < \infty \end{cases} \Rightarrow \int_{t=-\infty}^{\infty} \delta(t)\,\mathrm{d}t = 1 \tag{3.3}$$

with $\varepsilon \to 0$.

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 discrete Fourier series. This discrete Fourier transform (DFT) 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 yields

$$\sum_{n=-\infty}^{n=\infty} \delta(t - nT_s) = \sum_{k=-\infty}^{\infty} C_k \mathrm{e}^{jk2\pi f_s t}. \tag{3.4}$$

The coefficients C_k of the resulting discrete Fourier series are found by integrating over a period see Eq. 2.5. The additional factor 2 in Eq. 2.5 is due to the single sided range.

$$C_k = \frac{1}{T_s} \int_{t=-T_s/2}^{T_s/2} \sum_{n=-\infty}^{n=\infty} \delta(t - nT_s)\mathrm{e}^{-jk2\pi f_s t}\,\mathrm{d}t. \tag{3.5}$$

[1] Strange in the sense that many normal mathematical operations cannot be performed, e.g., $\delta^2(t)$ does not exist.

Within the integration interval there is only one active Dirac pulse at $t = 0$, so the complicated formula reduces to

$$C_k = \frac{1}{T_s} \int_{t=-T_s/2}^{T_s/2} \delta(t) \mathrm{e}^{-jk2\pi f_s t} \mathrm{d}t = \frac{1}{T_s} \mathrm{e}^{-jk2\pi f_s \times 0} = \frac{1}{T_s}. \tag{3.6}$$

Now the substitution of C_k results in the mathematical description of the DFT from the sequence of Dirac pulses in the time domain:

$$\sum_{n=-\infty}^{n=\infty} \delta(t - nT_s) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} \mathrm{e}^{jk2\pi f_s t} = \frac{1}{T_s} \int_{f=-\infty}^{\infty} \sum_{k=-\infty}^{k=\infty} \delta(f - kf_s) \mathrm{e}^{jk2\pi f_s t} \mathrm{d}f. \tag{3.7}$$

The last term is the standard inverse Fourier transform with a summation of frequencies as argument. This sum of Dirac functions in the discrete Fourier series is therefore the frequency domain counterpart of the time-domain Dirac series:

$$\sum_{n=-\infty}^{n=\infty} \delta(t - nT_s) \Leftrightarrow \frac{1}{T_s} \sum_{k=-\infty}^{k=\infty} \delta(f - kf_s). \tag{3.8}$$

The infinite sequence of short time pulses corresponds to an infinite sequence of frequency components at multiples of the sampling rate.

A fast Fourier transform (FFT) is a method to compute efficiently a DFT. This method quantizes the signal components on a grid with frequency bins of size $f_{\mathrm{bin}} = 1/T_{\mathrm{meas}}$ (see also Sect. 10.2.2). A time-discrete repetitive signal can be exactly analyzed with a DFT or FFT. If a time-continuous signal is analyzed by means of an FFT algorithm a form of frequency quantizing or discretization takes place, which can cause errors.

In Fig. 3.1 an example is presented[2] of the sampling of a signal $A(t)$ corresponding to $\mathbf{A}(\omega) = \mathbf{A}(2\pi f)$ with a bandwidth from $f = 0\,\mathrm{Hz}$ to $f = \mathrm{BW}$:

$$\mathbf{A}(\omega) = \int_{t=-\infty}^{\infty} A(t) \mathrm{e}^{-j2\pi f t} \mathrm{d}t. \tag{3.9}$$

Mathematically the sampling is performed by multiplying the time-continuous function $A(t)$ of Fig. 3.1a with the sequence of Dirac pulses, resulting in a time-discrete signal Fig. 3.1b. 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:

$$A_s(t) = \sum_{n=-\infty}^{n=\infty} A(t) \times \delta(t - nT_s) \Rightarrow \sum_{n=-\infty}^{n=\infty} A(nT_s). \tag{3.10}$$

[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.

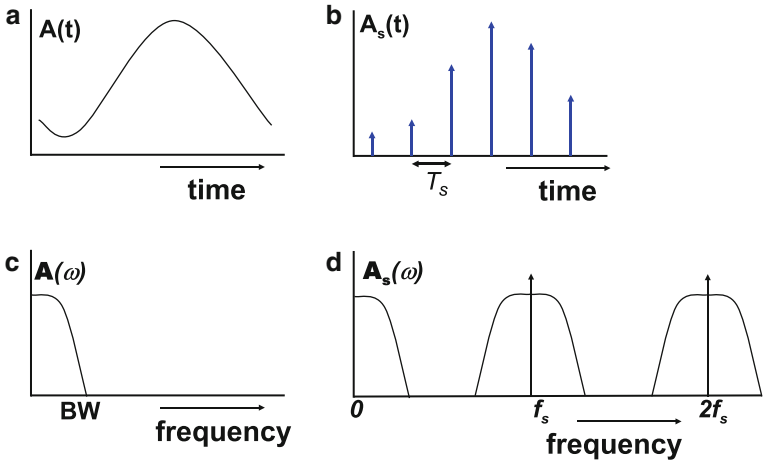


Fig. 3.1 Sampling an analog signal (**a**) in the time-continuous domain results in a series of analog signal samples (**b**) in the frequency domain the time-continuous signal (**c**) is folded around the sampling frequency and its multiples (**d**)

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 $\mathbf{A_s}(\omega)$:

$$\begin{aligned}\mathbf{A_s}(\omega) &= \int_{t=-\infty}^{\infty} \left[\sum_{n=-\infty}^{n=\infty} A(t)\delta(t-nT_s)\right] \mathrm{e}^{-j2\pi ft}\mathrm{d}t \\ &= \int_{t=-\infty}^{\infty} A(t)\frac{1}{T_s}\sum_{k=-\infty}^{\infty} \mathrm{e}^{jk2\pi f_s t}\mathrm{e}^{-j2\pi ft}\mathrm{d}t \\ &= \sum_{k=-\infty}^{\infty}\frac{1}{T_s}\int_{t=-\infty}^{\infty} A(t)\mathrm{e}^{-j2\pi(f-kf_s)t}\mathrm{d}t. \end{aligned} \tag{3.11}$$

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:

$$\mathbf{A_s}(\omega) = \sum_{k=-\infty}^{\infty}\frac{1}{T_s}\mathbf{A}(\omega - k\omega_s) = \sum_{k=-\infty}^{\infty}\frac{1}{T_s}\mathbf{A}(2\pi(f-kf_s)). \tag{3.12}$$

The original 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. Figure 3.1c, d depict the time-continuous signal and the sampled signal in the

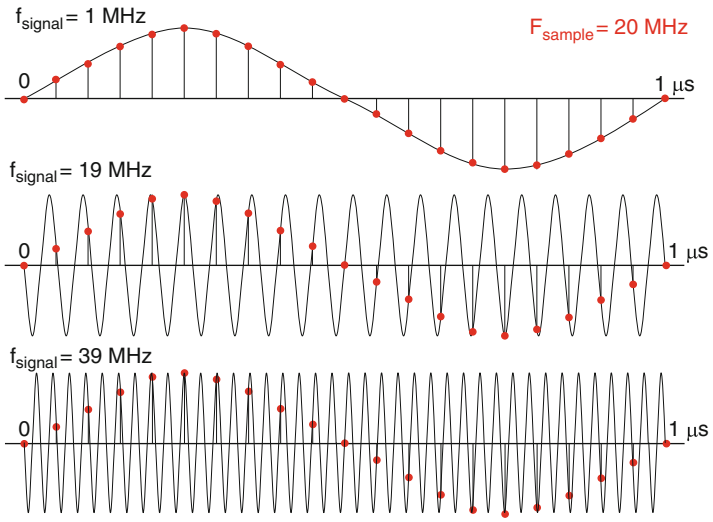


Fig. 3.2 Sampling three time-continuous signals: 1 MHz, 19 MHz, and 39 MHz sine waves result after sampling with 20 Ms/s in the same sampled data sequence (*dots*)

frequency domain. In the frequency domain of the sampled data signal, next to the original signal, also the upper bands are shown.

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 were mutually uncorrelated and could be converted in a physical quantity, there would indeed be a contradiction. However, in the reconstruction from a mathematical sequence of Dirac pulses to a physical quantity, there is an inevitable filtering operation, limiting the energy.

Frequency components in the time-continuous 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).

Every signal in the time-continuous domain is mapped into a base band set of samples. And different signals in the time-continuous domain can have the same representation in the time-discrete domain.

Example 3.1. The input stage of a sampling system is generating distortion. Plot the input signal of 3.3 MHz and its 4 harmonics in the band between 0 and the sampling frequency of 100 Ms/s. Change the sampling rate to 10 Ms/s and plot again.

Solution. Figure 3.3 shows on the left side the sampled spectrum at 100 Ms/s and on the right side the 10 Ms/s variant. Signal components will appear at frequencies: $i \times f_{\text{in}} \pm j \times f_s$, where $i = 1 \ldots$ number of harmonics and $j = 0 \ldots \infty$. Therefore in the last graph there are components as shown in Table 3.1 Again note that within each $f_s/2$ range there are exactly five components, corresponding to each of the original tones.

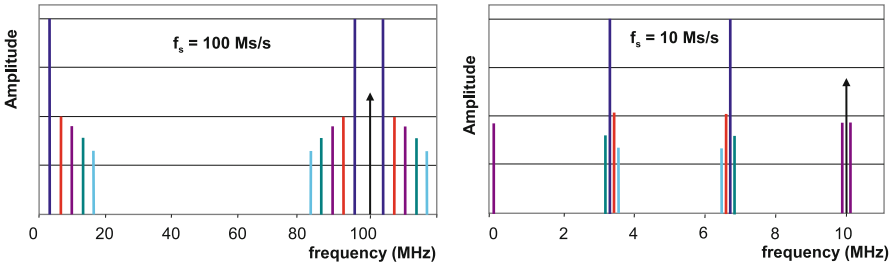


Fig. 3.3 The sample rate in the upper plot is 100 Ms/s and in the lower plot 10 Ms/s

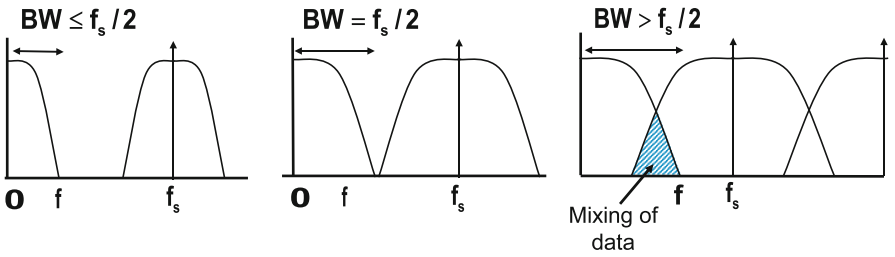


Fig. 3.4 The time-continuous bandwidth can be increased until half of the sample rate is reached

Table 3.1 Frequency components of a distorted 3.3 MHz sinusoid sampled at 10 Ms/s

0.1 MHz	$f_s - 3f_{\text{in}}$
3.2 MHz	$f_s - 4f_{\text{in}}$
3.3 MHz	f_{in}
3.4 MHz	$f_s - 2f_{\text{in}}$
3.5 MHz	$f_s - 5f_{\text{in}}$
6.5 MHz	$2f_s - 5f_{\text{in}}$
6.6 MHz	$2f_{\text{in}}$
6.7 MHz	$f_s - f_{\text{in}}$
6.8 MHz	$2f_s - 4f_{\text{in}}$
9.9 MHz	$3f_{\text{in}}$

3.1.1 Aliasing

If the bandwidth in the time-continuous domain increases, the mirror bands around the multiples of the sample frequency will follow. Figure 3.4 shows that this will lead to overlap of signal bandwidths after sampling. This phenomena is called "aliasing." The closest upper band directly adjacent to the original band is called: "the alias band".

A consequence of sampling is that the maximum useful signal bandwidth that can be handled in the discrete-time domain, is limited to BW $\leq f_s/2$.

Example 3.2. A 10 kHz sine wave is distorted with components at 20, 30, 40, and 50 kHz, and sampled with 44 ks/s. Draw the spectrum.

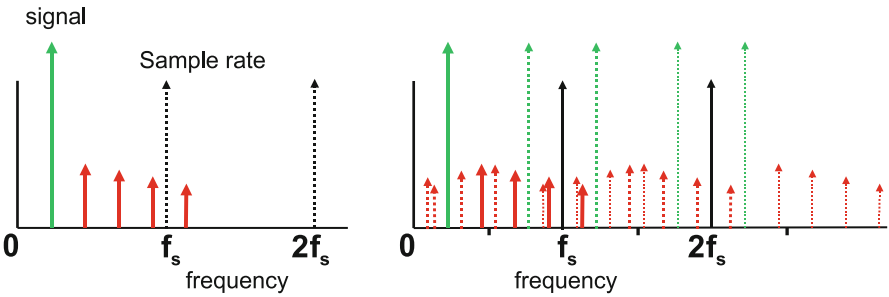


Fig. 3.5 The sample rate in this plot is 44 ks/s operating on a distorted sine wave of 10 kHz

Solution. The left part of Fig. 3.5 shows the input spectrum. The right part shows the result after sampling. The tones from the original spectrum are in full lines, the result of the folding and mirroring around the 44 ks/s sample rate are drawn in dashed lines, and the components originating from the second multiple at 88 ks/s are shown in thinner dashed lines. The 50 kHz component results after sampling in a (-6) kHz frequency. This component is shifted to the positive frequency domain, while keeping in mind that it differs in phase. In every interval of $f_s/2$ width there is exactly one copy of each originating component. So there is a simple check on the correctness and completeness of the spectrum: make sure the number of components exactly matches the number at the input.

3.1.2 SubSampling

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 "subsampling". 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 subsampling are used in communication applications, where sub-sampling is used as demodulation, (see Fig. 3.6). Nondeliberate forms of sub-sampling occur if undesired signals are present in the signal band. Examples are the following:

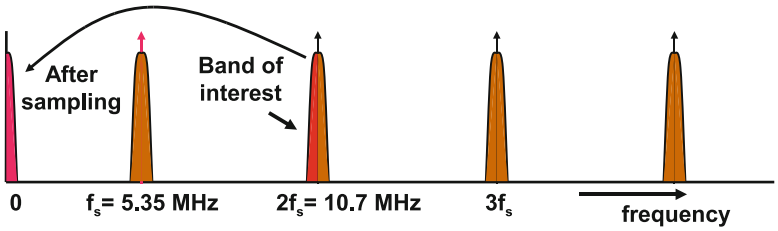


Fig. 3.6 Demodulation and sub-sampling of an IF-FM signal at 10.7 MHz by a 5.35 Ms/s sample pulse

- Harmonic distortion products of the base-band signal
- Thermal noise in the entire input band see Sect. 3.1.6
- Interference signals from other parts of the equipment or antenna

Example 3.3. Setup a subsampling scheme for an FM radio receiver.

Solution. In an FM radio the IF signal of 100 kHz is at a carrier frequency of 10.7 MHz. This signal can be down modulated and sampled at the same time by a 5.35 Ms/s signal (Fig. 3.6).

3.1.3 Sampling, Modulation and Chopping

Sampling of signals resembles 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. 3.7). Mathematically modulation is the multiplication of a signal with a pure sine wave of radial frequency ω_{local}:

$$G_{\text{mix}}(t) = A(t) \times \sin(\omega_{\text{local}} t), \tag{3.13}$$

$$A\sin(\omega t) \times \sin(\omega_{\text{local}} t) = \frac{A}{2}\cos((\omega_{\text{local}} - \omega)t) - \frac{A}{2}\cos((\omega_{\text{local}} + \omega)t). \tag{3.14}$$

In the result there are no components left at the input frequencies. Only two distinct frequencies remain. This formula is the basis for the first radio modulation technique: amplitude modulation.

If $A(t)$ is a band-limited spectrum mixing results in two frequency bands:

$$G_{\text{mix}}(t) = A(t) \times \sin(\omega_{\text{local}} t).$$

$$\mathbf{G}_{\text{mix}}(\omega) = \frac{1}{2}\mathbf{A}((\omega_{\text{local}} - \omega) - \frac{1}{2}\mathbf{A}(\omega_{\text{local}} + \omega),$$

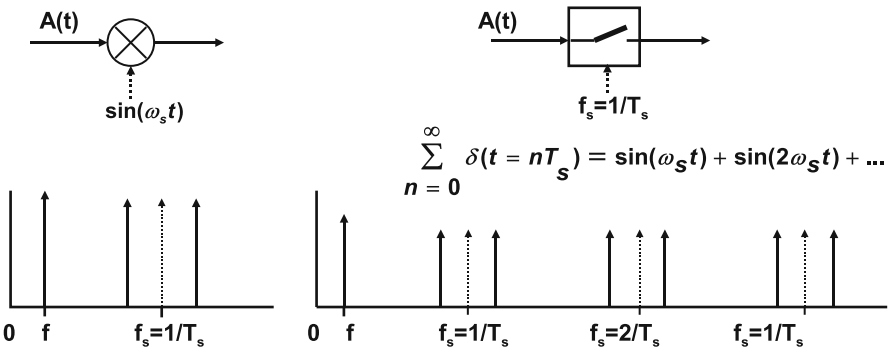


Fig. 3.7 Modulation and sampling of signals. Ideally the modulation and sampling frequencies disappear from the resulting spectrum. Here they are indicated for reference

the modulated bands appear as mirrored copies of each other around ω_{local}. Often one band is desired, and the other band is referred to as the "mirror image."

If the modulation principle is repeated the original sine wave can be recovered:

$$\begin{aligned} & G_{\text{mix}-\text{down}}(t) \\ & = G_{\text{mix}}(t) \times \sin(\omega_{\text{local}} t) \left(\frac{A}{2} \cos((\omega_{\text{local}} - \omega)t) - \frac{A}{2} \cos((\omega_{\text{local}} + \omega)t) \right) \\ & \quad \times \sin(\omega_{\text{local}} t) \\ & = \frac{A}{2} \sin(\omega t) - \frac{A}{4} \sin(2\omega_{\text{local}} t + \omega t) + \frac{A}{4} \sin(2\omega_{\text{local}} t - \omega t). \end{aligned} \tag{3.15}$$

The original component is accompanied by a pair of frequencies around $2\omega_{\text{local}}$. With a low-pass filter these components are removed.

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:

$$\mathbf{D_s}(\omega) = \frac{2\pi}{T_s} \sum_{k=-\infty}^{k=\infty} \delta \left(\omega - \frac{2\pi k}{T_s} \right). \tag{3.16}$$

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.

A particular form of sampling and mixing is called "self-mixing." A mixer can be seen as a device with two (mathematically) equivalent input ports. If a fraction of the signal on one port leaks into the other port, self-mixing will occur. In practical circuits mostly the large-amplitude local-oscillator frequency will leak into the low-amplitude port. If this leakage is described as $\alpha \sin(\omega_{\text{local}} t)$, the resulting output

component will be of the form: $\alpha/2+\sin(2\omega_{\text{local}}t)/2$. A noticeable DC component is the result, that can easily be mistaken for a circuit offset.

Chopping is a technique used for improving accuracy by modulating error-sensitive signals to frequency bands where the signal processing is free of errors (see also Sect. 7.7). First the signal is modulated to other frequency bands by multiplication with a chopping signal $f_{\text{chop}}(t)$. After signal processing, the signal is modulated back by multiplying again with $f_{\text{chop}}(t)$. The technique works well with a sine wave as modulator as $f^2_{\text{chop}}(t)$ contains a DC term and one frequency component at twice the chop frequency. Chopping can also be used to move unwanted signals out of the band of interest, e.g., chopping a DC-current source (also, dynamic element matching) will move mismatch and the $1/f$ noise to higher bands.

In differential circuits chopping is implemented by alternating between the differential branches. Mathematically this correspondents to a multiplication with a block wave with amplitude $+1,-1$. This block wave can be decomposed into a series of sines (see Table 2.9):

$$f_{\text{chop}}(t)=\sum_{n=1,3,5,\ldots}^{\infty}\frac{4\sin(n\pi/2)}{n\pi}\cos(\omega_{\text{chop}}t). \tag{3.17}$$

Now $f^2_{\text{chop}}(t)=1$ and a perfect restoration after chopping back is possible. Note that a signal $f_{\text{chop}}(t)$ composed of any sequence of $+1,-1$ transitions, fixed-frequency, or at arbitrary time moments, shows this property and can be used for chopping purposes. The spectrum of a fixed-frequency chop signal will be composed of a series of modulated spectra around odd multiples of the modulation frequency:

$$\mathbf{A_{chop}}(\omega)=\sum_{i=1,3,5,\ldots}^{\infty}\frac{4\sin(i\pi/2)}{i\pi}\mathbf{A}(i\omega_{\text{chop}}\pm\omega).$$

Chopping a spectrum from 0 to 1 MHz with a block-wave chopping signal of 10 MHz will remove the spectrum near DC and generate mirror bands at 9–11 MHz, 29–31 MHz, 49–51 MHz, etc.,

The higher bands of the chopped signal should not be removed. This would cause imperfections after chopping back. Any removed components can be regarded as a negative addition of signals to a perfectly chopped spectrum. These components will be treated as new input signals for the chopping back operation. So products of these components with the signal of Eq. 3.17 will appear. The removed parts of the spectrum $\mathbf{A}(n\omega_{\text{chop}}\pm\omega)$ will be modulated by the n-th harmonic of Eq. 3.17, resulting in an amplitude contribution at the position of the original signal with a relative strength of $1/n^2$. Also copies will appear at all odd multiple of the chopping frequency.

Example 3.4. A 135 MHz sine wave is sampled in a 150 Ms/s sampling system. Which frequency components will be in the sampled data spectrum? Is it possible to discriminate the result of this sampling process from sampling a 15 MHz sine wave?

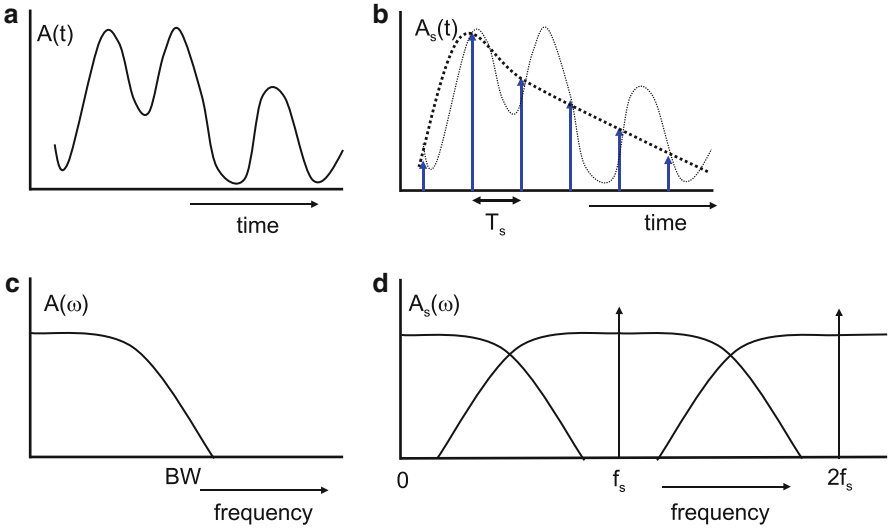


Fig. 3.8 The time-continuous signal contains higher frequency components and does not satisfy the Nyquist criterion. Both the sample series in the time domain (b) and the sampled spectrum that overlaps with the components around the sample frequency (d) do allow multiple reconstructions of the original time-continuous signal

Solution. If an input signal at frequency f_i is sampled by a sample rate of f_s then the sample data spectrum will contain the lowest frequency from the series: $f_i, (f_s - f_i), (f_s + f_i), (2f_s - f_i), (2f_s + f_i), \ldots, (nf_s - f_i), (nf_s + f_i), \ldots$ where $n = 0, 1, 2, \ldots, \infty$. In this case the second term delivers a 15 MHz component.

If directly a 15 MHz sine wave was sampled the sampled data sequence would be similar and even perfectly identical provided that the mutual phase shift is correct. In perfect conditions there is no way to tell from which time-continuous signal (in this case 15 or 135 MHz) this sequence originates. See also Example 3.10 on page 183.

3.1.4 Nyquist Criterion

Figure 3.8 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 3.8 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 are 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 Nyquist [85, Appendix 2-a], it was

Shannon who extended his mathematical theory of communication [86] in 1949 with a paper dealing with communication in the presence of noise. In that paper [87] 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:

$$f_s > 2\mathrm{BW}. \tag{3.18}$$

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 11–13 poles. Moreover such a filter will create a nonlinear 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 [88].

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. 3.1.3).

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. Comb filters are applied to separate the components. Still the net bandwidth of all useful components grouped together (Landau bandwidth) obeys the Nyquist criterion.

[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.

An even more advanced approach is observed in the theory of nonuniform sampling. In communication systems often only a limited number of carriers are simultaneously active. This is a sparse signal in a relatively wide bandwidth and 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 (Landau bandwidth) is still less than half of the effective sample rate.

Compressive sensing or compressive sampling [89] multiplies the signal with a high-rate random sequence, which is easier to implement in the analog domain than sampling. The relevant signals are again spread out over a large bandwidth. After bandwidth-limiting, a reconstruction ("L1" minimization) is possible if the random sequence is known, and the domain in which the signal is sparse. The theory is promising. Whether a real advantage can be obtained remains to be proven.

Example 3.5. Two sine wave signals at 3.2 and 4.8 MHz each modulated with a 0.1 MHz bandwidth signal are sampled at 1.1 Ms/s. Is the Nyquist criterion violated?

Solution. No, the total band occupied with relevant signal is 0.4 MHz, while the Nyquist bandwidth is 0.55 MHz. The sample rate must be carefully chosen not to mix things. Here the sampled bandwidths will span 0–0.2 and 0.3–0.5 MHz.

3.1.5 Alias Filter

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. 3.9).

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

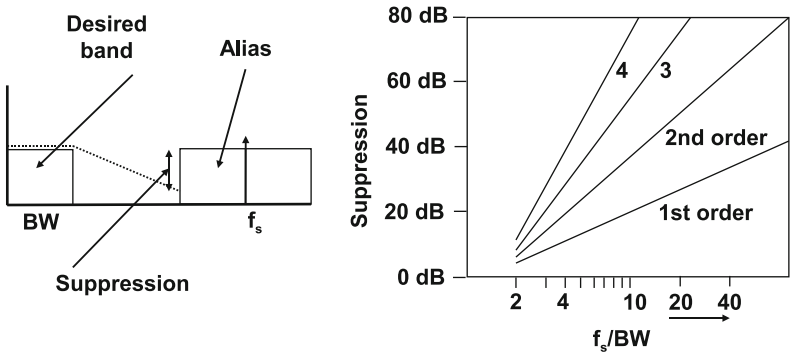
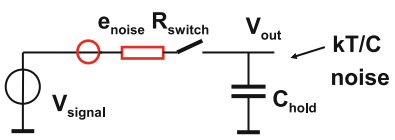


Fig. 3.9 The attainable suppression of the anti-alias filter depends on the number of poles in the filter and the ratio of the bandwidth to the sample rate



C_{hold}	V_{kTC}
1 pF	65 μV_{rms}
3 pF	35 μV_{rms}
10 pF	20.4 μV_{rms}

Fig. 3.10 Switched capacitor noise sampling: the series resistances act as a noise sources

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 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.

Example 3.6. A bandwidth of 2 MHz is sampled at 10 MHz. Determine the order of an anti-alias filter-suppressing alias components by 30 dB.

Solution. Aliasing will occur due to signals in bands around multiples of the sampling rate. In this case all signals between $f_s - \mathrm{BW} = 8$ MHz and $f_s + \mathrm{BW} = 12$ MHz will appear after sampling in the desired signal band. The task will be to design a filter that passes a bandwidth of 2 MHz, but suppresses at 8 MHz. With the help of Fig. 2.63, a seventh order filter is chosen. It is obvious that doubling the sample rate eases this requirement dramatically.

3.1.6 Sampling of Noise

Figure 3.10 shows an equivalent schematic of the basic sampling circuit consisting of a switch and a storage capacitor. Compared to the ideal situation two nonideal

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:

$$e_{\text{noise}} = \sqrt{4kTR\text{BW}} \tag{3.19}$$

with Boltzmann's constant $k = 1.38 \times 10^{-23}\,\text{m}^2\text{kgs}^{-2}\,\text{K}^{-1}$ and the absolute temperature T in Kelvin. This formulation expresses the noise in the positive frequency domain from 0 to ∞. 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,

$$v_{C,\text{noise}}^2 = \int_{f=0}^{f=\infty} \frac{4kTR\,\mathrm{d}f}{1+(2\pi f)^2R^2C^2} = \frac{kT}{C} \quad \Rightarrow v_{C,\text{noise}} = \sqrt{\frac{kT}{C}}. \tag{3.20}$$

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.

The same result follows from classical thermodynamics. The equipartition theorem says that, in thermal equilibrium, the thermal energy is equally distributed over all degrees of freedom. For a capacitor there is only one degree of freedom: its potential. Therefore energy contained in the thermal fluctuation of carriers $CV_{\text{noise}}^2/2$ equals the thermal energy for one degree of freedom: $kT/2$.

Equation 3.20 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 cutoff 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. 3.11).

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 $65\,\mu\text{V}_{\text{rms}}$ at room temperature. Unfortunately, a large capacitance value will require IC area, and will directly impact the power budget.

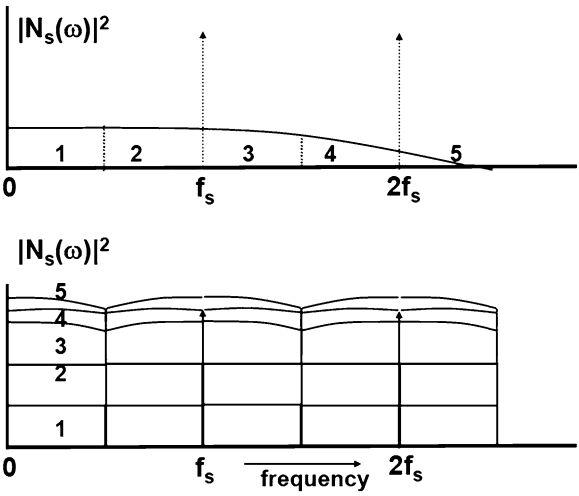


Fig. 3.11 Band-limited noise is sampled in a similar manner as normal signals. *Top*: this noise has a finite bandwidth, *bottom*: after sampling. The power spectra add up and are mirrored

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:

$$S_{ff,SH} = \frac{2kT}{Cf_s}. \tag{3.21}$$

The power noise density of the time-continuous network with the same resistor and capacitor having a cut-off frequency of $f_{\mathrm{RC}} = 1/2\pi\mathrm{RC}$ in its pass-band is:

$$S_{ff,rc} = 4kTR = \frac{2kT}{\pi Cf_{rc}}. \tag{3.22}$$

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. 3.11. 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 root-mean-square way. Also in situations where a switch discharges the charge of a capacitor into a fixed voltage, kT/C noise will appear ("reset-noise").

Example 3.7. An uncorrelated white noise with a total effective value of $1\,\mathrm{mV_{rms}}$ in the band limited to 120 MHz is sampled at 10 Ms/s. What is the noise density of the source? What is the noise density after sampling? What is the rms value of the noise signal after sampling?

Solution. The effective noise level of $1\,\mathrm{mV_{rms}}$ means that the noise has a power equal to the power of a 1 mV DC source. That allows a calculation of the noise

density of the noise source $S_{vv} = (1\text{mV})^2/120\,\text{MHz}$. After sampling all noise bands higher than $f_s/2$ are folded back to the base band. In this case the frequency range between DC and 5 MHz will contain 24 uncorrelated noise bands. The noise density is consequently: $S_{vv,s} = 24 \times (1\text{mV})^2/120\,\text{MHz}$. The total energy after sampling is found from integration of the noise density over the band, yielding again an effective noise level of: $1\,\text{mV}_{\text{rms}}$. What about the noise in the band beyond 5 MHz? The noise density in those bands is equally high and real, but during the reconstruction process no more energy can be retrieved than what is available in one band.

Example 3.8. In a process with a nominal supply voltage of 1.2 V a sinusoidal signal of 100 MHz and $500\,\text{mV}_{\text{peak–peak}}$ is sampled. A SNR of 72 dB is required. Calculate the capacitor.

Solution. $500\,\text{mV}_{\text{peak–peak}}$ corresponds to an rms voltage of $500/2\sqrt{(2)} = 177\,\text{mV}$. With a signal-to-noise ratio of $10^{72/20} = 4,000$ (corresponding to a 12 bit ADC) the kT/C noise must be lower than $177\,\text{mV}/4000 = 44\,\mu\text{V}_{\text{rms}}$, and a minimum capacitor of 2.15 pF is needed. A sinusoidal signal with a frequency of 100 MHz requires a current of $+/-\,0.34\,\text{mA}$. This charge on the capacitor has to be supplied, and a bias current of 1 mA must be used to avoid slew-rate behavior. In first order this circuit requirement will consume 1.2 mW.

3.1.7 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. 3.12. If noise changes the switching level of a buffer, 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 $30\text{–}100\,\text{ps}_{\text{rms}}$.[4]

[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.

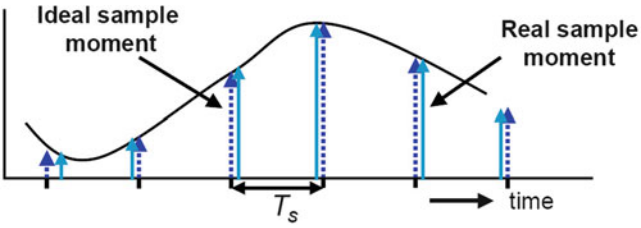


Fig. 3.12 The ideal sampling moments (*dashed*) shift in an arbitrary fashion in time if the sample clock is disturbed by jitter

See Sect. 2.7.16 for some theoretical background and the relation to phase noise.

Figure 3.12 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:

$$A(nT_s + \Delta T(t)) = \hat{A}\sin(\omega \times (nT_s + \Delta T(t)) \tag{3.23}$$

$$\Delta A(nT_s) = \frac{\mathrm{d}\hat{A}\sin(\omega t)}{\mathrm{d}t} \times \Delta T(nT_s) = \omega\hat{A}\cos(\omega nT_s)\Delta T(nT_s). \tag{3.24}$$

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:

$$\sigma_A = \sqrt{\frac{1}{T}\int_{t=0}^{T}(\omega\hat{A}\cos(\omega t)\sigma_{\mathrm{jit}})^2\mathrm{d}t} = \frac{\omega\hat{A}\sigma_{\mathrm{jit}}}{\sqrt{2}}. \tag{3.25}$$

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:

$$\mathrm{SNR} = \left(\frac{1}{\omega\sigma_{\mathrm{jit}}}\right)^2 = \left(\frac{1}{2\pi f\sigma_{\mathrm{jit}}}\right)^2 \tag{3.26}$$

or in decibel (dB):

$$\mathrm{SNR} = 20\,{}^{10}\log\left(\frac{1}{\omega\sigma_{\mathrm{jit}}}\right) = 20\,{}^{10}\log\left(\frac{1}{2\pi f\sigma_{\mathrm{jit}}}\right). \tag{3.27}$$

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.

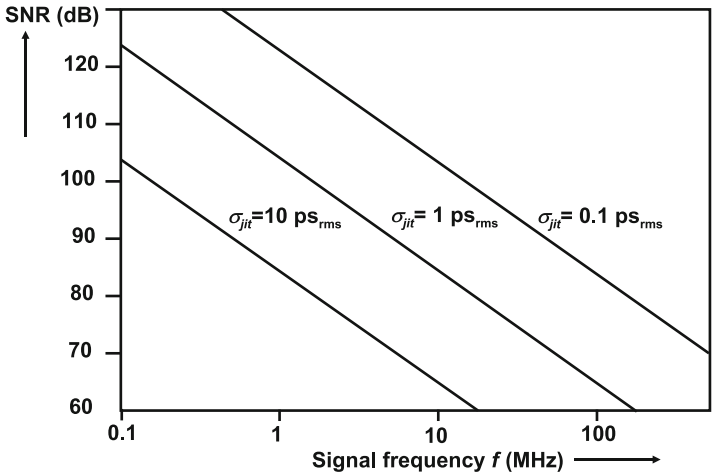


Fig. 3.13 The signal-to-noise ratio depends on the jitter of the sampling signal and the frequency of the time-continuous signal

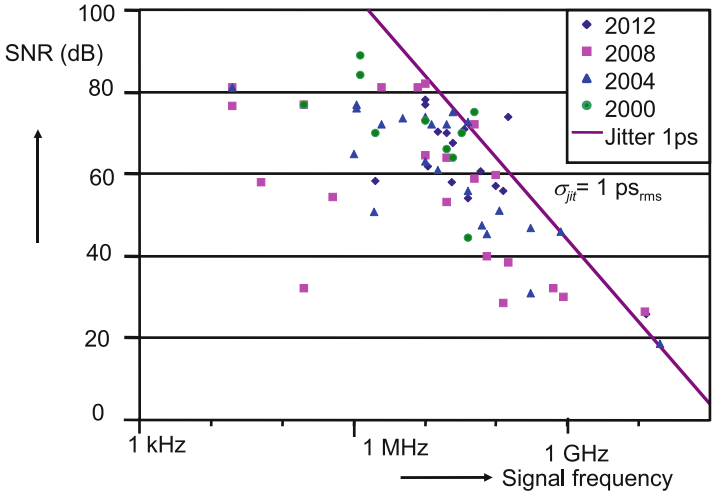


Fig. 3.14 The signal-to-noise ratio versus the signal frequency of analog-to-digital converters reported on the International Solid-State Circuits conferences in the years 2000, 2004, 2008, and 2012. (From: B. Murmann, "ADC Performance Survey 1997–2012," Online: http://www.stanford.edu/ murmann/adcsurvey.html)

Figure 3.13 shows the signal to noise ratio as a function of the input frequency for three values of the standard deviation of the time jitter. Figure 3.14 compares the jitter performance of analog-to-digital converters published on the International Solid-State Circuits conferences in the years 2000, 2004, 2008, and 2012. It is

Table 3.2 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	$> 50\,ns_{rms}$

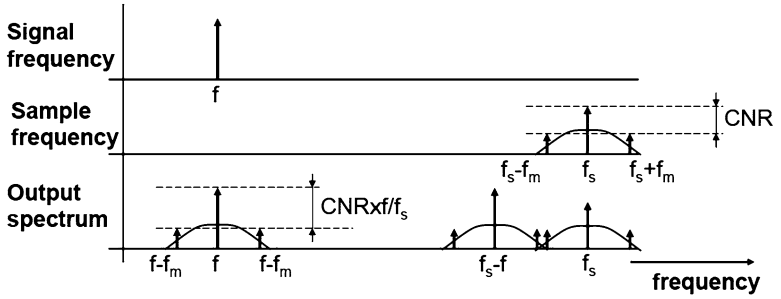


Fig. 3.15 Jitter around the sampling frequency will produce side spectra around the input tone

obvious that a jitter specification better than $\sigma_{\text{jitter}} < 1\,\text{ps}$ is a challenge. The comparison of the best converters in every year shows that little progress was made.

Table 3.2 indicates some jitter numbers from commercial timing components.

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 [90].

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. 2.99) such as the following:

- 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. 3.15. 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. 3.16, the jitter can also 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. 3.15 and are the source for spurious components and

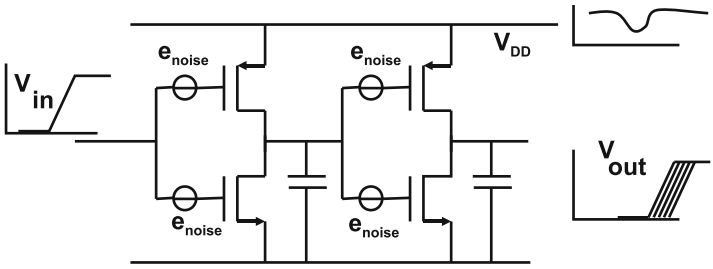


Fig. 3.16 A clock buffer for generating the digital sample signal can add to an ideal sample signal some noise of the buffer transistors. Also fluctuations on the power supply will affect the switching behavior of the buffer, causing uncertainty on the edges and jitter in the sampling

signal distortion. Also the digital circuits that generate and propagate the sample pulse must be treated as if these were analog blocks.

Example 3.9. The clock buffer in Fig. 3.16 has edge transition times of 70 ps. How much jitter can be expected if a random noise of 60 mV_{rms} is present on the power supply of 1.2 V?

Solution. Due to voltage changes on the power supply lines, the currents inside the buffer will change, in first-order proportional to the voltage change. As a consequence the slope of the transition will vary linearly. The mid-level switching point is now reached after 35 ps delay from the input mid-level passing. A voltage change of 60 mV/1.2 V = 5 % will create a 5 % delay variation on the slope and the delay of 35 ps. So the expected jitter is 1.75 ps_{rms} per edge. As the same voltage variation applies to two inverters, the overall jitter is 3.5 ps_{rms}.

Example 3.10. In Example 3.4 the sample sequence is distorted by a random jitter component of 5 ps_{rms}. Is it possible to discriminate in the sampled data domain between a 15 MHz input sinewave or a 135 MHz input sine wave?

Solution. With perfect sampling both signals will result in equivalent wave forms in the sampled-data domain. However, the presence of jitter allows to discriminate, as the resulting SNR for a 15 MHz input signal is $\text{SNR} = 1/2\pi f_i \sigma_{\text{jitter}} = 66.5\,\text{dB}$, while the SNR for 135 MHz equals 47.5 dB.

Example 3.11. Calculate the jitter due to thermal noise that an inverter with dimensions N/P of 0.2/0.1 and 0.5/0.1 in a 90-nm CMOS process adds to an edge of 50 ps (bottom-top).

Solution. Every transistor adds noise that is related to the transconductance in the channel: $i_{\text{noise}} = \sqrt{4kT\text{BW}g_m}$. If the inverter is at its mid-level point (0.6 V) both transistors will be contributing to a total noise current of $i_{\text{noise},n+p} = \sqrt{4kT\text{BW}(g_{m,n} + g_{m,p})}$. This noise corresponds to an input referred noise voltage of $v_{\text{noise},n+p} = i_{\text{noise},n+p}/(g_{m,n} + g_{m,p}) = \sqrt{4kT\text{BW}/(g_{m,n} + g_{m,p})}$. With the help of Table 2.20, an input noise voltage is found of 0.62 mV in a 10 GHz bandwidth. This

bandwidth is an approximation based on the observation that a rising edge of 50 ps followed by a similar falling edge limits the maximum frequency of the inverter to 10 GHz. The jitter is estimated as $\tau_{\text{jitter}} : \tau_{\text{edge}} = v_{\text{noise},n+p} : V_{\text{DD}}$ and $\tau_{\text{jitter}} = 25$ fs.

3.2 Time-Discrete Filtering

Time-discrete filtering is used for designing up- and down-sampling of spectra in digital-to-analog converters of Sect. 9.1 for the design of time-discrete sigma–delta modulators. Time discrete filtering forms a subset of time-discrete signal processing (see e.g., [83, 84]). In analog-to-digital converters, especially the sigma–delta modulators of Sect. 9.4 apply time-discrete filters, both in their architecture and 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 flip-flops. Frequency filters are realized by combining the time-delayed sampled with specific weighting factors or multiplying coefficients.

Operations and functions in the discrete-time domain are described in the z-domain (see Sect. 2.1.5). A delay of one basic sample period is transformed into the function z^{-1}. A frequency sweep from 0 to $f_s/2$ results in a circular movement of the z vector in a complex plane from $+1$, via $0+j$ to -1. For the frequency range $f_s/2$ to f_s the z vector will turn via the negative imaginary plane and return to $z = 1$. Figure 3.17 shows two integrators described in the z-domain. The left structure adds the present sample to the sum of the previous samples. After the sample pulse the output will be the integral over all past sample moments. The right topology does the same; however, now the sum is directly available. The transfer functions for both structures are

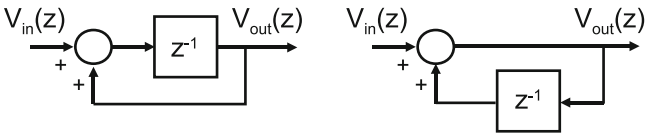


Fig. 3.17 Two integrators in the z-domain

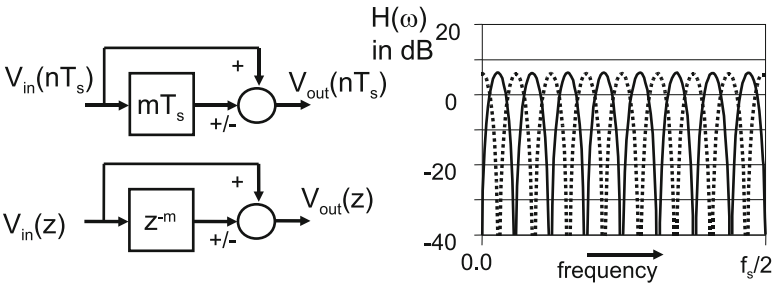


Fig. 3.18 The comb filter as sampled-data structure and in the z-domain. The frequency response shows in a *solid line* the sine response (minus sign at the summation), while the *dotted line* represents the cosine response (plus sign)

$$H(z) = \frac{z^{-1}}{1 - z^{-1}} \qquad H(z) = \frac{1}{1 - z^{-1}}.$$

The first integrator has a pole at DC: $z = 1$ and a zero at $z = 0$. The most simple filter in the time-discrete domain is the comb filter. The addition of a time-discrete signal to an m-sample period delayed signal gives a frequency transfer that can be evaluated in the z-domain:

$$H(z) = 1 \pm z^{-m}.$$

This function has zeros at all frequencies where $z^{-m} = \pm 1$, resulting in m zeros distributed over the unity circle. Using the approximation $z \leftrightarrow \mathrm{e}^{sT_s}$ results in:

$$\begin{aligned} H(s) &= 1 \pm \mathrm{e}^{-smT_s} = \mathrm{e}^{-smT_s/2}(\mathrm{e}^{+smT_s/2} \pm \mathrm{e}^{-smT_s/2}) \\ |H(\omega)| &= 2|\cos(\omega m T_s/2)|, \quad \text{addition} \\ |H(\omega)| &= 2|\sin(\omega m T_s/2)|, \quad \text{subtraction,} \end{aligned} \tag{3.28}$$

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. 3.18). In this plot the zeros are observed in the frequency domain.

Comb filters are mostly applied in systems where interleaved 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. 3.19). 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 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

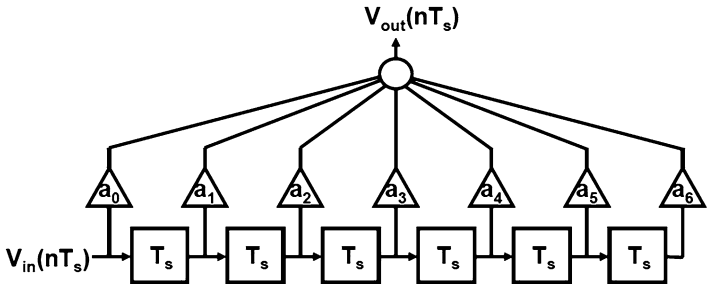


Fig. 3.19 The general structure of a finite impulse filter

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:

$$y(nT_s) = \sum_{k=0}^{k=\infty} h(k)x(nT_s - kT_s). \tag{3.29}$$

Applied to the filter in Fig. 3.19, this gives

$$V_{\text{out}}(nT_s) = \sum_{k=0}^{k=N-1} a_k V_{\text{in}}((n-k)T_s). \tag{3.30}$$

The z-transform results in a description of the transfer of a FIR filter:

$$H(z) = \sum_{k=0}^{k=N-1} a_k z^{-k}. \tag{3.31}$$

In order to transform this transfer function from the discrete-time domain to the frequency domain, the term z^{-1} is substituted by $\mathrm{e}^{-j\omega T_s}$, which results in

$$H(\omega) = \sum_{k=0}^{k=N-1} a_k \mathrm{e}^{-jk\omega T_s}. \tag{3.32}$$

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 hold 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]:

$$e^{-j\omega T_s} = \cos(\omega T_s) - j\sin(\omega T_s). \tag{3.33}$$

After moving the average delay $NT_s/2$ out of the summation, real and imaginary terms remain:

$$H(\omega) = e^{-j\omega NT_s/2} \sum_{k=0}^{k=N/2-1} (a_k + a_{N-k})\cos(k\omega T_s/2) - j(a_k - a_{N-k})\sin(k\omega T_s/2). \tag{3.34}$$

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.

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") [91]. This routine optimizes the transfer based on a number of specifications.

Figure 3.20 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 3.21 shows the impulse response for a somewhat more elaborate filter with 48 coefficients. The transition between the passband 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 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. 3.22).

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

[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.

[7] An equivalent analog filter would require 10–12 poles.

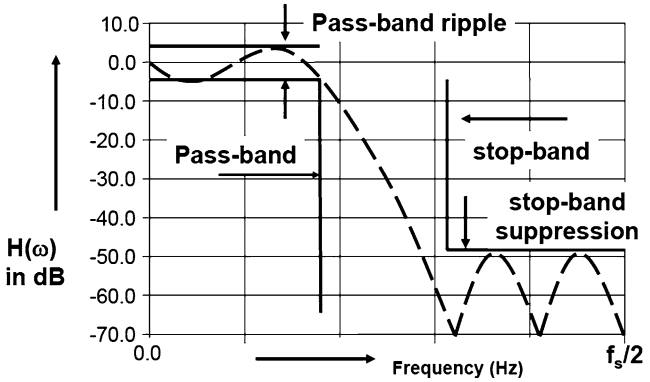


Fig. 3.20 Definition scheme of a filter response

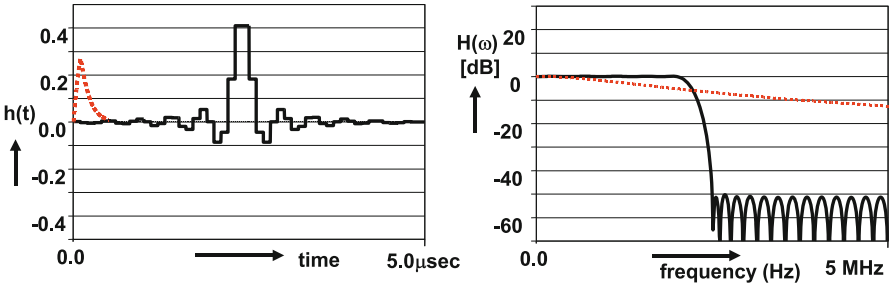


Fig. 3.21 A low-pass FIR filter with 48 coefficients; *left* is the impulse response of the filter and its analog realization (*dashed*). *Right* is the frequency response of both

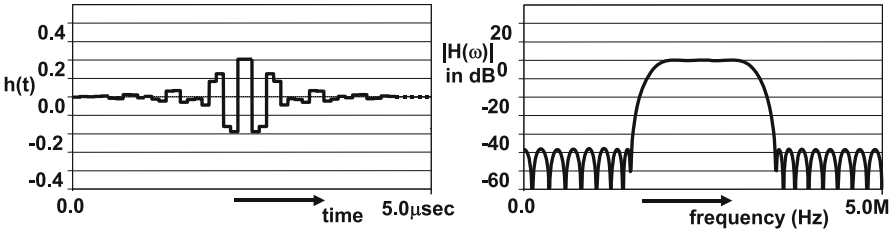


Fig. 3.22 A band-pass FIR filter with 48 coefficients; *left* is the impulse response of the filter and its analog realization. *Right* is the frequency response of both

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.

Table 3.3 Coefficient values for the low-pass FIR filter of Fig. 3.23

Coefficient	Value
$a_0 = a_6$	−0.06
$a_1 = a_5$	0.076
$a_2 = a_4$	0.36
a_3	0.52

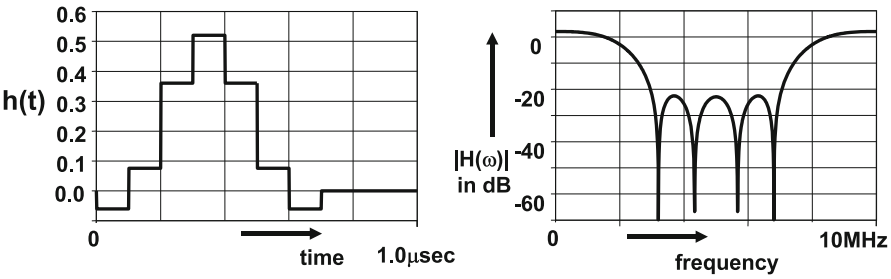


Fig. 3.23 The impulse response and the frequency transfer function of a seven coefficient filter from Fig. 3.19 at a 10 MHs/s sample rate

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 [92].[8] In digital-to-analog conversion, the semi-digital filter uses digital delays with analog coefficients, (see Sect. 7.3.5).

Example 3.12. Determine with a suitable software tool the coefficients for the structure in Fig. 3.19 to create a low-pass filter.

Solution. If the transition for the low-pass filter is chosen at approximately $f_s/4$, coefficients as in Table 3.3 are found.

Figure 3.23 shows the time response and the frequency transfer function from Fig. 3.19 with the coefficients of Table 3.3. The time response is an impulse response: the input for the filter is a pulse with an amplitude of "1" during one sample 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.

[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.

3.2.2 Half-Band Filters

In order to reduce the complexity of digital FIR filters additional constraints are needed. Introducing the symmetry requirement

$$H(\omega) + H(\omega_s/2 - \omega) = 1 \tag{3.35}$$

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:

$$\begin{aligned} a_m &= 0.5 \\ a_{m+i} = a_{m-i} &= C_i, \qquad i = 1,3,5,\ldots \\ a_{m+i} = a_{m-i} &= 0, \qquad i = 2,4,6,\ldots \end{aligned}$$

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:

$$a_{m+i} = \frac{\sin(i\pi/2)}{i\pi}, \qquad i = -m, \ldots, -2, -1, 0, 1, 2, \ldots, m. \tag{3.36}$$

Table 3.4 lists the coefficients for four half-band filters designed for a passband from 0 to $f_s/8$ and a stop band from $3f_s/8$ to $f_s/2$. Figure 3.24 compares these four half-band filter realizations. The filter with the least suppression has three nonzero 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.

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 $f_s/2$.

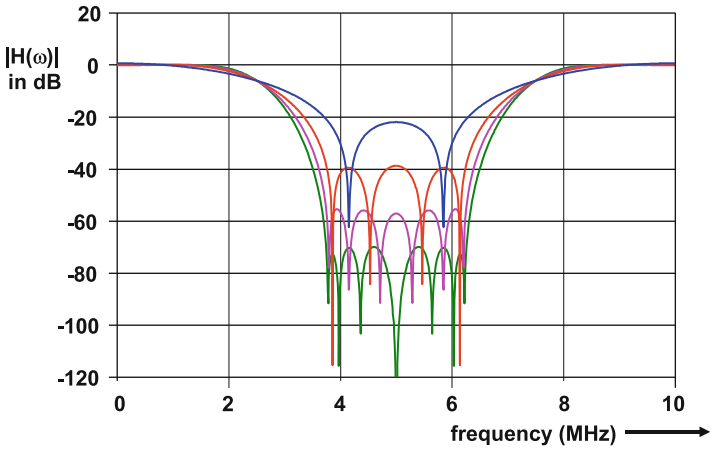


Fig. 3.24 Four half-band filters, with 3, 5, 7, and 9 nonzero coefficients (Courtesy: E.E. Janssen)

Table 3.4 The nonzero coefficients for four half-band filters in Fig. 3.24

Coefficients	Suppression	Ripple
$a_m = 0.5$ $a_{m-1} = a_{m+1} = 0.2900$	20 dB	0.8 dB
$a_m = 0.5$ $a_{m-1} = a_{m+1} = 0.2948$ $a_{m-3} = a_{m+3} = -0.0506$	38 dB	0.1 dB
$a_m = 0.5$ $a_{m-1} = a_{m+1} = 0.3016$ $a_{m-3} = a_{m+3} = -0.0639$ $a_{m-5} = a_{m+5} = 0.0130$	55 dB	0.014 dB
$a_m = 0.5$ $a_{m-1} = a_{m+1} = 0.3054$ $a_{m-3} = a_{m+3} = -0.0723$ $a_{m-5} = a_{m+5} = 0.0206$ $a_{m-7} = a_{m+7} = -0.0037$	72 dB	0.002 dB

Courtesy: E.E. Janssen

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. 3.25) (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. 3.25 (lower-left)) [84]. An efficient alternative is to split the filter in two sections for the odd and even samples (Fig. 3.25 (right)).

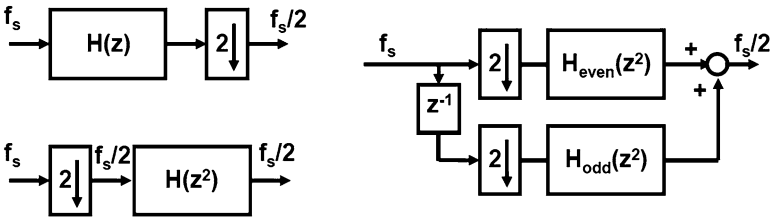


Fig. 3.25 Down-sampling can be realized by selecting every second sample (*upper-left*). Sample selection and decimation can be interchanged (*lower-left*). More advanced methods split the filter in an even-sample and odd-sample portion

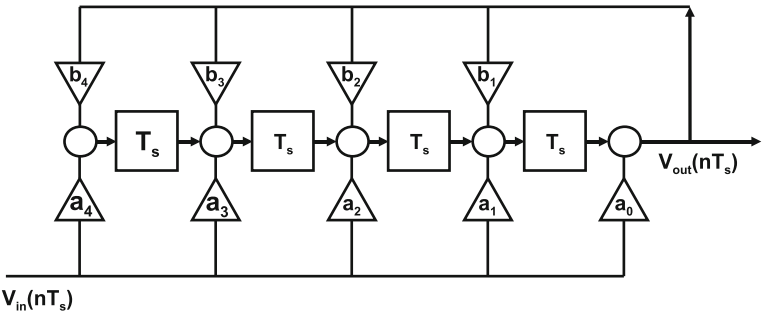


Fig. 3.26 The structure of an infinite impulse response (IIR) filter containing a feedback path from output to the summation nodes

3.2.4 IIR Filters

A drastic solution to the hardware problem of FIR filters is the "infinite impulse response" IIR filter.

Figure 3.26 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 $z = \mathrm{e}^{j\omega T_s}$ is applied):

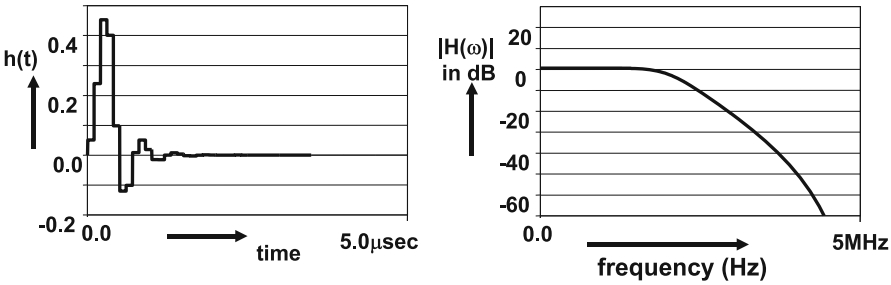


Fig. 3.27 The impulse response of this 4-tap IIR continues beyond four sample periods. The frequency response of this 4-tap filter is much steeper than the response of the 7-tap FIR filter

Table 3.5 Comparison of discrete filter realization techniques

Implementation	Switched capacitor	Semi-digital filter Sect. 7.3.5	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

$$H(z) = \frac{\sum_{k=0}^{k=N-1} a_k z^{-k}}{1 - \sum_{k=1}^{k=N-1} b_k z^{-k}} \Leftrightarrow H(\omega) = \frac{\sum_{k=0}^{k=N-1} a_k \mathrm{e}^{-jk\omega T_s}}{1 - \sum_{k=1}^{k=N-1} b_k \mathrm{e}^{-jk\omega T_s}}. \tag{3.37}$$

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 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. 3.27 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 3.5 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.

Exercises

3.1. A sinusoidal signal of 33 MHz is distorted and generates second and third harmonics. It is sampled by a 32 Ms/s system. Draw the resulting spectrum.

3.2. An IF-television signal occupies a bandwidth from 33 to 39 MHz. Propose a sampling frequency that maps the 39 MHz component on DC. Consider that the power consumption of the following digital circuit is proportional to the sampling frequency and must be low.

3.3. A signal bandwidth from DC to 5 MHz must be sampled in the presence of an interferer frequency at 20 MHz. Choose an appropriate sampling rate.

3.4. An image sensor delivers a sampled-and-held signal at a fixed rate of 12 Ms/s. The succeeding digital signal processor can run at 10 MHz. Give an optimal theoretical solution. What is a (nonoptimal) practical solution?

3.5. What is a stroboscope? Explain its relation to sampling.

3.6. Must the choice for a chopping frequency obey the Nyquist criterion?

3.7. Set up a circuit where the signal is stored as a current in a coil. What is the meaning of $i_{\text{noise}} = \sqrt{KT/L}$?

3.8. The signal energy of the luminance (black-and-white) signal of a television system is concentrated around multiples of the line frequency (15,625 Hz). Although the total television luminance signal is 3 MHz wide, a sample rate of around 4 Ms/s is possible. Give an exact sample rate. Why will this sampling not work for a complete television signal with color components added?

3.9. An audio system produces samples at a rate of 44.1 ks/s. With a maximum audio signal of -6 dB of the full-scale between 10 and 20 kHz, propose an alias filter that will reduce the frequency components over 20 kHz to a level below -90 dB.

3.10. How much SNR can be obtained if a signal of 10 MHz is sampled with a sample rate of 80 Ms/s with 5 ps_{rms} jitter. What happens with the SNR if the sample speed is increased to 310 Ms/s at the same jitter specification. Compare also the SNR in a bandwidth between 9 and 11 MHz.

3.11. Design a half-band filter with 19 nonzero coefficients to get a pass-band stop band ratio of 100 dB. Use a computer program.

3.12. An analog-to-digital converter is sampling at a frequency of just 2.5$\times$ the bandwidth. Due to the large spread in passive components, the problem of alias

filtering is addressed by placing before the converter a time-discrete filter running at twice the sample rate and, before that time-discrete filter, a time-continuous filter. Is this a viable approach? There are twice as many samples coming out of the filter than the converter can handle. Is there a problem?

3.13. Make a proposal for the implementation of the filters in the previous exercise if the bandwidth is 400 kHz and an attenuation of better than 60 dB must be reached starting from 500 kHz.

Chapter 4
Sample-and-Hold

Abstract The sample-and-hold circuit or track-and-hold circuit performs the sampling operation. These circuits have to operate at the highest signal levels and speeds, which makes their design a challenge. The chapter discusses first the specific metrics for these circuits, such as pedestal step, droop time, and hold-mode feedthrough. The different elements, switch, capacitor, and buffer, are discussed. Some architectures and often applied implementation schemes are shown. The trade-off between noise and distortion requires a careful balance to achieve the optimum performance.

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 an 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;

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_4,

© Springer Science+Business Media, LLC 2013

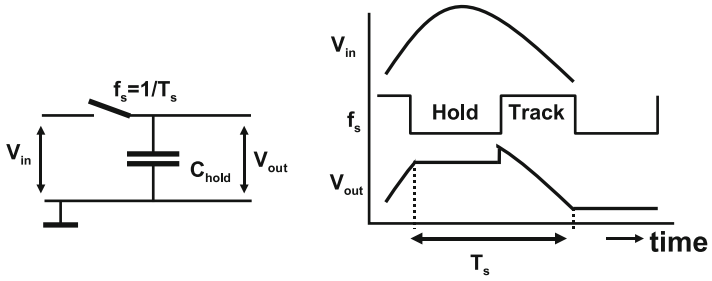


Fig. 4.1 A switch and a storage element form a track-and-hold circuit

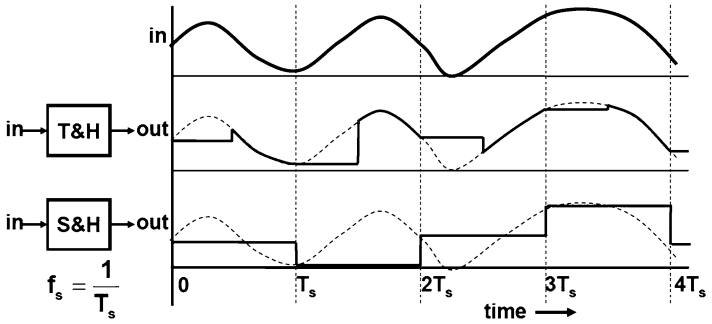


Fig. 4.2 The input signal (*above*) is tracked by a track-and-hold circuit during one phase of the clock period and held during the other phase. A sample-and-hold circuit holds the signal during the full period of the clock signal

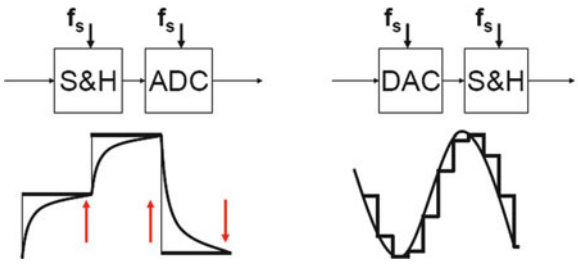


Fig. 4.3 A sample-and-hold circuit is used as an input sampler for an analog-to-digital converter or as a de-glitch circuit in an digital-to-analog converter

see Fig. 4.3 (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. For example, 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.

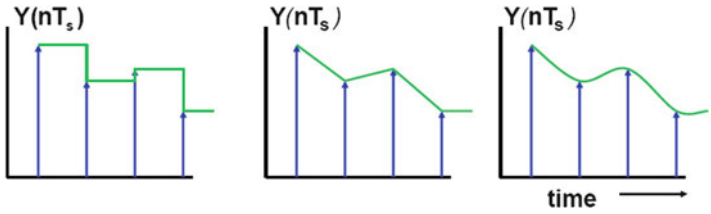


Fig. 4.4 A Dirac sequence can be reconstructed via a zero-order hold, a first-order interpolation, or higher-order reconstruction algorithms

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; 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).

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., Sect. 7.8.4). 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.

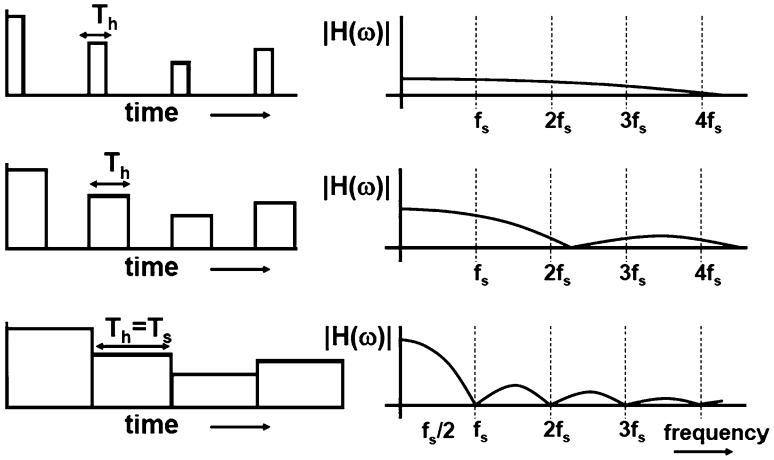


Fig. 4.5 The hold time determines the filter characteristics of a sample-and-hold

$$h(t) = \begin{cases} 1, & 0 < t < T_h \\ 0, & \text{elsewhere.} \end{cases} \tag{4.1}$$

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:

$$H(\omega) = \int_{t=0}^{t=\infty} h(t) \times \mathrm{e}^{-j\omega t} \mathrm{d}t = \frac{1}{T_s} \int_{t=0}^{t=T_h} 1 \times \mathrm{e}^{-j\omega t} \mathrm{d}t = \frac{\sin(\pi f T_h)}{\pi f T_s} \mathrm{e}^{-j\omega T_h/2}. \tag{4.2}$$

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 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

[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".

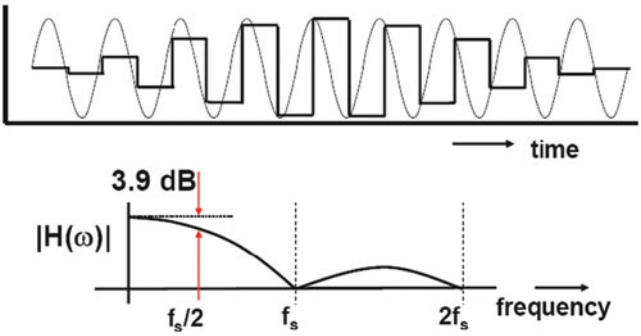


Fig. 4.6 Time and frequency response of a sample-and-hold signal close to half of the sampling rate

of the transfer curve is often summarized to "$\sin(x)/x$" behavior. Some authors use "$\mathrm{sinc}(x)$." The integral $\int \sin(x)/x \mathrm{d}x$, see Table 2.7, belongs to the mathematical class of Dirichlet integrals.

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. 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 depending 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.

Example 4.1. Can the $\sin(x)/x$ frequency behavior of a T&H circuit be compensated by an analog filter?

Solution. In the frequency domain the amplitude loss can (partially) be compensated by means of a first-order high pass filter, e.g., a voltage divider of two resistors, where the top resistor is shunted with a capacitor. However, beyond $f_s/2$, deviations occur. Moreover the frequency response is an average over infinite time, and the instantaneous time-domain response will show at certain phase relations large excursions. Finally in this setup the high-frequency noise will be amplified.

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 (SNR) 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

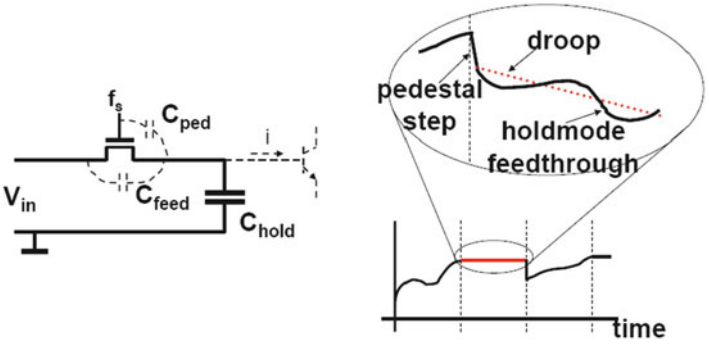


Fig. 4.7 Three artifacts in a sample-and-hold: droop, hold-mode feedthrough, and pedestal step

the most critical components in realizing an optimum conversion performance. The specification of the performance of an S&H circuit consists of the standard analog ingredients: distortion (THD), SNR, power consumption (P), etc. Next to these standard requirements, a number of specific specifications exist 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 feedthrough 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 feedthrough is minimum in integrated solutions, a T-switch may further reduce this effect, Fig. 4.12.
- The aperture time is the time period in which the sample value is determined. Normally this is a relatively small time period in which the sample pulse makes the transition to switch off. During this clock edge still some tracking of the signal can continue. This will add a small additional fixed delay to the actual sampling moment 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. 3.1.7. Deviations of the aperture time can arise because an NMOS switch remains longer conductive when the signal is at a low level than at high levels.

Aperture distortion is introduced when the slew rate of the signal is comparable to the transition time of the switch. A signal change that follows the edge of the switching pulse postpones the sampling moment, while a signal change in opposite direction of the switching pulse will speed up the effective sampling moment.

One analog-to-digital architecture has a deliberately long aperture time: the dual-slope converter; see Sect. 8.7. Here the signal is integrated during the aperture time. The effect on the transfer characteristic is a filtering of high-frequency components (noise), which is a desired effect in applications of this converter.

Example 4.2. What is the minimum frequency on which a track-and-hold circuit can operate if the hold capacitor of 1 pF experiences a leakage current of 1 nA? The output must be accurate within 1 mV.

Solution. The voltage on the hold capacitor will see a voltage change of $\Delta V_{C_{\text{hold}}} = I_{\text{leakage}} \times T_{\text{hold}}/C_{\text{hold}}$. With the above numbers $T_{\text{hold}} \leq 1\,\mu\text{s}$. If the hold time is dominant the minimum sample rate is 1 Ms/s.

Example 4.3. While measuring a sample-and-hold circuit an unexpected signal appears in the band of interest at frequency f_{oeps}. How can the origin of this component be determined?

Solution. Many unwanted frequency components in a measurement originate from a combination of unexpected signals and folding back by means of sampling. Distortion is a potential cause. In that case the resulting component will be observable at frequency: $n \times f_s \pm m \times f_i$, with f_s and f_i as the sample rate and the input test frequency. In order to determine the factors n, m, the sample rate is increased by, e.g., 1 kHz. If the unwanted frequency component is created by the sampling process at $n \times f_s$, a shift in f_{oeps} of n kHz will occur. In the same way the relevant harmonic can be traced: the test frequency is increased by 1 kHz, and if f_{oeps} shifts m kHz, the mth harmonic is involved.

In case f_{oeps} shifts with an f_s increase but not with f_i, probably an external frequency is being modulated into the baseband, perhaps an oscillation in a power supply, an oscillation in supporting circuits (high-frequency buffers over the edge of stability), feedthrough of a digital signal, a test setup of a colleague, a mobile phone in the test room, etc.

In case there is no reaction nor on sample rate neither on input frequency shifts the printed-circuit board must pick up an external frequency: refrigerators, air-conditioning, and any of the above.

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 SNRs 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 Sects. 2.3.6 and 2.4.11:

- Diffusion capacitances are not suited for most specifications. Leakage, nonlinearities, and low capacitance per unit area generate too many problems.
- The gate-to-channel capacitor, see Sect. 2.4.7, 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 and has a good linearity and low parasitics, see Fig. 2.39.
- In some process variants a so-called metal–insulator–metal capacitance option is present with excellent properties.
- In older processes a double polysilicon capacitor option is offered. The depletion of the polysilicon can lead to voltage nonlinearities.

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.

Example 4.4. A signal is sampled on two equal parallel connected capacitors. After sampling the two capacitors are connected in series doubling the signal component. Does this method improve the SNR?

Solution. The value of a single capacitor is C. So the root-mean-square noise after sampling is $v_{\text{noise}} = \sqrt{kT/C}$ per capacitor, and $\sqrt{kT/2C}$ for the parallel connection. With a signal amplitude of V_A and a root-mean-square value of $V_A/\sqrt{2}$, the SNR is

$V_A/\sqrt{2}/\sqrt{kT/2C} = V_A/\sqrt{kT/C}$. Now both capacitors are disconnected and placed in series. The signal rms value doubles to $V_A\sqrt{2}$. The two noise contributions in series give $\sqrt{2kT/C}$, the same noise when sampling with a capacitor of value $C/2$, the equivalent capacitor for the series connection. The overall SNR is now: $V_A\sqrt{2}/\sqrt{2kT/C} = V_A/\sqrt{kT/C}$, the same as for the parallel connection.

How about all the switching needed for the rearrangement of capacitors? Practically that will indeed deteriorate the actual performance, but in theory, there is no charge moved, so no additional noise is picked up.

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 used as a switch depends on the gate-to-channel voltage minus the threshold voltage; see Sect. 2.5:

$$\begin{aligned} R_{\text{on,NMOS}} &= \frac{1}{(W/L)_N \beta_{N\square}(V_{\text{DD}} - V_{\text{in}} - V_{T,N})} \\ R_{\text{on,PMOS}} &= \frac{1}{(W/L)_P \beta_{P\square}(V_{\text{in}} - |V_{T,P}|)}. \end{aligned} \tag{4.3}$$

The NMOS switch is conductive if the input voltage is lower than the gate voltage minus the threshold voltage. The maximum gate voltage often equals the power supply voltage. The PMOS switch is conductive with input voltages more than a threshold voltage above the gate voltage (mostly zero or ground level). At low supply voltages and large signal excursions the voltage-dependent resistance of the switch can lead to aperture time differences causing 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 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:

$$R(V_{\text{in}}(t)) = R_0 + \frac{V_{\text{in}}(t)}{V_{\text{in,peak-peak}}}\Delta R.$$

With an input signal $V_{\text{in}}(t) = 0.5\ V_{\text{in,peak-peak}} \sin(\omega t)$ the current is mainly determined by the capacitor: $I(t) \approx \omega C 0.5\ V_{\text{in,peak-peak}} \cos(\omega t)$. The voltage drop over

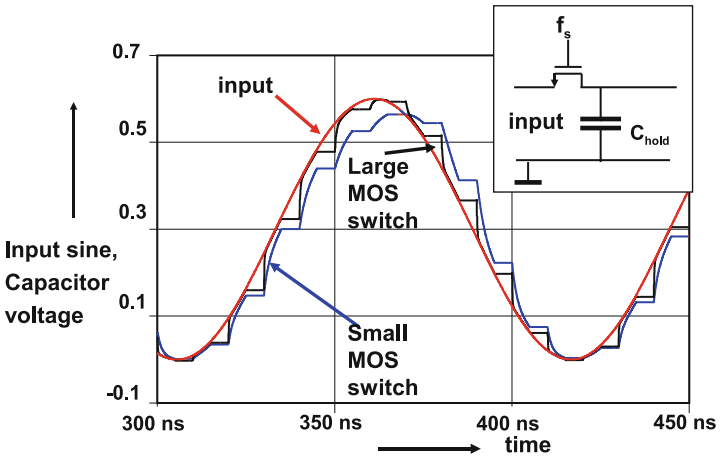


Fig. 4.8 The resistance of a small NMOS switch causes significant distortion in the hold signal

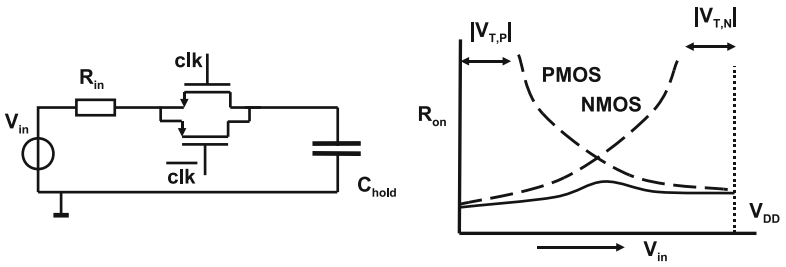


Fig. 4.9 The complementary PMOS–NMOS switch

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

$$\mathrm{HD2} = \frac{\omega \Delta RC}{4}. \tag{4.4}$$

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 other's 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

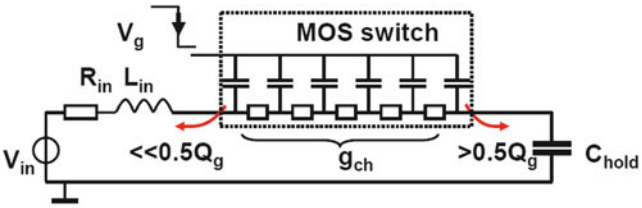


Fig. 4.10 The channel charge in a MOS device acts as a transmission line when the transistor is switched off. If the impedances on either side differ, the charge splitting is asymmetrical

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 $V_{T,N} + |V_{T,P}| < V_{\mathrm{DD}}$, an input region exists where none of the two transistors will conduct. Crols and Steyaert [93, 94] 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 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 $V_{\mathrm{DD}} - V_T - V_{\mathrm{in}}$, 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 an amount of signal charge and a constant charge term:

$$V_{\mathrm{hold}} = V_{\mathrm{in}} + \frac{Q_g}{2C_{\mathrm{hold}}} = V_{\mathrm{in}} + \frac{WLC_{\mathrm{ox}}(V_{\mathrm{in}} - V_{\mathrm{DC}})}{2C_{\mathrm{hold}}} = V_{\mathrm{in}}\left(1 + \frac{WLC_{\mathrm{ox}}}{2C_{\mathrm{hold}}}\right) - \frac{WLC_{\mathrm{ox}}V_{\mathrm{DC}}}{2C_{\mathrm{hold}}}. \tag{4.5}$$

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 contains much less charge and therefore charge splitting is less an issue.

When the switch turns to its isolating mode, 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

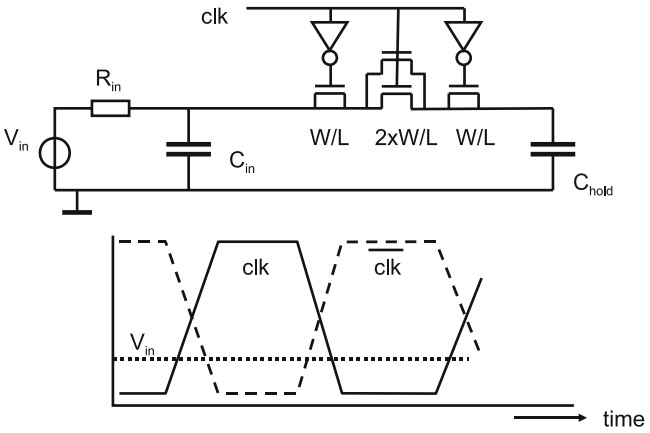
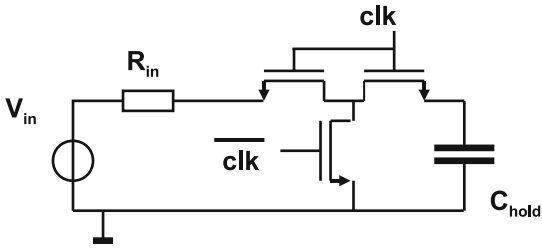


Fig. 4.11 The inherent charge splitting can be compensated by dummy switches

Fig. 4.12 The T-switch reduces the signal feed-through



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 matching of all parasitics and 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.

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

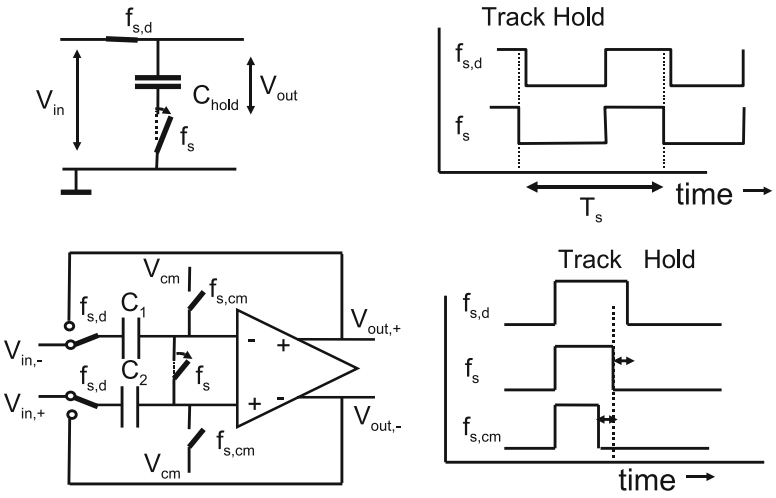


Fig. 4.13 Top: principle of bottom plate sampling, the switch to the bottom plate driven by f_s opens before $f_{s,d}$ disconnects the top plate. Bottom: Differential variant of bottom plate sampling. The switches in both schemes are drawn at the moment the bottom plate disconnects via f_s [95]

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.

Example 4.5. What will happen to a 10 MHz sine wave if the PMOS switch in Fig. 4.9 switches off 50 ps later than the NMOS switch?

Solution. A delay between the two switches will result in a time shift of the upper and lower part of the signal. In theory a complex signal processor can correct, but in practice the resulting sample sequence will show a distorted behavior. As the error is in the time domain, and resembles jitter behavior, a first-order calculation will give a rough estimate of the resulting distortion. THD$\approx 1/(2\pi f \Delta T) = -50$ dB.

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

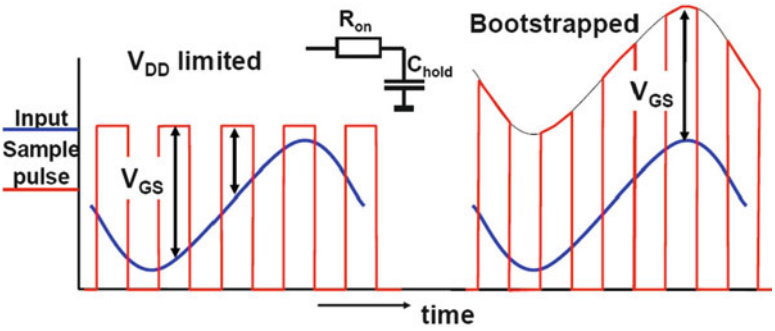


Fig. 4.14 The drive voltage of the switch varies if the power supply voltage is the limiting factor (*left*). If the switched is bootstrapped, ideally the drive voltage is constant over the input range

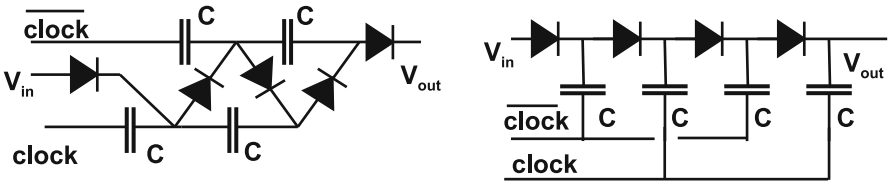


Fig. 4.15 Two techniques. *Left*: the series connected Cockcroft-Walton voltage multiplier (1932). *Right*: the parallel-connected Dickson multiplier [96]

parasitic capacitors connected to the bottom plate becomes a problem, this bottom-plate sampling will partly reduce the modulation problems with the sample switch.

4.3.4 CMOS Bootstrap Techniques

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 breakdown voltage. A limitation of this scheme is that the effectiveness

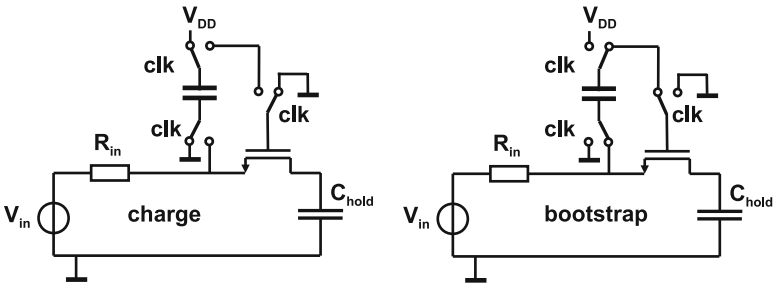


Fig. 4.16 The bootstrapping principle

depends on the ratio between the series connection of capacitors and the parasitic capacitors.

Dickson [96] 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

$$V_n = n\left(\frac{C}{C + C_{\text{par}}} V_{\text{clock}} - V_{\text{diode}}\right). \tag{4.6}$$

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 [97] 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 [98]. 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

[2] A relation seems likely with the tales on Baron von Mŭnchhausen, who pulled himself up by his bootstraps, and "booting" of computers.

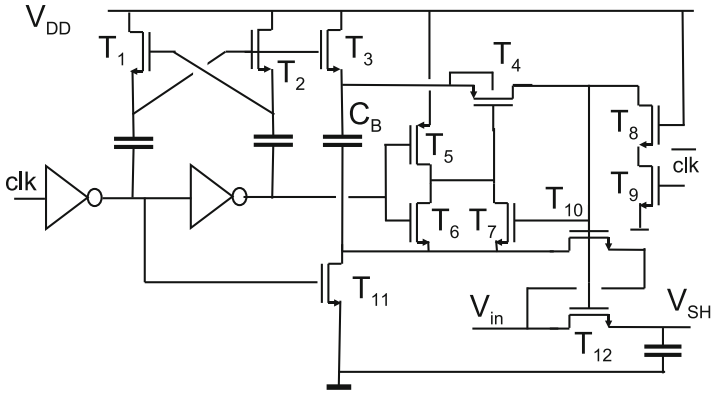


Fig. 4.17 The bootstrapping circuit according to Abbo [98]

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 $1.9 \times V_{DD}$. The requirements for 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 cannot 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. 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.

The previous bootstrap circuit can be simplified, see Fig. 4.18. The charge pump circuit has been removed and transistor T_3 has been changed into a PMOS device. As the *n*-well of this device must be connected to the highest bootstrapped voltage, extra loading of the bootstrap capacitor C_B will result.

4.3.5 Buffering the Hold Capacitor

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 [99] and degenerated differential pairs [56] are used. The inherent speed advantage must be traded off versus the nonlinearity. Still an 8–10 bit performance level is possible.

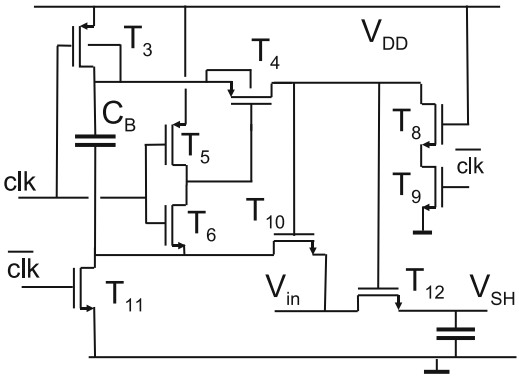


Fig. 4.18 A simplified bootstrapping circuit modified from [99]

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 lose 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 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

$$A_{\mathrm{DC}} > \frac{1}{\varepsilon} = 2^N, \tag{4.7}$$

  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 ($\mathrm{e}^{2\pi} = 527 \approx 2^9$), 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 [99, 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. 4.19. 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 in a protection circuit. A series resistor and some diffusion capacitor load the input 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. 4.20 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

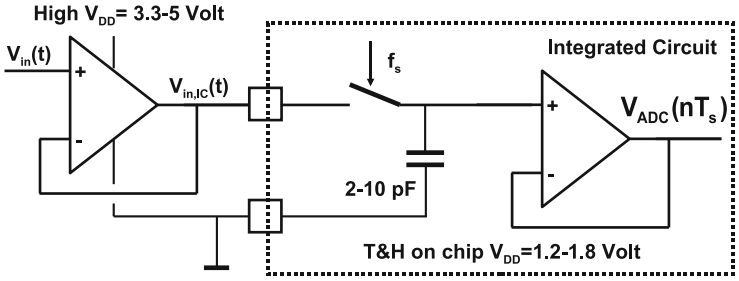


Fig. 4.19 An off-chip buffer drives the switch which is directly connected to the bondpad

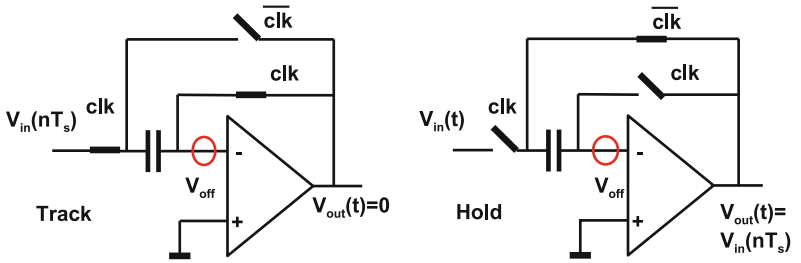


Fig. 4.20 A track-and-hold circuit with offset cancellation in track (*left*) and hold mode

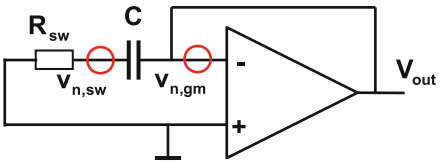


Fig. 4.21 The equivalent T&H circuit during track mode [100]

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

$$V_{\text{out,error}} = V_{\text{off}}(z)(1 - z^{-0.5}).$$

The exponent -0.5 assumes a 50 % duty cycle of the switching signal. The transfer is characterized in the frequency domain by a term

$$H(f) = 2\sin(\pi f/2f_s). \tag{4.8}$$

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 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. 2.5.10. As $1/g_m \gg R_{\text{sw}}$ and as the noise of the switch and sample capacitor spreads out over a much wider bandwidth, the opamp contribution is relevant. $v_{g_m,n}$ models this contribution in Fig. 4.21. The resulting output noise $v_{\text{out},n}$ 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 transistors in the input pair and the unity-gain transfer gives

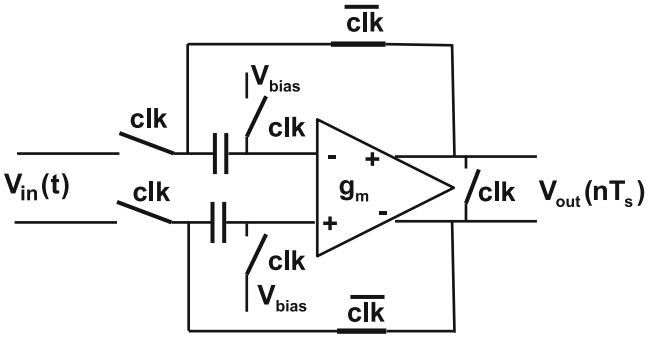


Fig. 4.22 A track-and-hold circuit based on a transconductance in differential design [101]

$$v_{\text{out},n}^2 = 2 \times \frac{4kT}{g_m} \int_{f=0}^{f=\infty} \frac{1}{1+(f/f_{\text{UGBW}})^2} \mathrm{d}f = \frac{4kT}{g_m} \frac{f_{\text{UGBW}}\pi}{2} \tag{4.9}$$

with the transconductance noise correction factor $\alpha = 1$. The factor two is caused by the sampling of the noise at the disconnect moment from the track mode and once more from sampling after the reconnect in the hold mode.

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 ($f_{\text{UGBW}} = g_m/2\pi C_{\text{load}}$ 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 4.22 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 feedback the capacitors are connected to the output

of the transconductance. This variant does not cancel the input offset; however the kT/C noise contributes only once and the noise of the transconductance during the track phase is not sampled.

Example 4.6. Design a track-and-hold circuit based on Fig. 4.20. The process is 0.18 µm CMOS with a nominal power supply of 1.8 V. 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 60-dB noise and accuracy is needed at the end of the sampling phase. Only NMOST switches can be used without bootstrapping. A generator with an impedance of 50 Ω generates a 0.4 $V_{peak-peak}$ signal, and any DC input level is allowed.

The NMOST has a threshold voltage of 0.4 V and a $\beta_\square = 350\,\mu A/V^2$. The PMOST has a threshold of 0.45 V and a $\beta_\square = 80\,\mu A/V^2$. Determine the size of the hold capacitor, the switches, and type of opamp.

Solution. The architecture of Fig. 4.20 limits the input voltage range to a maximum of $V_{in,max} = V_{DD} - V_{T,NMOST} - V_{drive}$. If the drive voltage is chosen at 0.2 V, the input will vary from 0.8–1.2 V.

The total noise floor is calculated from the input root-mean-square signal level $0.2V/\sqrt{(2)}$ and the required SNR: 60 dB or 1,000×, yielding 0.14 mV_{rms}. In the topology of Fig. 4.20 kT/C_{hold} noise is added during the disconnect from the input source. Also during the hold mode (time-continuous) noise is produced of a similar power level. For calculating the necessary hold capacitor a safety factor of 2 is used: 0.14/2 mV_{rms} results in $C_{hold} = 1$ pF.

At 10 MHz the track phase and the hold phase will last half a clock period or 50 ns. The settled value must be reached within 60 dB or $10^{-3} \approx e^{-7}$. Moreover, during the next track phase, the capacitor must be discharged to a similar level. So settling time constant τ equals 50 ns/7≈ 7 ns. This sets the requirement for the switches and the unity gain bandwidth of the opamp. With $R_{tot}C_{hold} = 7$ ns, $R_{tot} = 7\,k\Omega$. Choosing a maximum switch resistance of $R_{sw} = 1.5\,k\Omega$ leaves some margin: $R_{sw} = L/W\beta_\square V_{drive}$ results in $W/L = 1.8/0.18\,\mu m$.

The unity gain bandwidth of the opamp must exceed this speed level, so $2\pi f_{UGBW}\tau = 1$ and $f_{UGBW} > 20$ MHz. The settling level can only be achieved if the DC gain of the opamp is sufficient: $A_{DC} > 1{,}000$. These numbers can be achieved with a Miller opamp; see Sect. 2.7.11.

A simulation of the circuit with these values will still reveal some distortion caused by the switches. Larger switches and more drive voltages are needed.

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. 4.23. 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

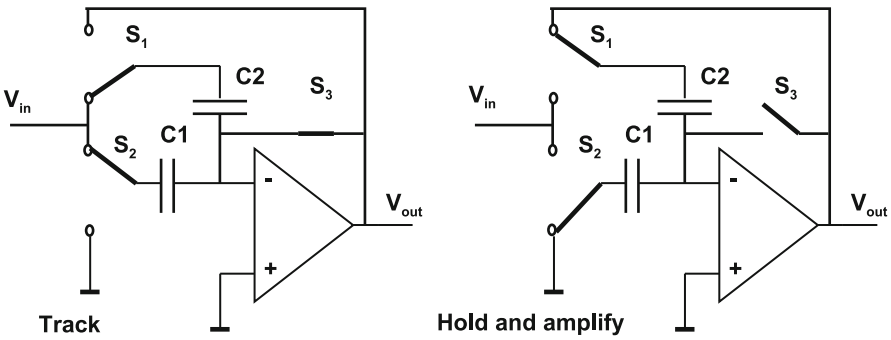


Fig. 4.23 A track-and-hold with multiplying capacitors as is used for pipeline converters

in parallel to the input signal. When the transition to the hold-phase occurs (right), switch S_1 creates via C_2 the feedback 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

$$V_{\text{out}} = \frac{V_{\text{in}}(C_1 + C_2)}{C_2}. \tag{4.10}$$

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. The 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. In some analog-to-digital converter architectures, the INL and DNL depends on the accuracy of the charge transfer. A limited DC gain will leave some charge on C_1 and impair this specification. Equation 4.7 sets a boundary condition.

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 $1/2\pi f_{\text{UGBW}}$. 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. 4.23 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.

For high-accuracy analog-to-digital converters the gain of the opamp must follow Eq. 4.7 to reduce incomplete charge transfer. In [102] the authors observe that the error voltage on the input of the opamp is equal to $V_{\text{out}}/A_{\text{DC}}$. Reducing V_{out} will reduce the error, and a low opamp gain can be tolerated. Figure 4.24 shows the basic

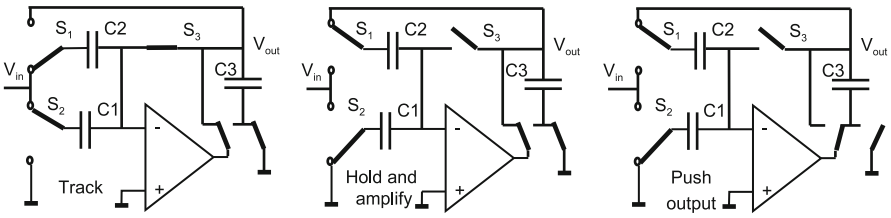


Fig. 4.24 An additional clock phase is used to reduce the requirements on opamp gain [102]

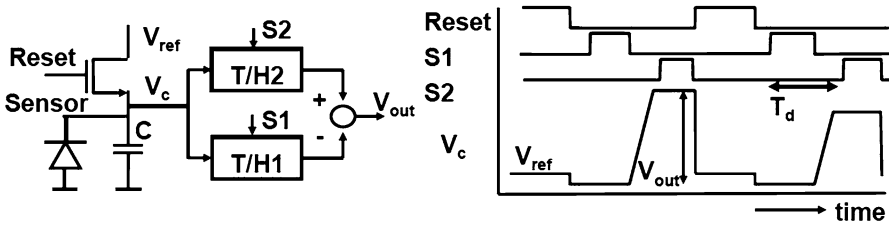


Fig. 4.25 The correlated double sampling technique

concept. Both the track and amplify phases are identical to the standard solution. At the end of these modes capacitor C_3 is charged to the output voltage. In a third phase of operation, this capacitor is switched in series with the output terminal of the opamp. The opamp output voltage is reduced to virtually zero and consequently also the input related DC-offset voltage. The charge transfer is now nearly perfect. This technique can also be applied inside a two-stage opamp, where it is inserted between the first and second stage, thereby allowing the output stage to deliver DC current. Also various multiplexing forms are conceivable.

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 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. 4.25). 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 4.25 shows the basic functionality. A single capacitor can be used for sampling both the "zero-signal"

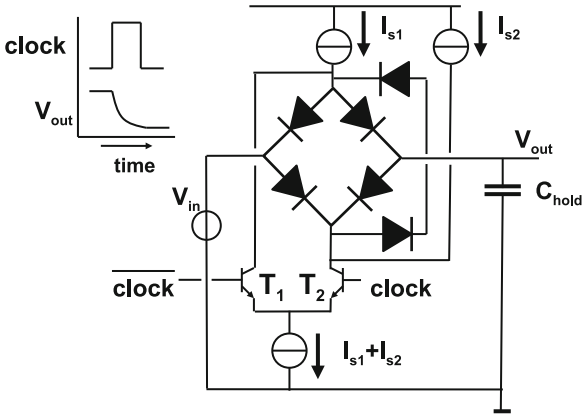


Fig. 4.26 A track-and-hold circuit with diode bridge [103]

and the wanted signal. The transfer function for the unwanted components is

$$|H(s)| = |1 - e^{-sT_d}| = |e^{-sT_d/2} 2 \sin(sT_d/2)|$$
$$|H(\omega)| = |2 \sin(\omega T_d/2)| = |2 \sin(\pi f T_d)| \quad (4.11)$$

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 Bipolar Examples

Bipolar design is still frequently used for analog realizations. In contrast to complementary MOS technology (CMOS), commercial bipolar processes are not always equipped with two complementary device types with similar performance levels. The design style is therefore focussed on passing the signal only via high-performance npn transistors.

Implementing a track-and-hold function with bipolar transistors requires a suitable switch topology. An often employed solution uses a diode bridge. Figure 4.26 shows a basic circuit. When T_2 is conductive the current I_{s1} flows through the four-diode bridge. Every diode will show a differential resistance of $2KT/qI_{s1}$ and the path from input to output sees a resistance of KT/qI_{s1}; see Eq. 2.78. With currents between 100 µA this corresponds to 250 Ω. In hold mode T_1 is active. The current I_{s1} is directly drained and the current I_{s2} flows via the two additional diodes. The bridge diodes are now reverse biased and nonconductive.

The nonlinear resistance in the conductive phase forces to reduce the signal swing. The leakage via the diodes is comparable to the leakage of diffusions in CMOS T&H circuits.

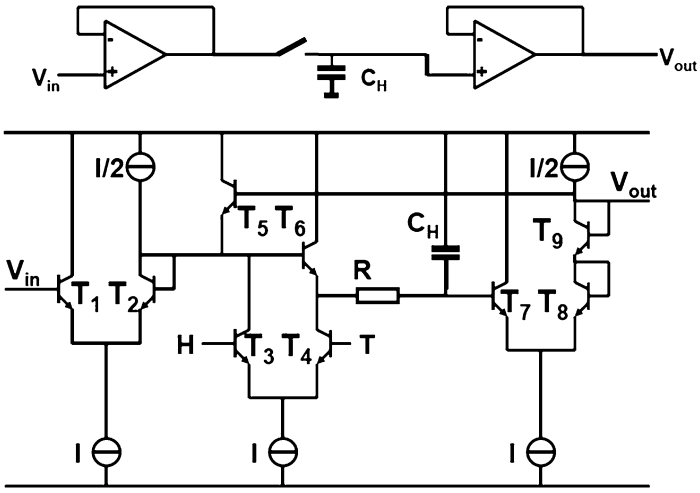


Fig. 4.27 A track-and-hold circuit in bipolar technology [104]

Vorenkamp and Verdaasdonk [104] published in 1992 an 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. 4.27 uses a pre-stage, the actual switch and capacitor, and an output buffer. T_1 and T_2 form a 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. This voltage is applied to the base of T_5 which consequently has a zero base-emitter voltage and is not conducting.

Similar to the diode bridge switch, this design also uses current rerouting to perform the switch function. 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 nonconductive. If this circuit is designed as a pseudo-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 pseudo-differential circuit [104] is a pair of cross connected compensation capacitors for reducing hold-mode feedthrough.

Example 4.7. Can the hold capacitor in Fig. 4.27 be connected to the negative supply rail?

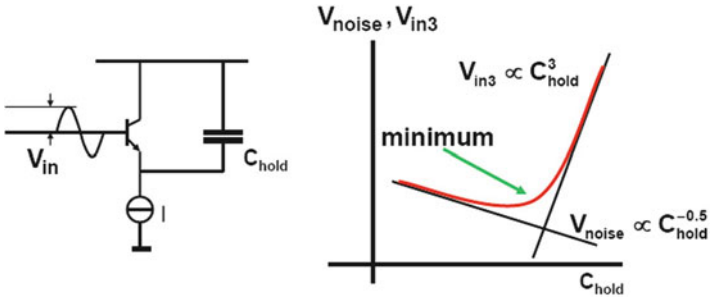


Fig. 4.28 This simple example shows that the hold capacitor cannot optimize the noise and distortion simultaneously

Solution. With ideal supply voltages it does not matter whether the hold capacitor is connected to the positive or negative supply. Here the positive rail is used as a signal reference. The power supply rejection for the negative power supply variations is now high because only current sources connect the circuit to the negative power supply. So all internal voltages are referenced to the positive power supply rail. Consequently the hold capacitor is referenced to that rail as well, and the impact of variations on the negative power supply will be suppressed. These variations would directly add up to the held voltage on the capacitor if that capacitor is connected to the negative supply.

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 Eq. 3.20. The distortion component can be calculated from the ideal circuit in Fig. 4.28. 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 $i_C = j\omega C V_a \sin(\omega t)$ will be taken away from the current in the bipolar transistor. This variable current will modulate the base-emitter voltage:

$$I - i_C = I_0 \mathrm{e}^{q(V_{\mathrm{BE}} - \Delta V_{\mathrm{BE}})/kT}.$$

Reversing the exponent to a logarithm and Taylor expansion for the first three terms gives

[3] A bipolar circuit is simple to analyze; of course, the same holds for a MOS circuit.

$$\Delta V_{\mathrm{BE}} = \frac{kT}{q}\left[\frac{i_C}{I} - \frac{1}{2}\left(\frac{i_C}{I}\right)^2 + \frac{1}{3}\left(\frac{i_C}{I}\right)^3\right].$$

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-order harmonics will however remain.

Substitution of the sine function leads to an estimate[4] of the fundamental and third-order distortion component:

$$v_{C,1} = V_a - \frac{kT}{q}\left(\frac{V_a j\omega C}{I}\right),$$
$$v_{C,3} = \frac{1}{12}\frac{kT}{q}\left(\frac{V_a \omega C}{I}\right)^3. \tag{4.12}$$

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 SNR. In the right-hand side of Fig. 4.28 the distortion term is compared to the noise contribution. For a given set of frequency, signal amplitude, and bias current parameters there will be an optimum capacitor value for a minimum combination of noise and distortion.

Example 4.8. If the third-order signal component of a sample-and-hold is given by $v_3 = AC_{\mathrm{hold}}^3$ and the noise signal level as $v_n = B/\sqrt{C_{\mathrm{hold}}}$, what is the optimum C_{hold} for best SINAD performance?

Solution. Combining both components in the power domain gives $v_{\mathrm{tot}} = \sqrt{(AC_{\mathrm{hold}}^3)^2 + (B/\sqrt{C_{\mathrm{hold}}})^2}$. Taking the derivative of v_{tot} with respect to C_{hold} and setting the result to 0 give the minimum value: $6A^2C_{\mathrm{hold}}^7 = B^2$ where either the ratio $A/B = \sqrt{6C_{\mathrm{hold}}^7}$ or the value of C_{hold} results.

Exercises

4.1. An FM signal of 100 kHz bandwidth at 10.7 MHz carrier is subsampled. Unfortunately a neighboring digital circuit running at 13.5 MHz will generate unwanted frequency components at 13.5 MHz and its second and third harmonic

[4] Contributions from higher-order terms are not included.

frequencies. Define a subsample frequency that maximizes the distance between the wanted and the spurious signals.

4.2. If the track-and-hold from Example 4.6 on p. 217 is used to de-glitch a digital-to-analog converter and drive in a time-continuous mode an analog circuit, what should change?

4.3. A simple T&H circuit as in Fig. 4.7 is designed in a 90-nm CMOS process with a switch transistor of dimensions $W/L = 5\,\mu m/0.1\,\mu m$. The hold capacitor is 100 fF and the input range is from 0 to 0.5 V. Give the overall amplification, the systematic offset, and the random offset of the capacitor voltage.

4.4. Discuss the impact of dummy switches in Fig. 4.11 on the random variation of the output offset voltage on the hold capacitor.

4.5. In an analog-to-digital converter the sampling speed is 250 Ms/s with an input range of $1.41\,V_{peak-peak}$. A performance of 90 dB is requested in a band from 30 to 40 MHz. The input sampling capacitor may contribute half of the noise energy. Calculate its size.

4.6. In sampling system the sample rate is 500 Ms/s with an input signal of range of $1\,V_{peak-peak}$ and a clock jitter of $1\,ps_{rms}$. A performance of 60 dB is required for all signals: the Nyquist baseband. Calculate its size.

4.7. Consider in Fig. 4.11 the case where the compensation pulse is too early and the case where it is too late. Although both cases are unwanted, which will be less harmful to the performance?

4.8. An FM signal of 100 kHz bandwidth at 10.7 MHz carrier is subsampled. Unfortunately a neighboring digital circuit running at 13.5 MHz will generate unwanted frequency components at 13.5 MHz and its second and third harmonic frequencies. Define a subsample frequency that maximizes the distance between the wanted and the spurious signals.

4.9. Use the technique in Fig. 4.24 between the first and second stage of the Miller opamp in Sect. 2.7.11. What is the optimum position and what input error reduction can be achieved?

4.10. A signal is sampled on two parallel connected capacitors: C_1 and C_2, with $C_2 = 2C_1$. After sampling the capacitors are connected in series. Does the SNR change between the parallel and series connection?

4.11. Modify Fig. 4.22 with a bottom-plate sampling technique.

4.12. A sine wave is sampled. There is a second-order distortion present in the sampling chain. Is it possible to distinguish from the resulting output spectrum, whether the distortion is affecting the original signal or appears after the sampling process?

4.13. Calculate in a 0.18 µm process the resistance of a complementary switch as in Fig. 4.9 with $W/L = 5/0.18\,\mu\text{m}$ at input voltages of 0, $0.5V_{\text{DD}}$, and V_{DD}. Use the data from Table 2.20.

Chapter 5
Quantization

Abstract Quantization is the second main process in conversion. This chapter deals with the mathematical derivation of quantization in several resolution ranges. Quantization results in several specific parameters: integral and differential linearities and derived problems such as monotonicity.

The signal-to-noise ratio is also affected by quantization. Some special topics are the effect of dither and the relation between differential nonlinearity and signal-to-noise.

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 [105]. 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.

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_5,

© Springer Science+Business Media, LLC 2013

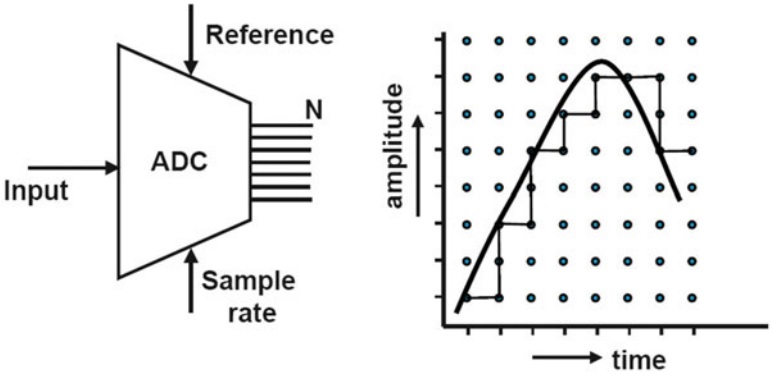


Fig. 5.1 An analog-to-digital converter rounds the amplitude of the continuous signal on the time-discrete sampling moments to the closest value of the discrete amplitude scale

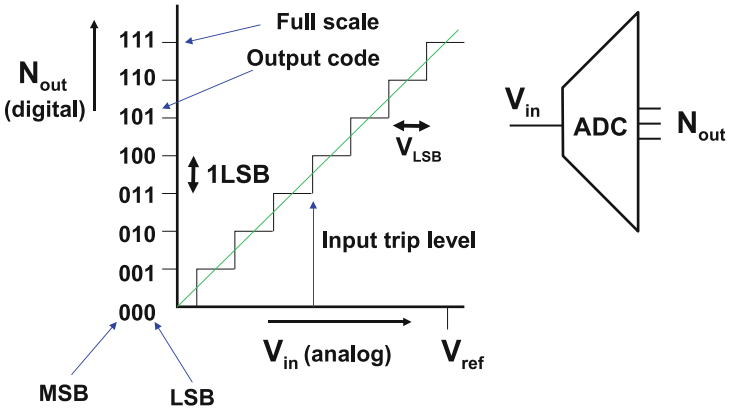


Fig. 5.2 Definition of analog-to-digital conversion parameters

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 [106] and IEEE 1241 [107, 108]. 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 ... 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 V 0.14, 0.24, ..., 0.74 V, 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 two 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:

$$\frac{\text{full scale}}{2^N} = \mathrm{LSB} \Leftrightarrow A_{\mathrm{LSB}} = \frac{\text{physical reference}}{2^N}, \tag{5.2}$$

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."

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 $A(i) = i \times A_{\mathrm{LSB}}$. The integral nonlinearity (INL) of a converter shows its deviation from the ideal conversion function:

$$\mathrm{INL} = \frac{A(i) - i \times A_{\mathrm{LSB}}}{A_{\mathrm{LSB}}}, \ \forall i = 0 \ldots (2^N - 1). \tag{5.3}$$

[1] Other digital code formats are discussed in Paragraph 7.1.1.

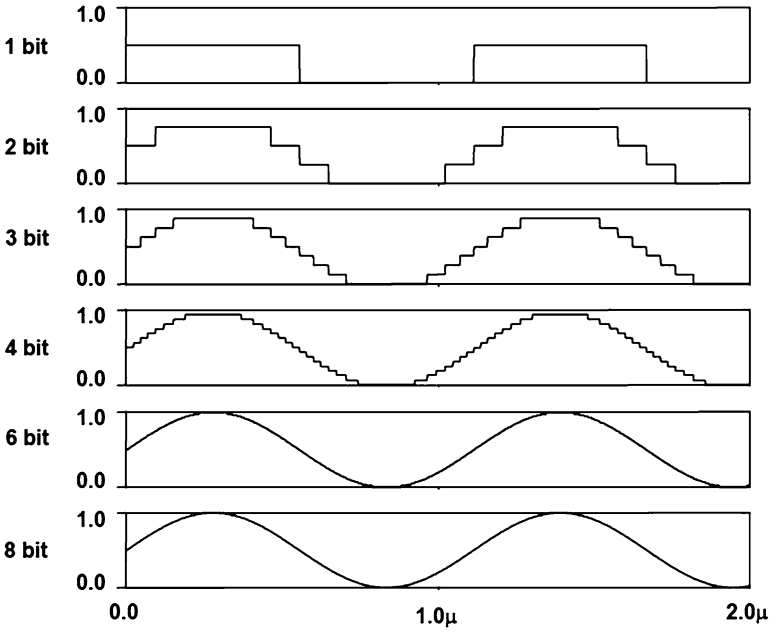


Fig. 5.3 Representation of a sine wave quantized with 1-,2-,3-,4-,6-, and 8-bit resolution

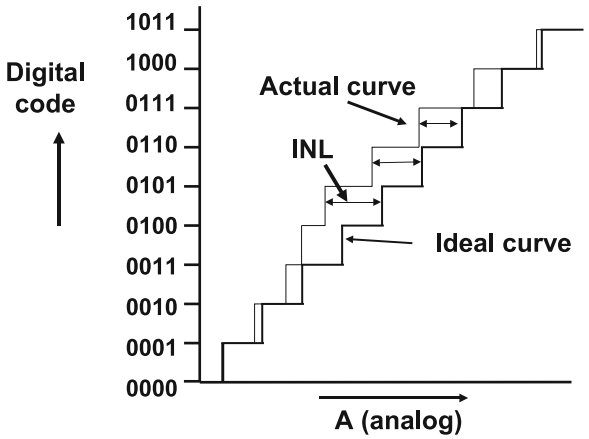


Fig. 5.4 Definition of the integral linearity error

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:

$$\mathrm{INL} = \max \left| \frac{A(i) - i \times A_{\mathrm{LSB}}}{A_{\mathrm{LSB}}} \right|, \ \forall i = 0 \ldots (2^N - 1). \tag{5.4}$$

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 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 not 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:

$$\mathrm{THD} = 10^{10} \log \left(\frac{\text{Power of second and higher harmonics}}{\text{Power of first harmonic}} \right). \tag{5.5}$$

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 differential nonlinearity (DNL) is the deviation of each step with respect to the ideal LSB size. The mathematical formulation is:

$$\mathrm{DNL} = \frac{A(i+1) - A(i)}{A_{\mathrm{LSB}}} - 1, \; \forall i = 0 \ldots (2^N - 2) \tag{5.6}$$

or as a single maximum number:

$$\mathrm{DNL} = \max \left| \frac{A(i+1) - A(i)}{A_{\mathrm{LSB}}} - 1 \right|, \; \forall i = 0 \ldots (2^N - 2). \tag{5.7}$$

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

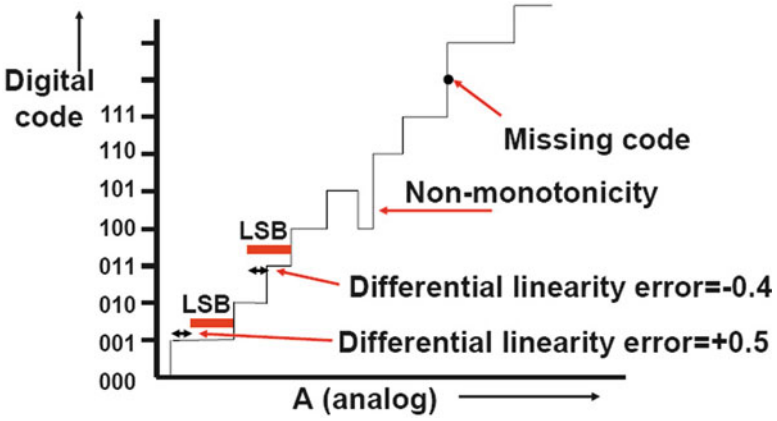


Fig. 5.5 Definition of the differential linearity

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 behavior can result in catastrophic problems if this converter is part of a control loop. Therefore some systems explicitly demand for monotonic behavior: an increase of the input level must lead to a zero or positive increase of the 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 sawtooth-shaped pattern in the INL curve indicates a subranging architecture, while the overall "sine" pattern is an indication that a third-order distortion is dominant (compare Fig. 2.3). Obviously a rigorous differential design has eliminated the second-order component, which would have resulted in a "half sine wave."

Example 5.1. Suppose an INL spec of 1 bit is given, is there a limit to the DNL spec?

Solution. The ideal INL curve is shown in Fig. 5.7 with on each side the curves corresponding to a deviation of $\pm$ 1 LSB. The maximum error is artificially created by pulling one trip level 1 LSB to the left and the neighboring trip level 1 LSB to the right. For clarity the shift is here only 0.9 LSB. The maximum DNL error where the INL stays within $\pm$ 1 LSB is therefore $+2$ LSB.

A more mathematical approach is to split the DNL formula in the difference of two INL's:

[2]Dutchmen often confuse monotonic with monotonous which means "vervelend" (=boring)

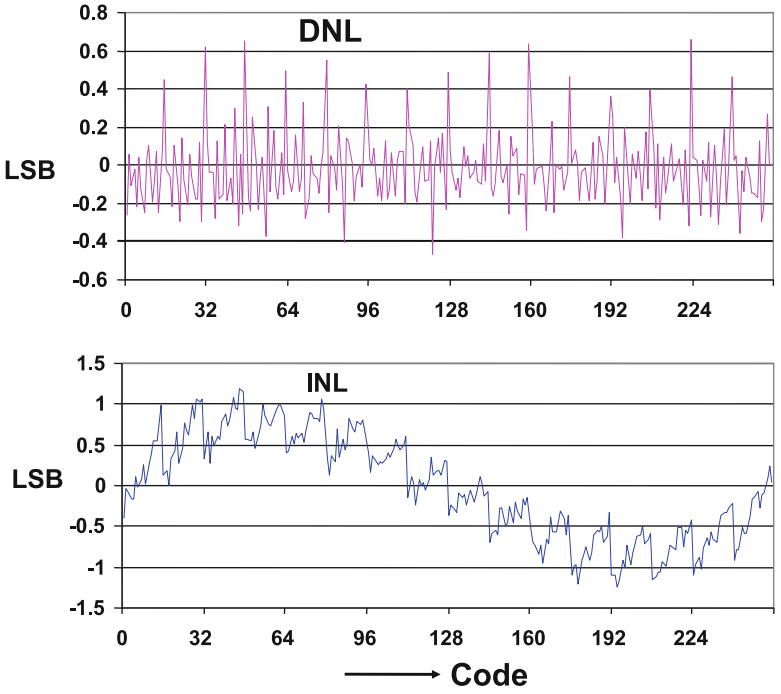


Fig. 5.6 An example of a DNL curve (*upper*) and an INL curve (*lower*) of an eight bit converter

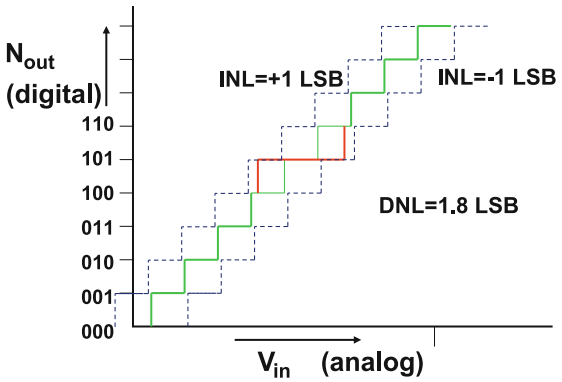


Fig. 5.7 INL curve with INL= +1 and INL= −1 curves

$$\begin{aligned}\mathrm{DNL} &= \frac{A(i+1)-A(i)}{A_{\mathrm{LSB}}} - 1 = \frac{A(i+1)-A(i)}{A_{\mathrm{LSB}}} - \frac{(i+1)\times A_{\mathrm{LSB}} - i\times A_{\mathrm{LSB}}}{A_{\mathrm{LSB}}} \\ &= \frac{A(i+1)-(i+1)\times A_{\mathrm{LSB}}}{A_{\mathrm{LSB}}} - \frac{A(i)-i\times A_{\mathrm{LSB}}}{A_{\mathrm{LSB}}} = \mathrm{INL}(i+1)-\mathrm{INL}(i) \\ &= (+1)-(-1) = 2.\end{aligned}$$

5.2 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 [109–111].[3] Figure 5.8 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 that were calculated in Sect. 2.1.2:

$$f(t) = \frac{4}{\pi}\sin(2\pi ft) + \frac{4}{3\pi}\sin(3 \times 2\pi ft) + \frac{4}{5\pi}\sin(5 \times 2\pi ft) + \cdots \tag{5.8}$$

and the total energy ratio between fundamental and harmonic components amount 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 quantized signal $\hat{A}\sin(2\pi ft)$, where A_{LSB} is equal to 1 [110]:

$$\begin{aligned} y(t) &= \sum_{p=1,3,5,..}^{\infty} A_p \sin(2\pi pft). \\ A_p &= \hat{A} \ \text{ for } p = 1, \\ A_p &= \sum_{n=1}^{\infty} \frac{2}{n\pi} J_p(2n\pi\hat{A}) \ \text{ for } p = 3,5,\ldots, \end{aligned} \tag{5.9}$$

[3] N. Blachman has mathematically analyzed many processes around quantization. His publications from 1960 to 1985 are a good starting point

[4] As the simulator for Fig. 5.8 requires to sample the signal, also alias components are visible

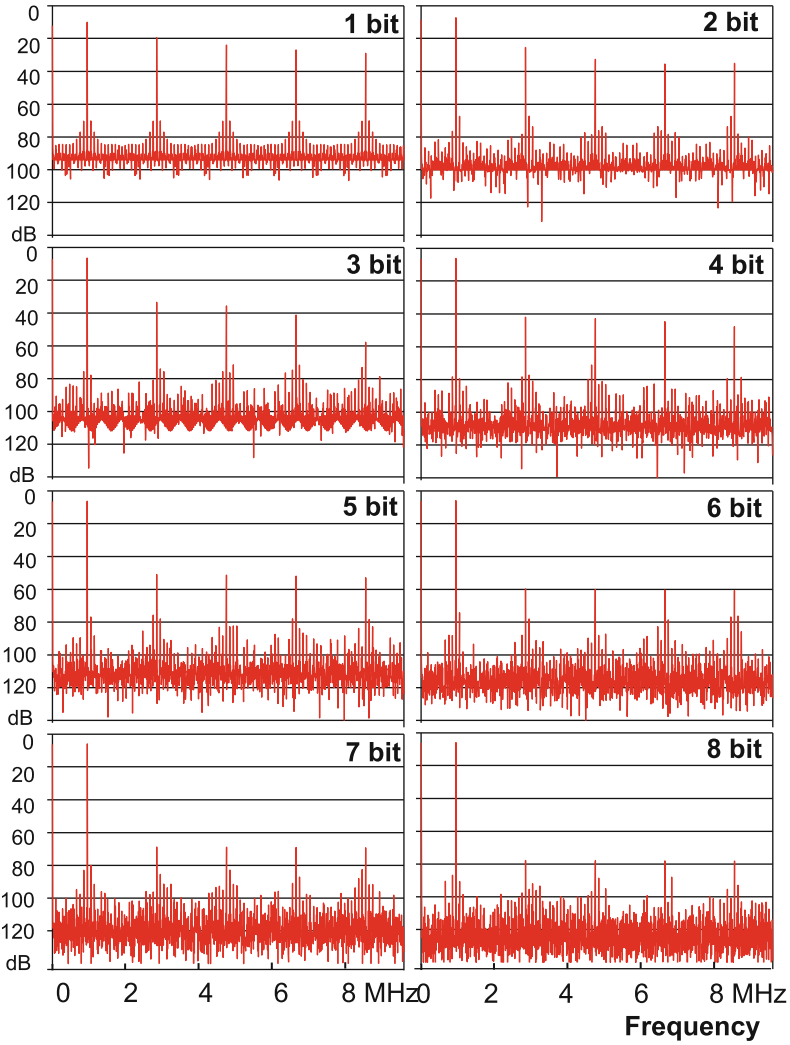


Fig. 5.8 The frequency spectrum of a 900 kHz signal quantized at 100 Ms/s with resolutions of $N = 1$, to $N = 8$

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:

$$A_p = \hat{A} \text{ for } p = 1,$$

$$A_p = (-1)^{(p-1)/2} \frac{h(\hat{A})}{\sqrt{\hat{A}}} \text{ for } p = 3, 5, \ldots,$$

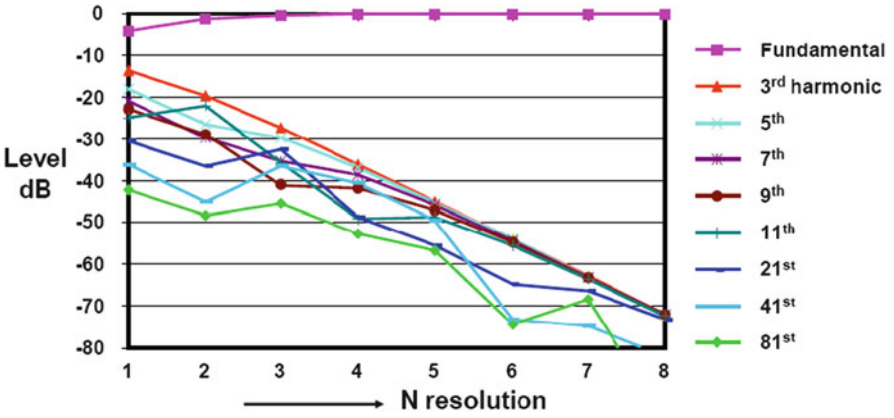


Fig. 5.9 The frequency components of a 900 kHz signal quantized at 100 Ms/s with resolutions of $N = 1$ to $N = 8$. The fundamental and the harmonic at 3,5,7,9,11,21,41, and 81 times the fundamental frequency are shown as a function of the quantization resolution. A decay of 8–9 dB/bit is visible

For $\hat{A} = 2^3, 2^4, 2^5, \ldots$ the value of $h(\hat{A})$ is more or less constant [110]. 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 [111] finds a value closer to 8 dB/bit.

Figure 5.9 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.

5.2.3 7-Bit and Higher Quantization

The quantization error is a nonlinear phenomena and an approximation can be used for quantization higher than 6-bit quantization 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 5.10 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 $+0.5$ 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 [112, 113] or the lowest error power is reached if the quantization levels are uniformly spaced. The power

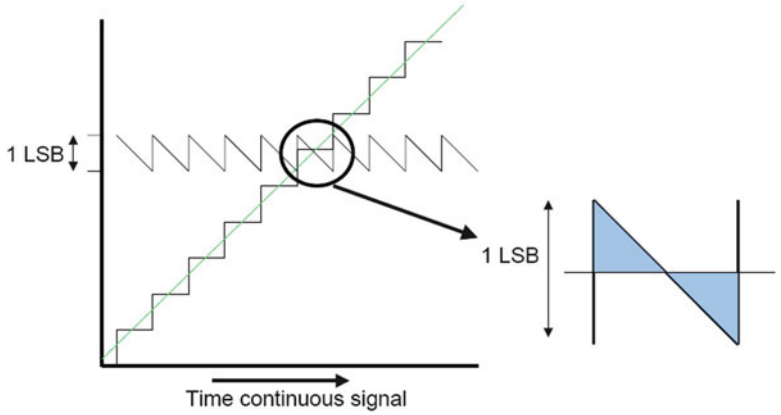


Fig. 5.10 As the analog-to-digital converter rounds the signal to the closest level, a triangular-shaped error signal will arise. This saw-tooth signal is the fundamental quantization error

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. 5.8). However, for a sufficiently large signal with a frequency that is not linked to the sample rate, 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.

The probability density of the error signal between -0.5 and $+0.5$ LSB is assumed constant and uniformly distributed. The power in that range, see Fig. 5.10, is determined by calculating the estimation value for the variance. The integration of the amplitude squared times the probability density function yields the equivalent power for the quantization. The quantization power now equals [109][5]

$$Q^2 = \frac{1}{A_{\mathrm{LSB}}} \int_{\varepsilon=-0.5A_{\mathrm{LSB}}}^{\varepsilon=0.5A_{\mathrm{LSB}}} A_{\mathrm{error}}^2(\varepsilon)\mathrm{d}\varepsilon = \frac{1}{A_{\mathrm{LSB}}} \int_{\varepsilon=-0.5A_{\mathrm{LSB}}}^{\varepsilon=0.5A_{\mathrm{LSB}}} \varepsilon^2 \mathrm{d}\varepsilon = \frac{A_{\mathrm{LSB}}^2}{12}. \tag{5.10}$$

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

[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.

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	Distortion 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 $\mu = 0$, $\sigma = v_{\text{rms}}$	Amplitude is uniformly distributed $[-A_{\text{LSB}}/2, A_{\text{LSB}}/2]$
$v_{\text{rms}} = \sqrt{4kTR\Delta\text{BW}} = 4\,\text{nV}$ @ $R = 1\,\text{k}\Omega, \Delta\text{BW} = 1\,\text{Hz}$	
$v_{\text{rms}} = \sqrt{kT/C} = 0.22\,\text{mV}$ @ $C = 1\,\text{pF}, \text{BW} = 0 \ldots f_s/2$	$A_{\text{rms}} = \dfrac{A_{\text{LSB}}}{\sqrt{12}}$ BW $= 0 \ldots f_s/2$

$$\text{Quantization energy} = \frac{A_{\text{LSB}}^2}{12} = \left(\frac{\text{full-scale}}{2^N\sqrt{12}}\right)^2, \tag{5.11}$$

where "full scale" is the analog range of the conversion and N the resolution.

Example 5.2. Compare the quantization "noise" with thermal noise.

Solution. Table 5.1 compares a few characteristics of thermal noise and quantization energy.

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 $-54\,\text{dB}$. In specific systems still special attention must be paid to the distortion side of the 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 wave-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:

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	

$$\text{SNR} = 10\,^{10}\log\left(\frac{\text{Signal power}}{\text{noise power}}\right) = 20\,^{10}\log\left(\frac{V_{\text{signal,rms}}}{V_{\text{noise,rms}}}\right). \tag{5.12}$$

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 SNR:

$$\text{Quantization power} = \frac{A_{\text{LSB}}^2}{12},$$

$$\text{Signal power} = \frac{1}{T}\int_{t=0}^{t=T} \hat{A}^2 \sin^2(\omega t)\mathrm{d}t = \frac{\hat{A}^2}{2} = \frac{2^{2N}A_{\text{LSB}}^2}{8}$$

$$\text{SNR} = \frac{\text{Signal power}}{\text{Quantisation power}} = \frac{3}{2}2^{2N}$$

$$\text{SNR} = 10\,^{10}\log\frac{3}{2}2^{2N} = 1.76 + N \times 6.02\,\text{dB}. \tag{5.13}$$

This last formula is an approximation that is often used in designing an analog-to-digital converter:

The maximum SNR represented by a digital word of N bits in a bandwidth of $f_s/2$ is $1.76 + 6.02 \times N$ dB.

The signal transported via 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 signa-to-quantization 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.:

$$\text{SINAD} = 10\,^{10}\log\left(\frac{\text{Power of first harmonic}}{\text{Power of all unwanted components}}\right). \tag{5.14}$$

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 Eq. 5.13:

$$\text{ENOB} = \frac{\text{SINAD} - 1.76}{6.02}. \tag{5.15}$$

Table 5.3 Combinations of resolutions and sample rates resulting in the same SNR

N	f_s
5	3,840 Ms/s
6	960 Ms/s
7	240 MS/s
8	60 Ms/s
9	15 Ms/s
10	3.75 Ms/s
11	2 Ms/s

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.

Example 5.3. What SNR can be reached by a 10-bit 50 Ms/s analog-to-digital converter in the bandwidth from DC to half of the sampling rate? And in a bandwidth between 2.0 and 2.5 MHz?

Solution. A converter of 10-bit resolution can reach a SNR in the full conversion band of 0 to 25 MHz of 62 dB.

In a bandwidth limited to 0.5 MHz, only the proportional fraction of the total quantization energy will appear: 0.5/25. In power ratios this factor corresponds to $^{10}\log 50 = 17$ dB. Therefore the signal-to-quantization error in a limited bandwidth is $\mathrm{SNR}_Q = 62 + 17 = 79$ dB.

Example 5.4. 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 full signal to quantization energy ratio in this band? 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?

Solution. The signal-to-quantization-noise of an ideal 8-bit quantizer is given by the following: $\mathrm{SNR} = 1.76 + 6.02 \times N = 49.9$ dB. The quantization energy is spread out evenly over the band from 0 to $f_s/2$ and then mirrored to higher bands. With $f_s = 60$ Ms/s the quantization energy is contained in a 30 MHz wide band. In a 1 MHz band there will a factor 30 less energy or in dB: $10^{10}\log(30) = 14.8$ dB. That brings the SNR in the 1 MHz band to 64.7 dB.

The generalized formula linking the sample rate to the SNR in 1 MHz is:

$$\mathrm{SNR}(1\,\mathrm{MHz}) = 1.76 + 6.02 \times N + 10^{10}\log\left(\frac{f_s/2}{1}\right).$$

Keeping the SNR fixed to 64.7 results in the values of Table 5.3. Note that for a resolution of 11 bit or higher the sample rate is fixed by the lower boundary of the Nyquist criterium.

5.3.2 Nonuniform 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 midpoint 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: $v_{\text{rms}} = \sqrt{v_1^2 + v_2^2 + \ldots}$; however, every now and then a time moment may occur where a significant number of carriers are in phase and their amplitudes add up: $v_{\text{top}} = v_1 + v_2 + \ldots$. 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 nonuniform quantization schemes, to improve the conversion quality at lower hardware costs, e.g., for normal amplitude distribution [113]. However, in practice, this means to design a specific analog-to-digital converter for each signal type. 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 $g(t)$:

$$g(t) = \text{sign}(f(t)) \frac{\ln(1 + \mu|f(t)|)}{\ln(1 + \mu)}, \tag{5.16}$$

where μ determines the amount of compression. It is obvious that after compression and analog-to-digital conversion a decompression step is required with the exact inverse function. Any intermediate signal processing must take the complex structure of the compressed signal into account. Other solutions implement floating-point architectures such as variable-gain amplifiers or two parallel analog-to-digital converters (see Sect. 8.10.6).

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 [114] see (Fig. 5.11). This dither signal has typically

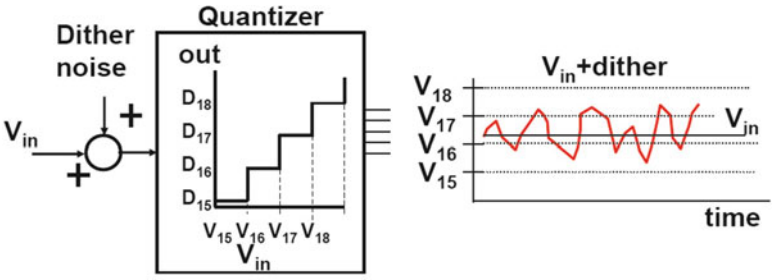


Fig. 5.11 The addition of a random signal allows to determine the value of a DC-signal at greater accuracy than the quantization process allows

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

$$V_{\mathrm{in}} = (B_{\mathrm{integer}} + Q_{\mathrm{fraction}}) \times V_{\mathrm{LSB}} \tag{5.17}$$

is composed of the quantized digital signal $B_{\mathrm{integer}} \times V_{\mathrm{LSB}}$ and the (lost) quantization fraction of the signal $Q_{\mathrm{fraction}} \times V_{\mathrm{LSB}}$. The addition of the uniform noise dither will also activate the quantization trip level at $B_{\mathrm{integer}} + 1$. 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

$$\begin{aligned} &\frac{(M - M_{\mathrm{fraction}})B_{\mathrm{integer}} + (M_{\mathrm{fraction}})(B_{\mathrm{integer}} + 1)}{M} \\ &= B_{\mathrm{integer}} + \frac{M_{\mathrm{fraction}}}{M} \approx B_{\mathrm{integer}} + Q_{\mathrm{fraction}}\big|_{M\to\infty}. \end{aligned} \tag{5.18}$$

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.

[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

5.4 Modeling INL and DNL

Figure 5.12 shows some typical shapes that can be found in an INL plot of a converter. Mostly a combination of elements is present with some elements more dominating than others. The upper side of the plot shows a second- and third-order distortion curve resulting from transfer curves as in Fig. 2.3.

The consequences of these INL plots on the signal transfer can be understood in a visual manner. When an input wave completes a full period, the INL curve is passed from left to right and then from right to left. So any input signal is multiplied by an (ideal) linear transfer characteristic and on top of that with two times the corresponding range of the INL curve. So an input sine wave passing through a converter with transfers as in Fig. 5.12 is multiplied by a patterns of a double repetition rate, leading to higher-order distortion products.

The starting point of a formal analysis is a transfer curve of the form $y = x + ax^2$ as in Sect. 2.1.3. The peak-to-peak amplitude of the INL curve for an analog-to-digital converter in Fig. 5.12 (upper, left) is approximated as $\mathrm{INL}_{\mathrm{peak-peak}} = aV_{\mathrm{ref}}^2/4$ in Volt or $aV_{\mathrm{ref}}2^N/4$ in LSBs. V_{ref} is the signal range. The derivation in Sect. 2.1.3 of a sine wave with an amplitude of $V_{\mathrm{ref}}/2$ resulted in a ratio between signal and second-order distortion component of $aV_{\mathrm{ref}}/4$. Measuring the second-order component in an INL plot therefore directly predicts the expected low-frequency distortion.

The sinusoidal-looking shapes in Fig. 5.12 (upper, right) and (lower, left) lead to an approximation $y = x + a\sin(2k\pi x/V_{\mathrm{ref}})$, where k is the number of full periods visible in the input range. Expanding this function via the Bessel function of Eq. 2.217 results in harmonic components in the output at frequencies $(k-1)f_{\mathrm{in}}$ and $(k+1)f_{\mathrm{in}}$.

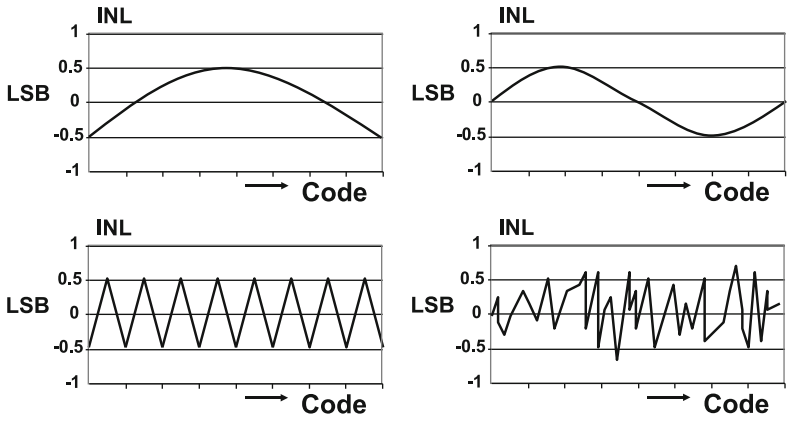


Fig. 5.12 Basic shapes of four typical INL patterns. *Upper* curves show patterns that are caused by second- and third-order distortion. The *left* lower pattern is often linked to sub-ranging architectures. The random pattern is typical for flash conversion

The transfer in Fig. 5.12 (lower, right) is typical for a flash converter and array-based digital-to-analog converters. Now the Integral and DNL represent, the nonuniformity 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:

$$A(i) = i \times A_{\mathrm{LSB}} + A_G(i), \tag{5.19}$$

where $A_G(i)$ is a sample from a Gaussian distribution with $\mu = 0$ and variance σ_A^2. 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 λ in the Gaussian probability function, for which the probability that $2^N - 1$ instances will remain within the interval (−DNL,+DNL) is acceptable see (Table 2.11). In a normal distribution λ is in the range $\lambda = 3.0\sigma - 3.4\sigma$ 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

$$\begin{aligned} \mathrm{SNR}_{\mathrm{Q+DNL}} &= 10\,{}^{10}\log 2^{(2N-3)} - 10\,{}^{10}\log\left(\frac{1}{12} + \frac{\sigma_A^2}{V_{\mathrm{LSB}}^2}\right) \\ &= 6.02N - 9.03 - 10\,{}^{10}\log\left(\frac{1}{12} + \frac{\mathrm{DNL}^2}{2\lambda^2}\right). \end{aligned} \tag{5.20}$$

For $n = 0$, DNL $= 0$, the formula results in the well-known SNR $= 6.02N +$ 1.76 dB. In Fig. 5.13 the SNR has been plotted in dB versus the DNL for $N = 8$ and $\lambda = 3.0$. The squares indicate the results of Monte Carlo computer simulations, 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 midrange. The triangles indicate measurement points.

Figure 5.13 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.

Example 5.5. A 10-bit ADC is not perfect: at the desired signal and sampling frequency, the DNL is 0.7 bit, while a second-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 ENOB of this ADC.

Advise whether the LSB can be removed, so that the succeeding processing runs with 9-bit samples.

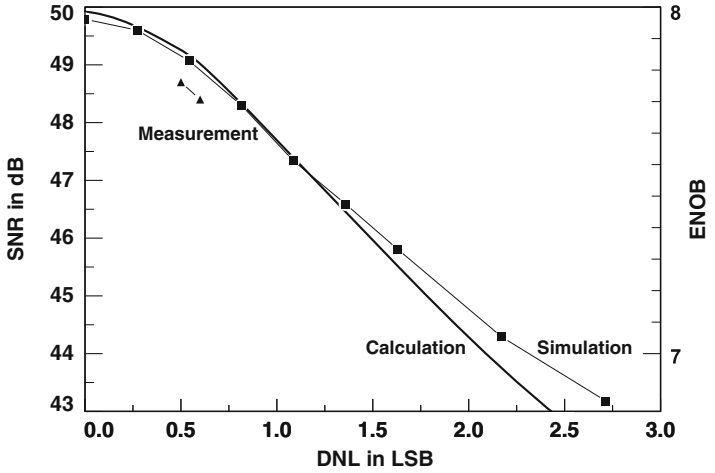


Fig. 5.13 Signal-to-noise ratio for quantization with Gaussian-distributed random offset errors versus the expected value of the DNL. The *squares* indicate Monte-Carlo simulations, the *triangles* refer to measurements obtained using the A/D converter described in Sect. 8.9

Table 5.4 List of powers of unwanted components

$\text{SNR}_{\text{Q+DNL}} = 60.7\,\text{dB}$	$\text{P}_{\text{Q+DNL}} = 10^{-60.7/10} P_{\text{sig}}$	$= 0.85 \times 10^{-6} P_{\text{sig}}$
$\text{THD} = -56\,\text{dB}$	$\text{P}_{\text{THD}} = 10^{-56/10} P_{\text{sig}}$	$= 2.5 \times 10^{-6} P_{\text{sig}}$
$\text{Tone} = 60\,\text{dB}$	$\text{P}_{\text{tone}} = 10^{-60/10} P_{\text{sig}}$	$= 1.0 \times 10^{-6} P_{\text{sig}}$
	P_{total}	$= 4.35 \times 10^{-6} P_{\text{sig}}$

Solution. Four components contribute to the loss of signal-to-noise performance: the quantization error, the DNL, the distortion, and the clock component. All these components must be added up in the power domain. First the combined effect of quantization and DNL is estimated with Eq. 5.20 and $\lambda = 3$:

$$\text{SNR}_{\text{Q+DNL}} = 6.02 \times 10\text{–}9.03 - 10\,^{10}\log\left(\frac{1}{12} + \frac{\text{DNL}^2}{2\lambda^2}\right) = 60.7\,\text{dB}.$$

Now all components can be related to the power of the signal, see Table 5.4.

So the SINAD=$^{10}\log(4.35 \times 10^{-6}) = 53.6\,\text{dB}$ or 8.61 ENOB.

The same ENOB can be achieved by a 9-bit converter. What happens after the 10-bit is removed? First, it will be assumed that the errors causing the DNL = 0.7 LSB will have the same absolute magnitude. Or on a 9-bit level the DNL = 0.35 LSB. Substituting this DNL in the equation gives

$$\text{SNR}_{\text{Q+DNL}} = 6.02 \times \mathbf{9} - 9.03 - 10\,^{10}\log\left(\frac{1}{12} + \frac{\mathbf{0.35}^2}{2\lambda^2}\right) = 55.6\,\text{dB}.$$

Now all components can again be related to the power of the signal, see Table 5.5.

Table 5.5 List of powers of unwanted components after the 10th bit is removed

$\text{SNR}_{\text{Q+DNL}} = \mathbf{55.6}\,\text{dB}$	$\text{P}_{\text{Q+DNL}} = 10^{-55.6/10} P_{\text{sig}}$	$= \mathbf{2.75} \times 10^{-6} P_{\text{sig}}$
$\text{THD} = -56\,\text{dB}$	$\text{P}_{\text{THD}} = 10^{-56/10} P_{\text{sig}}$	$= 2.5 \times 10^{-6} P_{\text{sig}}$
$\text{Tone} = 60\,\text{dB}$	$\text{P}_{\text{tone}} = 10^{-60/10} P_{\text{sig}}$	$= 1.0 \times 10^{-6} P_{\text{sig}}$
	P_{total}	$= \mathbf{6.25} \times 10^{-6} P_{\text{sig}}$

Removing the 10-th bit causes the SINAD to drop to $^{10}\log(6.25 \times 10^{-6}) = 52.0\,\text{dB}$ or 8.35 ENOB. Whether this is acceptable, depends on the application. The effect of the tenth bit will be more pronounced at lower input levels: For a 6-dB lower input signal the second-order distortion component will be 12 dB or 16× in power reduced, thereby changing the above balance considerably to the advantage of using the tenth bit.

Exercises

5.1. Suppose an DNL spec of 1 bit is given, is there a minimum or maximum limit to the INL spec?

5.2. The output signal of an FM intermediate frequency circuit has a bandwidth of 100 kHz at a frequency of 10.7 MHz. An analog-to-digital converter samples this signal at 5.35 Ms/s. What resolution is needed to obtain an SNR due to quantization of 14 bit in 100 kHz bandwidth?

5.3. A white-noise dither signal is applied with an in-band root-mean-square level of $0.289 \times V_{\text{LSB}}$. Give an expression for the best obtainable SNR.

5.4. A 14-bit ADC has to convert a 2 $\text{V}_{\text{peak–peak}}$ signal at a quality level corresponding to ENOB = 12.5 bit. The sampling capacitor is 1 pF. How much in-band distortion can you tolerate?

5.5. A SNR of 90 dB is required in a 20 kHz bandwidth. In the system a 650 kHz sampling rate is available and a 650 MHz rate. At Which combinations of resolution and sample rate can be used to achieve the desired SNR? Which converter has your preference?

5.6. What is the maximum spurious-free dynamic range of a perfect 8-bit analog-to-digital converter?

5.7. What is the maximum allowable INL to reach a spurious-free dynamic range of 80 dB with a 12-bit converter.

5.8. A sine wave is quantized with 2-bit. Is it possible to find an amplitude of the sine wave that will result in zero third-order distortion?

5.9. A sine wave of 10 kHz is sampled at 60 ks/s and quantized with 8-bit resolution. Make a drawing of the resulting waveform. Is this a favorable situation? What happens if the sine wave contains odd harmonics? Draw the spectrum.

5.10. An INL curve of a 10-bit digital-to-analog converter shows a value of -0.4 LSB at the ends of the range and at $+0.4$ LSB in the middle. Give an estimate for the second order distortion ratio.

5.11. Derive a relation between the INL and the third-order distortion in Fig. 5.12.

5.12. A full-range sine wave is quantized by an analog-to-digital converter with an INL curve as in Fig. 5.12 (lower, left). At what frequencies do you expect harmonic components? Now the sine wave is 3/4 range, at what frequencies will now harmonic components appear?

5.13. A 6-bit converter has a transfer characteristic described by $y = x + x(x - 0.25)(x-0.75)(x-1)$ where $x = 0 \ldots 1$. Determine the INL curve and the harmonics if this converter is driven by a signal $x(t) = 0.5 + 0.5\sin(\omega t)$. Repeat with a signal of $x(t) = 0.5 + 0.25\sin(\omega t)$.

Chapter 6
Reference Circuits

Abstract Every converter needs a reference quantity to link the numerical values on one side of the converter to a physical quantity on the other side. Several requirements are posed on the reference generator.

The most common reference is based on the bandgap energy of silicon. The general scheme for a bandgap circuit is extensively discussed as well as the limitations of the various components. The effect of mismatch in this circuit is described at the hand of an example. Finally a few implementations of low-voltage bandgap circuits are shown.

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 1 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

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_6,

© Springer Science+Business Media, LLC 2013

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 $-20\,^{\circ}\mathrm{C}$ 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 (pacemaker).
- Limited drift as a function of time. Aging of components is often related to temperature and overvoltage. 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. For example, 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 malfunctioning 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 feedback 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.

A reference circuit is based on some physical quantity. There are many physical quantities to chose from; see Table 6.1:

[1] Do not 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

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 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 °C to 70 °C. The bandgap reference idea was originally developed by Hilbiber [115] in 1964. Widlar [116], Kuijk [117], and Brokaw [118] refined the idea in the early 1970s.

Example 6.1. A diode with a temperature coefficient of $-2\,\mathrm{mV/°C}$ is supplied via a resistor with a current of 1 µA at 27 °C, linearly proportional to the absolute temperature (PTAT). At what value will the voltage over this combination show a temperature independent behavior?

Solution. The current can be written as $I_{\mathrm{PTAT}} = 10^{-6} \times \mathrm{Temp}/300$ [Ampere]. Consequently the voltage over the resistor R will show a positive temperature coefficient that must balance the diode behavior:

$$\frac{I_{\mathrm{PTAT}} R}{\mathrm{Temp}} = 2\,\mathrm{mV/°C},$$

resulting in a resistor value of 600 kΩ with a voltage drop of 0.6 V at room temperature. Summed with the diode forward voltage, the combination will show an overall voltage drop of approximately 1.25 V.

6.2 Bandgap Reference Circuits

6.2.1 Bandgap Principle

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 Eq. 2.79. The current over a pn-junction is analyzed in detail in Sects. 2.4.4 and 2.4.5 and is written as

$$I_{\mathrm{pn}} = C_0 T^{\eta} \mathrm{e}^{\frac{q(V_{\mathrm{pn}} - V_{\mathrm{bg}})}{kT}}. \tag{6.1}$$

All terms in Eq. 2.77 that have no temperature dependence have been collected in the term C_0. Note that $V_{\mathrm{pn}} < V_{\mathrm{bg}}$ 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 $-2\,\mathrm{mV/°C}$. The slope is mainly determined by the temperature dependence of the intrinsic carrier concentration n_i (see Eq. 2.64) 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 [115]: 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

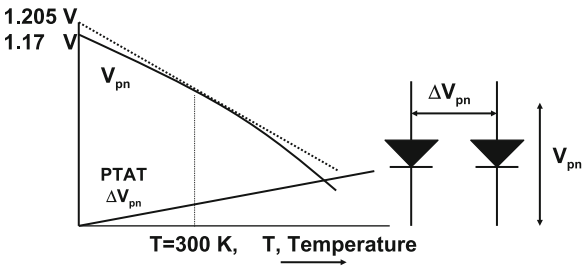


Fig. 6.1 The diode voltage V_{pn} decreases with increasing temperature at roughly 2 mV/°C. The differential voltage between two diode voltages ΔV_{pn} (e.g., created by different areas of the diodes) is proportional to the absolute temperature. By amplifying this last slope to a level where the differential voltage temperature coefficient is equal but opposite to the diode temperature coefficient, a temperature stable voltage can be obtained

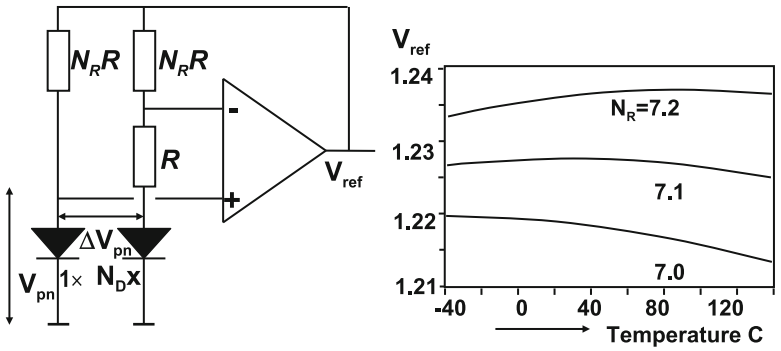


Fig. 6.2 The bandgap reference circuit [117]. The graph shows a simulation of this circuit with $N_D = 24$ and $N_R = 7.0, 7.1, 7.2$

the n-region aligns with the valence band and the Fermi level of the p-region aligns with the conduction band, so $V_{\text{pn}} = V_{\text{bg}}$.

The task of a bandgap circuit is to 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 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.

$$\Delta V_{\text{pn}} = V_{\text{pn},1} - V_{\text{pn},N_D} = \frac{kT}{q} \ln(N_D). \tag{6.2}$$

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 $N_R \Delta V_{\text{pn}}$ 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

$$V_{\text{ref}} = V_{\text{pn}} + N_R \Delta V_{\text{pn}}. \tag{6.3}$$

[2]Temperature dependence of the resistors or other components is canceled during the fine tuning of the simulation.

Of course the choice for N_R is essential for an exact 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:

$$V_{\mathrm{ref}}(T) = V_{\mathrm{pn}} + N_R \Delta V_{\mathrm{pn}} = V_{\mathrm{bg}} + \frac{kT}{q} \ln\left(\frac{I_{\mathrm{pn}}}{C_0 T^{\eta}}\right) + N_R \frac{kT}{q} \ln(N_D).$$

After differentiation for the temperature T and setting the result to zero, this maximum value is found at a temperature T_0:

$$\ln\left(\frac{I_{\mathrm{pn}}}{C_0 T_0^{\eta}}\right) - \eta + N_R \ln(N_D) = 0.$$

Substituting this result in the original formula gives

$$V_{\mathrm{ref}}(T) = V_{\mathrm{bg}} + \frac{kT}{q}\eta\left(1 - \ln\left(\frac{T}{T_0}\right)\right).$$

The maximum point in the curvature is found at T_0 with a value

$$V_{\mathrm{ref}}(T = T_0) = V_{\mathrm{bg}} + \frac{kT_0}{q}\eta.$$

The choice of N_R determines T_0 and the value of $V_{\mathrm{ref}}(T = T_0)$; see Fig. 6.2. Series expansion of the logarithmic function with an argument close to "1" requires to use $\ln(1+\alpha) = \alpha$ and yields

$$V_{\mathrm{ref}}(T) = V_{\mathrm{bg}} + \frac{kT_0}{q}\eta\left(1 - \left(\frac{T - T_0}{T_0}\right)^2\right). \tag{6.4}$$

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 $kT/q \approx 26\,\mathrm{mV}$. The typical output voltage is 1.23–1.28 V.

6.2.2 *Artifacts of the Standard Circuit*

The circuits in Figs. 6.2 and 6.5 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 need an additional start-up circuit to avoid this 0 V condition. A pull-up

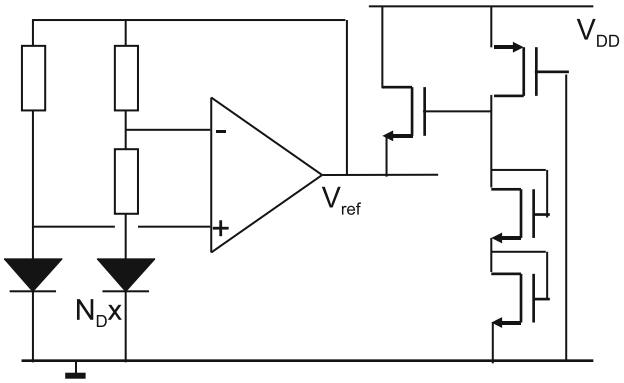


Fig. 6.3 A simple start-up circuit forces the bandgap circuit from the 0-V output. The transistor connected to the reference voltage will switch off if the reference voltage rises above one threshold voltage

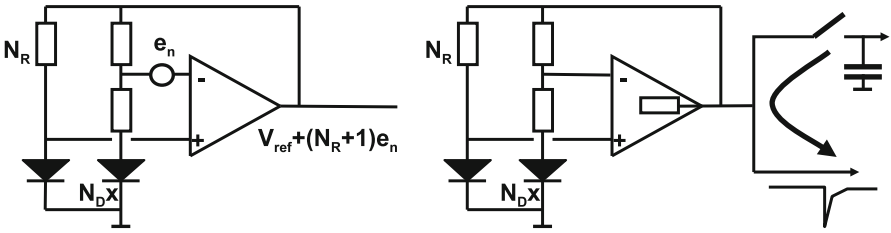


Fig. 6.4 A bandgap reference circuit amplifies unwanted input signals. If the output impedance is too high unwanted coupling will occur between connected circuits

resistor on the reference voltage is often sufficient. In Fig. 6.3 an NMOS transistor with the source connected to the output, the drain to the power supply, and the gate to a voltage generated by two stacked threshold voltages will become active if the reference voltage is below one threshold voltage. More elaborate schemes switch off the start-up after sufficient output voltage is detected.

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 operational amplifier (input-related circuit effects) will be multiplied with this factor; see Fig. 6.4 (left). The bandwidth of the operational amplifier can be reduced to suppress the 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 to 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.

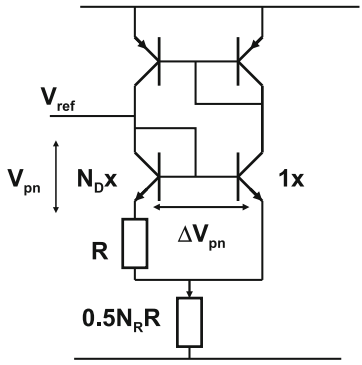


Fig. 6.5 A bandgap reference circuit with bipolar transistors

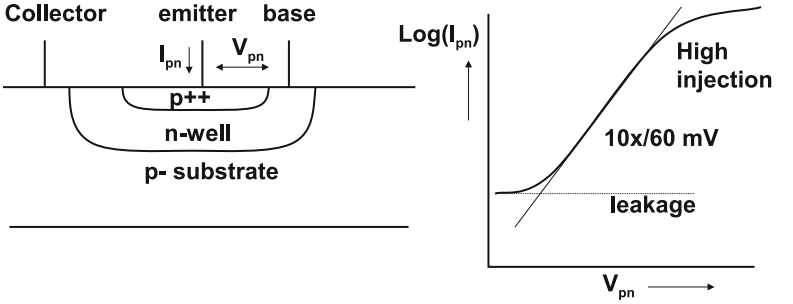


Fig. 6.6 The pn-diode or pnp transistor in a CMOS process and its I–V curve

6.2.3 Bipolar Bandgap Circuit

Only four bipolar transistors are needed to design a bandgap reference circuit [118]; see Fig. 6.5. 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., [119].

6.2.4 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 [120]; see Fig. 6.6. In fact this structure is a parasitic pnp transistor where the substrate serves as a collector; see Sect. 2.4.5. 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. 2.5.11. 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 cannot be large in this structure. High injection will cause deviation from the ideal I–V curve. On the lower side leakage limits the usable current range. Typical designs aim at a 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 in a CMOS process 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 2.12. These temperature coefficients need to be part of the resistor model in order to achieve a correct result on silicon.

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 6.7 shows the input stage for the operational amplifier of a 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 NMOS devices in inversion.

The input stage of Fig. 6.7 is build up as a folded cascode stage. As indicated in Fig. 6.4 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.

Next to these hand calculations it is well possible to use circuit simulation for estimating the input-referred mismatch voltage. Figure 6.8 shows a comparison of a

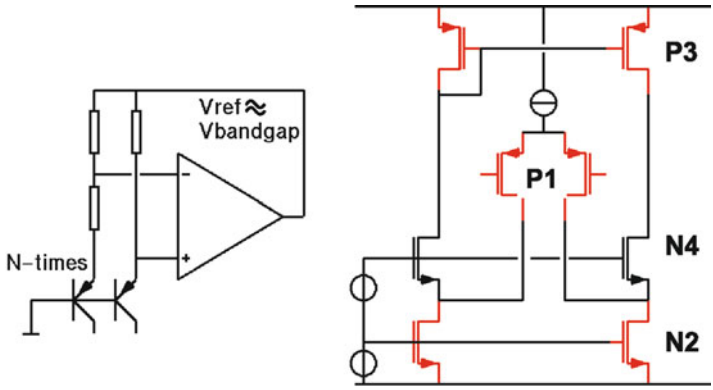


Fig. 6.7 The input stage of an operational amplifier for a bandgap reference circuit

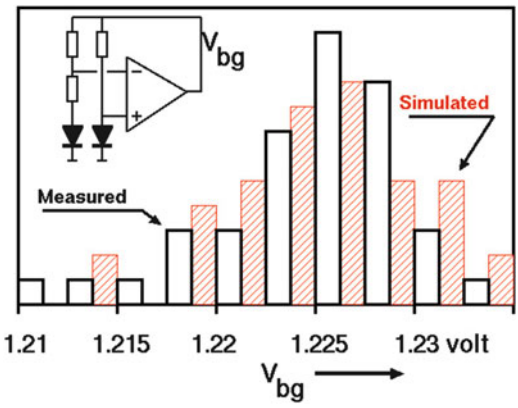


Fig. 6.8 Measured output voltage histogram of a bandgap reference circuit compared to simulation

typical simulation and a measured batch of reference circuits. Both the mean value and the standard deviation agree reasonably well for this sample size.

At higher temperatures the mismatch remains of equal importance.

Example 6.2. Calculate the input-referred mismatch of the design in Fig. 6.7, when the transistor dimensions are P1: 90/5, N2: 28/7, and P3: 60/4, with $A_{V_{Tn}} = 14.2\,\text{mV}\mu\text{m}$ and $A_{V_{Tp}} = 20.5\,\text{mV}\mu\text{m}$. In what way can the output variation of a group of bandgap circuits be reduced?

Solution. The input-referred mismatch of pair P1 is easily found as it equals the threshold-voltage mismatch of P1 using Eq. 2.110. The threshold-voltage mismatches of pairs N2 and P3 are found in the same way; however, their input-referred contribution needs to be calculated. The threshold-voltage mismatch of N2 translates in a current mismatch via its transconductance. 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

Table 6.2 Calculation of input-referred mismatch of the bandgap design, with $A_{V_{Tn}} = 14.2\,\text{mV}\mu\text{m}$, $A_{V_{Tp}} = 20.5\,\text{mV}\mu\text{m}$, $\mu_n = 2.5 \times \mu_p$

First design			Second design		
Transistor Dimensions	σ_{V_T} (mV)	σ_{Vin} (mV)	Transistor Dimensions	σ_{V_T} (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	Total input referred:		1.35
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

current an input-referred voltage source in the gate connection of P1 is needed. This is the input-referred mismatch voltage from N2, $\sigma_{V_{\text{in}},\text{N2}}$. This voltage is found from equating the mismatch current from N2 to the input-referred mismatch current in P1, via the transconductances of both devices:

$$i_{\text{mismatch,N2}} = g_{m,\text{N2}} \times \sigma_{V_{\text{T}},\text{N2}} = g_{m,\text{P1}} \times \sigma_{V_{\text{in}},\text{N2}}. \tag{6.5}$$

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 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 $N_R + 1$ 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 to reduce N_R from 10 to 6.8. The second design is summarized on the right hand side of Table 6.2.

6.2.5 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 [121] presented an elegant solution, Fig. 6.9. 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 $V_{\text{be}}/N_R R$ which has the negative temperature coefficient of the base-emitter junction. After

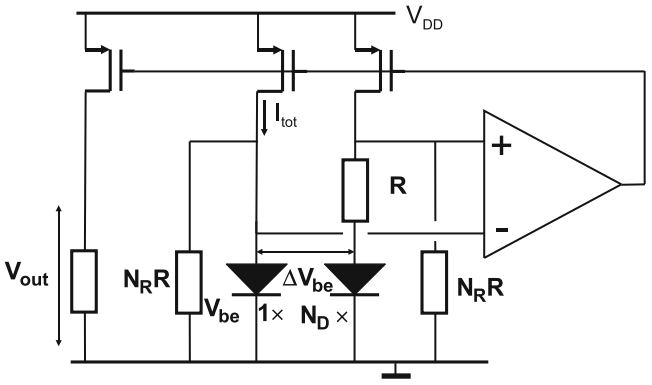


Fig. 6.9 A bandgap reference circuit for operation below 1.2 V [121]

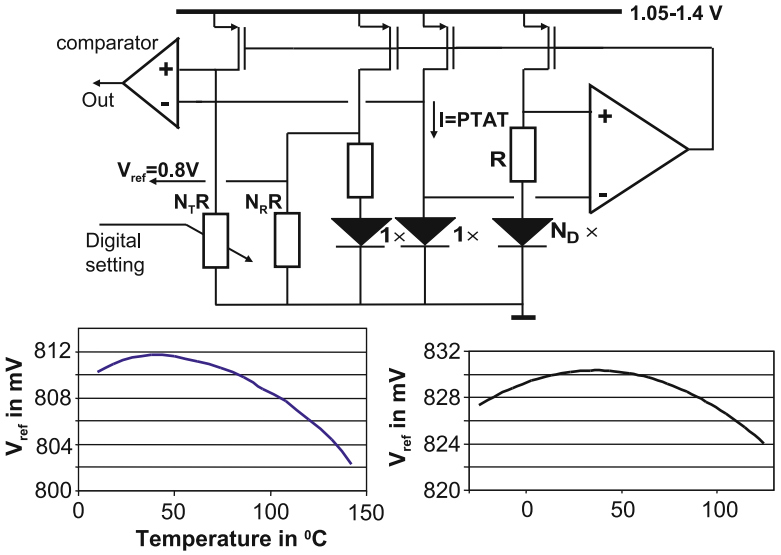


Fig. 6.10 A bandgap reference circuit of 0.8 V extended with temperature monitoring [122, 123]

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. 6.10). Proper tuning 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.

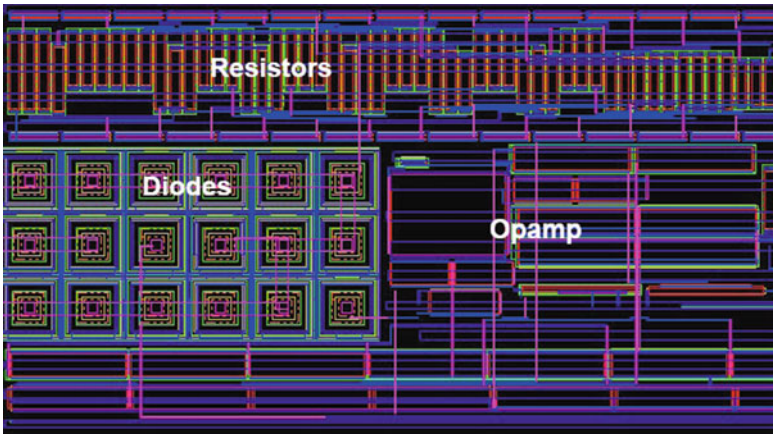


Fig. 6.11 The lay-out of a bandgap reference circuit in a 65-nm CMOS process

Figure 6.11 shows a portion of the layout of a bandgap reference circuit. The characteristic 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 [124].

Example 6.3. 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 %?

Solution. The diode ratio of 8 leads to a $\Delta V_{\mathrm{BE}} = (kT/q)\ln(8) = 53\,\mathrm{mV}$. With $V_{\mathrm{BE}} \approx 0.7\,\mathrm{V}$, the voltage over the resistor will be $V_{\mathrm{OUT}} - V_{\mathrm{BE}} \approx 0.6\,\mathrm{V}$. So the amplification ratio is in the range $0.6/0.053 \approx 11$. The maximum allowed output spread is $\sigma_{\mathrm{OUT}} = 0.01 \times V_{\mathrm{OUT}} \approx 12\,\mathrm{mV}$. This s.d. is due to the input-referred mismatch that is amplified by the opamp and resistor configuration. The input-referred mismatch must therefore be less than 12 mV/11=1.1 mV. As $A_{V_{\mathrm{T}}}$ for 0.18 µm technology is in the range of 5 mVµm, the minimum area of the input transistors is found from

$$\sigma_{V_{\mathrm{T}}} = \frac{A_{V_{\mathrm{T}}}}{WL} \rightarrow WL > 25\,\mu\mathrm{m}^2.$$

The topology of the opamp will determine whether the input-referred mismatch is determined by the input stage or whether more components play a role.

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.

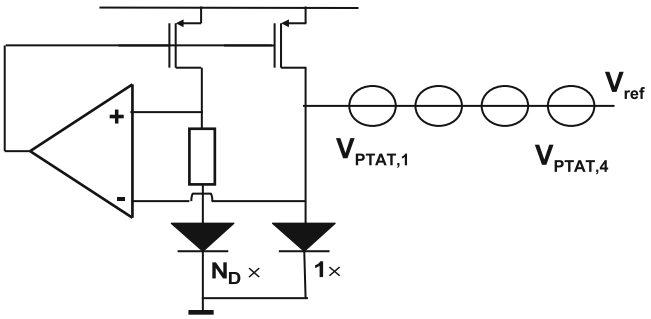


Fig. 6.12 A bandgap circuit with stacked PTAT sources [125]

- 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 form a series connection of several PTAT sources, Fig. 6.12. Instead of a linear multiplication of the error 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. This idea was followed in the design of a temperature-compensated current source [126]. Also circuits based on back-gate connected MOS devices (DTMOS transistor) are reported [125].
- 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

$$V_{\text{ref}} = V_{\text{T}} + \sqrt{\frac{2I}{\mu C_{\text{ox}} W/L}}. \tag{6.6}$$

  Both the threshold of the MOS transistor and the mobility show a considerably temperature coefficient [35]. 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 considered to be not reliable. Therefore some form of controlled compensation is needed, which is normally applied by making the current temperature dependent [127].
- The difference between two threshold voltages can be exploited in various ways. In the circuit of Fig. 6.13 (left) the W/L ratios and currents are equal. Now the output voltage is easily found by subtracting Eq. 6.6 for transistor T_2 from T_1. In first order the current-dependent part will cancel and $V_{\text{out}} = V_{\text{T},T_1} - V_{\text{T},T_2}$. 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 [128].

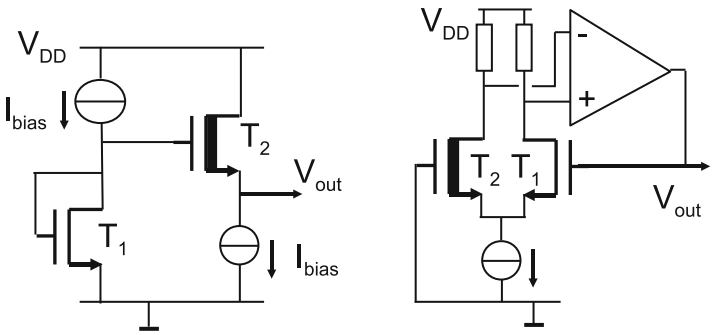


Fig. 6.13 Two ways of generating the difference between two threshold voltages of T_1 and T_2

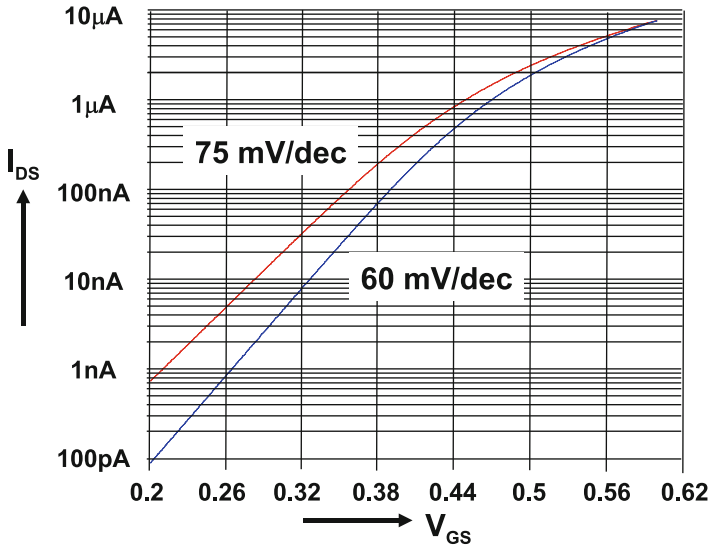


Fig. 6.14 A 4/1 µm PMOS transistor is operated in the weak inversion regime. The *upper curve* shows the current for an increasing gate voltage, the *lower curve* for a decreasing source voltage

- The threshold voltage of an EEPROM device can be trimmed during the programming process. This method and derived ideas require however some form of factory trimming.

Despite some interesting attempts these principles have never gained much acceptance compared to bandgap-based references.

Example 6.4. The subthreshold current in a MOS transistor shows an exponential relation as a function of the gate-source voltage. Is a MOS transistor in weak inversion a viable replacement for the diffusion-well-substrate pnp device?

Solution. The injection of charge into the channel at the source to channel junction follows a process described by the Boltzmann equation (see Sect. 2.4.4 and Eq. 2.77) and behaves like a pn-junction.

Unfortunately only the source side of this junction can be directly accessed. The potential in the channel is formed by the electrostatic balance between gate potential and local substrate voltage. In case the gate potential is changed with respect to the other terminals, an electrostatic division between gate capacitance and substrate capacitance will determine this channel potential. This division is represented by the factor m in Eq. 2.101. The weak-inversion current generated by an increasing gate voltage will follow a 70–80 mV/decade slope, as shown in Fig. 6.14.

On the other hand, also the source potential can be varied with respect to the gate and substrate voltages. Now there is no voltage division, and a perfect diode curve is found; see Fig. 6.14. This idea is followed in a DTMOS transistor [125], where gate and local substrate are tied together. It is also very well possible to have the gate and local substrate on different DC potentials.

It is therefore possible to replace the diode by a MOS transistor with a 60 mV/decade slope. However, in an absolute sense, the threshold voltage will still be part of the total voltage. This causes sensitivity to global and local variations and to specific temperature dependencies of the work functions.

Exercises

6.1. A process allows to design resistors with a global spread of $\pm$ 20 %. What will be the impact of this spread on the reproducibility of a standard bandgap circuit?

6.2. What bandwidth is advisable for the opamp in a standard bandgap circuit of Fig. 6.1 if this bandgap has to feed a 25 Ms/s switched capacitor digital-to-analog converter?

6.3. Discuss other solutions to the above problem: decoupling, a buffer circuit, and multiple references.

6.4. A start-up method in Fig. 6.6 uses a diode between the pnp-diode and the npn-diode. Discuss the advantages and disadvantages.

6.5. Propose a start-up method for the circuit in Fig. 6.2 other than indicated in the text.

6.6. The total area available for the opamp in a CMOS process with $A_{V_T} = 4\,\text{mV}\mu\text{m}$ is $40\times40\,\mu\text{m}^2$. Propose an opamp topology and give an estimate for the resulting reproducibly of the bandgap voltage.

6.7. Estimate for all bandgap circuits in this chapter what the power supply requirements are.

Chapter 7
Digital-to-Analog Conversion

Abstract The two most important architectures for constructing a digital-to-analog converter are the unary and binary approach. Both approaches have their merits. Next to the architecture, the second choice is the domain in which the converter is realized: voltage, current, charge, or time.

Realizations in all these domains are discussed and their specific behavior is analyzed. The resistor string is an important conversion element as it constitutes the digital-to-analog function in a flash converter. Its dynamic behavior is essential for reaching high-speed performance. The binary counterpart of this converter is the R-2R architecture. The current-steering topology is the dominant realization for fast stand-alone digital-to-analog conversion. The properties of this converter are described and analyzed.

Charge domain converters are mostly applied in lower-performance, low-power applications. Various topologies allow to choose between low area or better performance.

A special section is dedicated to error sources and methods to improve the performance. The dynamic element matching, current calibration and data weighted averaging methods are explained.

A number of examples details the design considerations and choices. Lay-out examples of commonly used structures are presented.

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.

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_7,

© Springer Science+Business Media, LLC 2013

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 mostly much less challenging.

The 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 digital-to-analog converters 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 a quantity corresponding to an LSB. 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 representations. Figure 7.1 shows both techniques.

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:

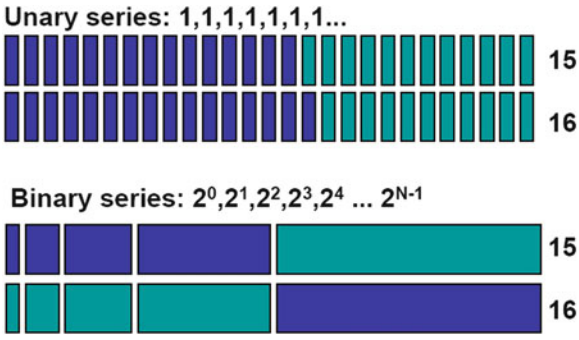


Fig. 7.1 Two basic techniques for digital–analog conversion: unary and binary representations. For both representations two signals are shown corresponding to a value of 15 and 16 LSBs

$$A = \sum_{i=0}^{i=2^N-1} b_i A_{\mathrm{LSB},i} = b_0 A_{\mathrm{LSB},0} + b_1 A_{\mathrm{LSB},1} + b_2 A_{\mathrm{LSB},2} + \cdots + b_{2^N-1} A_{\mathrm{LSB},2^N-1}. \tag{7.2}$$

A 1 LSB higher value can easily be created by adding one new element to the previous elements. The obvious advantage of this method is, that it 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:

$$B_b = \sum_{i=0}^{i=N-1} b_i 2^i = b_0 + b_1 2^1 + b_2 2^2 + \cdots + b_{N-1} 2^{N-1} \tag{7.3}$$

or in the physical domain:

$$A = \sum_{i=0}^{i=N-1} b_i A_{\mathrm{LSB}+i} = b_0 A_{\mathrm{LSB}} + b_1 A_{\mathrm{LSB}+1} + b_2 + \cdots + b_{N-1} A_{\mathrm{MSB}}. \tag{7.4}$$

As a switch has two positions, it is practical (but not necessary) to choose 2 as a base. The example in Fig. 7.1 shows a series of elements with the values: $A_{\mathrm{LSB}}, A_{\mathrm{LSB}+1}, \ldots, A_{\mathrm{MSB}}$= 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 lower part of Fig. 7.1). The implicit disadvantage of this method is the transition from one value to another value, at codes where many bits flip (e.g., 01111→10000). Although most transitions will result in a controlled LSB change, the midrange transition will cause the highest valued element to switch on and all other elements to switch off. Both values should differ the physical equivalent of one LSB; however, errors in the physical quantities $A_{\mathrm{LSB}}, A_{\mathrm{LSB}+1}, \ldots$ 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.

Binary-coded converters can be designed with low area and power consumption. But the problems with the transitions limit the performance. Another inherent potential issue with binary-coded circuits arises from the nonlinear relation between

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

the number of switched blocks and the output (e.g., for "15" four units are switched on, for "16" only one unit, for "17" and "18" two units.) Switching errors have no correlation with the code. Dynamic errors may occur as a result.

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 $2^{10}A_{\mathrm{LSB}}$. 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. 7.3 and 7.8.1

Most digital-to-analog (sub) schemes can be classified along the above architectural split. There are a few deviating forms such as ternary coding $(+1,0,-1)$, that are sometimes used in combination with sign/magnitude representation.

Many implementations have been reported in the older literature [129–134], illustrating the basic concepts.

7.1.1 Digital Representation

The unary and binary structures from the previous section 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 analog and digital processing. Table 7.1 shows 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" representation, the mid-level value is the virtual zero. 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 zeros and negative values with ones. 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. For example, 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 associated to this current. Obtaining a good signal-to-noise ratio for small signals is made difficult by the choice for two's-complement representation.

The "sign-and-magnitude" code is linked to the design style in analog hardware. The MSB of the code is the sign and the remaining part of the code is the amplitude of the signal. 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; see Fig. 7.2. In the case of rounding or truncation for a "straight binary" or "two's-complement" signal 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 toward zero. Such straightforward rounding or truncation will give a cross-over problem near zero and a distortion component.

[1] Many more representations exist; this table only lists the ones used in this book.

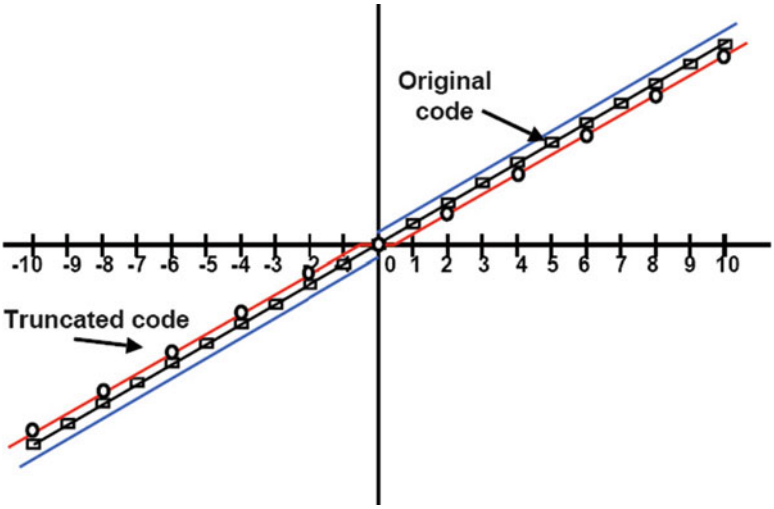


Fig. 7.2 Basic truncation or rounding creates distortion near the zero code

Table 7.2 Various forms of analog representation and physical domains

	Unary	Binary
Voltage	Resistor string *Flash ADC*	R-2R *Low-performance DAC*
Current	Current matrix *High bandwidth DAC*	Current splitting
Charge/ capacitor	Capacitor bank *Low-power DAC*	Capacitor bank
Time	PWM, SD mod *Low-bandwidth DAC*	Limited by distortion

In italic the main application area is indicated

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. 7.2.4), while in the current

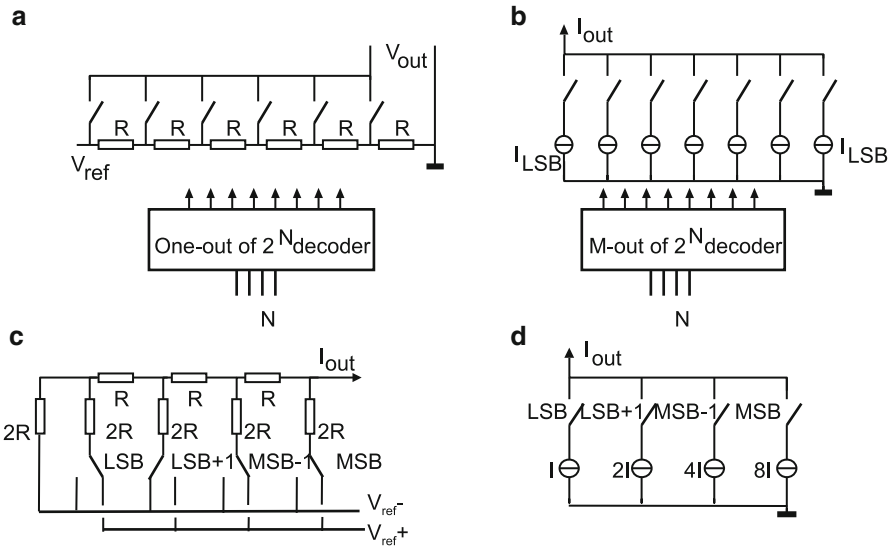


Fig. 7.3 Unary (*upper schemes*) and binary forms of resistor string and current source digital-to-analog conversion

domain currents can be split by means of transistor arrays. Capacitor arrays can use 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 brings.

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.

7.2 Digital-to-Analog Conversion in the Voltage Domain

7.2.1 Resistor Strings

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

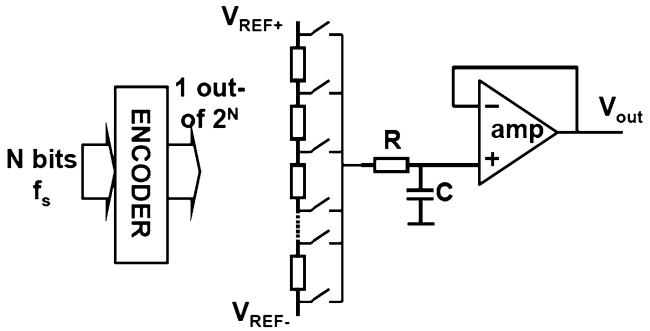


Fig. 7.4 A digital-to-analog converter based on a resistive divider

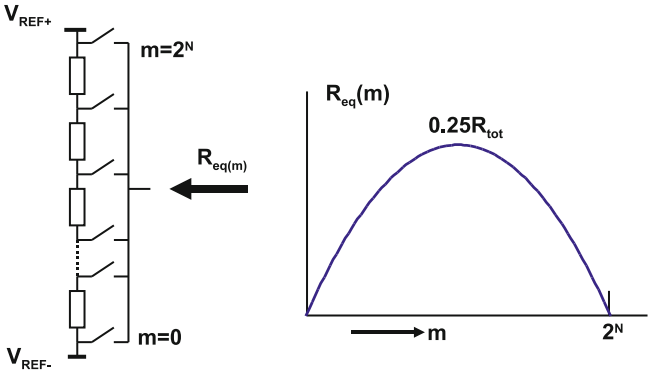


Fig. 7.5 The impedance of a resistor string various with the position

the position of the nodes in a ladder with $M = 2^N$ resistors, the impedance on each node is:

$$R_{\text{eq}}(m) = \frac{\frac{m}{M}R_{\text{tot}} \times \frac{M-m}{M}R_{\text{tot}}}{\frac{m}{M}R_{\text{tot}} + \frac{M-m}{M}R_{\text{tot}}} = \frac{m(M-m)}{M^2}R_{\text{tot}}. \tag{7.5}$$

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 needed by the ladder impedance to move charge will be position and signal-value dependent. High-frequency signals will show distortion.

The ladder impedance and its variation require 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.

7.2.2 Dynamic Behavior of the Resistor Ladder

The resistor string itself will show capacitive loading. This capacitive loading is distributed over the ladder. Excitations 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 resistance and capacitance 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:

$$\frac{\partial \quad \text{Temperature}(x,t)}{\partial t} = D\frac{\partial^2 \text{Temperature}(x,t)}{\partial x^2}, \tag{7.6}$$

where D is the thermal diffusion constant. This equation is used in 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

$$v(x,t) = \sum_{i=1}^{\infty} \mathrm{e}^{-i^2\pi^2 t/rcL^2} a_i \sin\left(\frac{i\pi x}{L}\right). \tag{7.8}$$

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:

$$v(x_m,t) \approx v_{\text{start}}(x_m)\mathrm{e}^{-\pi^2 t/rcL^2} \tag{7.10}$$

with a time constant: $\tau = rcL^2/\pi^2$.

[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.

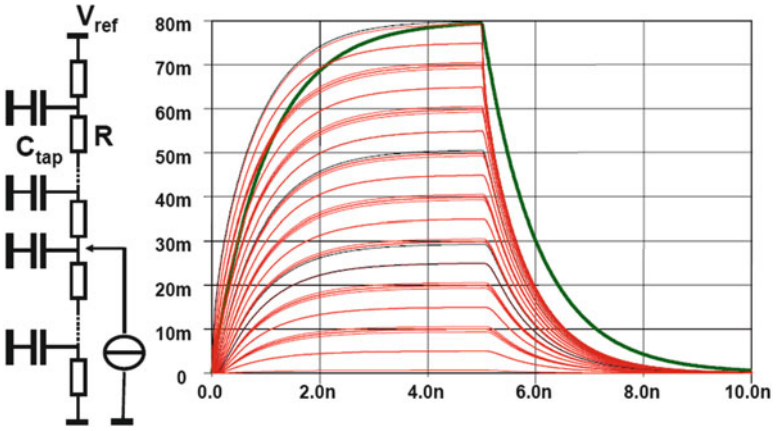


Fig. 7.6 In a ladder string with 256 resistors of 1.25 Ω and $C_{\text{tap}} = 0.1$ pF each, a current is injected at the middle tap and switched off after 5 ns. The *thin-line* plots show the time behavior at various nodes in the ladder. The *bold line* is the lumped RC approximation with $\tau = R_{\text{tot}}C_{\text{tot}}/8 = 80\,\Omega \times 12.8\,\text{pF}$

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 $R_{\text{tot}}/4 = rL/4$. 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: $\tau = R_{\text{tot}}C_{\text{tot}}/8 = rcL^2/8$, 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.

7.2.3 *Practical Issues in Resistor Ladders*

The actual design in the layout of a resistor ladder requires to consider various aspects. The available resistive material must show low random variation, a zero-voltage coefficient, and low parasitic capacitance. For high-speed conversion (>100 Ms/s) the time constant $\tau = R_{\text{tot}}C_{\text{tot}}$ must be low. As the total parasitic capacitance is significant, extremely low values for the tap resistances are necessary. Realizing tap resistances in the order of magnitude of 1 Ω poses considerable problems and requires sometimes to use special material layers on top of the chip.

In submicron CMOS processes a polysilicon layer has less parasitic capacitance and a lower temperature coefficient; see Table 2.16. A diffused layer gives a better

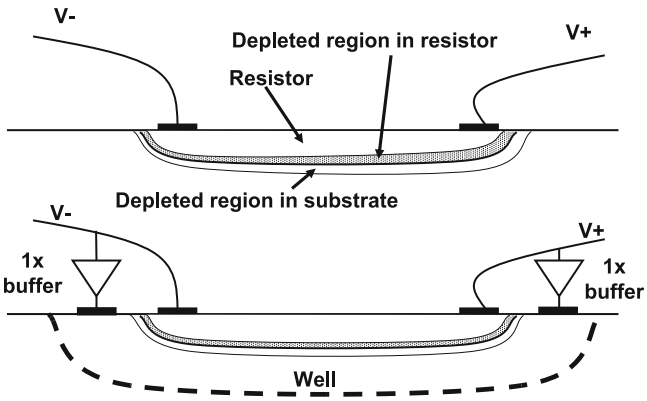


Fig. 7.7 Due to the voltage difference over a diffused resistor, the depth of the active resistor is less at higher voltages. Placing the resistor in a well of opposite dope and biasing the well with the same voltage difference avoids the voltage dependence

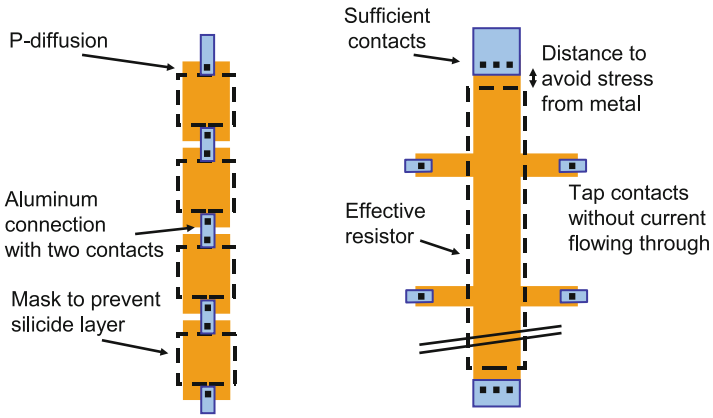


Fig. 7.8 On the *left-hand side* is a ladder built up with discrete resistors. On the *right-hand side* a ladder structure is shown for high accuracy; see Sect. 11.3

matching performance; see Table 11.9. 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 layout 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 current-supplying connections to the ladder are designed with sufficient contacts and are placed at some distance to reduce

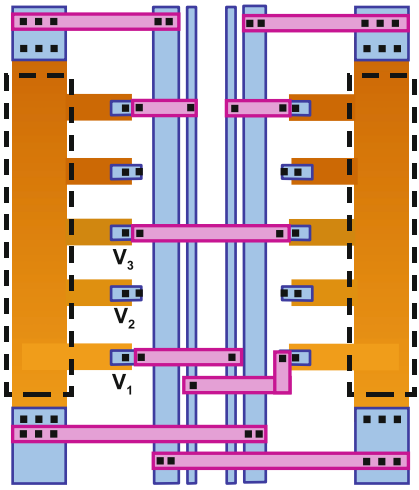


Fig. 7.9 A cross-coupled ladder structure

the material stress caused by the presence of aluminum wires; see also Fig. 11.20. 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. Gradients will create a non-linear voltage distribution over the ladder. If a ladder consists of 2^N resistors each nominally of value R with a gradient defined by a difference of ΔR between two adjacent resistors, then the ith resistor ($i = 1, \ldots, 2^N$) can be described as $R(i) = R + (i - 2^{N-1})\Delta R$. The current flowing through the ladder is constant, so the voltage on a node $m = 0, \ldots, 2^N$ can be described by a resistor ratio:

$$V(m) = \frac{\sum_{i=1}^{i=m} R + (i - 2^{N-1})\Delta R}{\sum_{i=1}^{i=2^N} R + (i - 2^{N-1})\Delta R} \approx m2^{-N}\left(1 - 2^{N-1}\frac{\Delta R}{R}\right) + m^2 2^{-N-1}\frac{\Delta R}{R} \quad (7.11)$$

and $V(0) = 0$. This formula is of the form: $y = x + ax^2$. Therefore the distortion and INL calculations of Sects. 2.1.3 and 5.4 apply.

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 technological background.

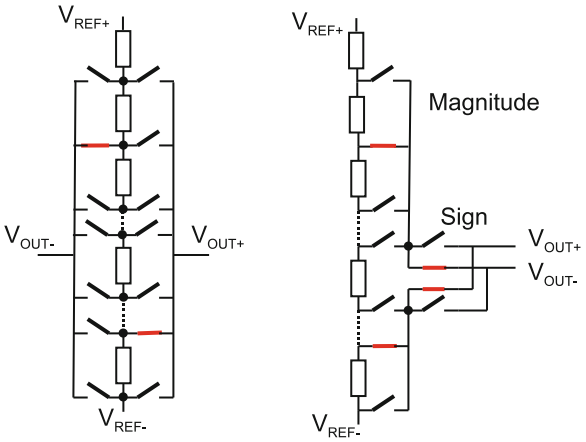


Fig. 7.10 A differential ladder structure and a sign-and-magnitude topology

An alternative way to eliminate a first-order gradient is to use a differential signal. Figure 7.10 shows two examples: a ladder structure with full-differential decoding and a structure based on a sign-and-magnitude decoder.

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 application, a careful design of the buffer is required; see Sect. 7.8.1. This topology allows an excellent static performance at a limited power level. The speed is limited by the performance of the buffer.

Example 7.1. Sixty-five 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 at one extreme
– The return line is connected to ground at both extremes
– The return line is connected to ground in the middle

Solution. This problem bears similarity to the calculation of the time constant on the distributed *RC* ladder. Starting at the open end of the return wire, it is clear that the current source causes a voltage drop IR over the first segment of the return wire. The voltage drop over the second segment is double, over the third segment triple, etc. The total voltage drop is therefore: $(1+2+3+\cdots+n)$IR. The sum of this series is found in Table 2.2: IR$n(n+1)/2$. With the data above the voltage drop at the open end is 20.8 mV.

If the open end is also grounded, the maximum voltage will appear in the middle. The middle current source contributes half to both sides, so the voltage drop is: $32 \times 0.5\text{IR} + \sum_{i=1}^{i=31} \text{IR} = 5.12$ mV.

One ground connection in the middle exactly splits the first problem in two, so at both open ends, a voltage appears of $\sum_{i=1}^{i=32} \text{IR} = 5.28$ mV.

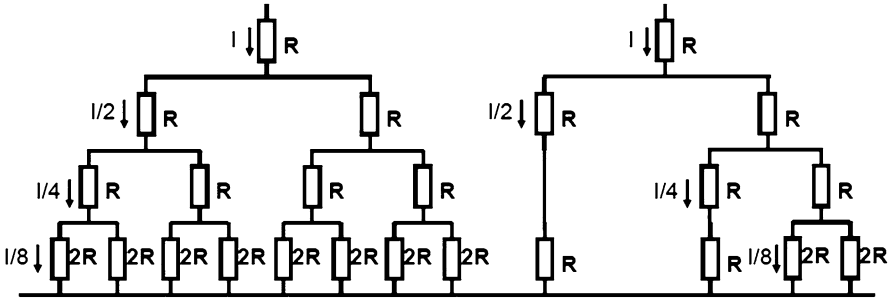


Fig. 7.11 Currents can be split in binary-weighted portions (*left*). The constant impedance in each branch allows to reduce the branches of the structure

Example 7.2. Show that a perfect linear gradient in Fig. 7.10 is canceled by the differential readout of a single ladder.

Solution. If $i = 0 \ldots (2^N - 1)$ is the digital code, then a perfect ladder will produce a voltage on every resistor node: $V_i = V_{\mathrm{REF}-} + iV_{\mathrm{LSB}}$, where $V_{\mathrm{LSB}} = 2^{-N}(V_{\mathrm{REF}+} - V_{\mathrm{REF}-})$. A linear gradient means that in the direction of increasing resistivity every next LSB is slightly (ΔV_{LSB}) larger. Now the voltage on every node is $V_i = V_{\mathrm{REF}-} + i(V_{\mathrm{LSB}} + i\Delta V_{\mathrm{LSB}})/S$. The term between brackets must be scaled back with a factor $S = (1 + 2^N \Delta V_{\mathrm{LSB}}/V_{\mathrm{LSB}})$ to fit to the reference voltages.

Now a differential output voltage can be formed by choosing the node voltage connected to i and the complementary code $2^N - i$. This gives us a differential output voltage:

$$\begin{aligned} V_{\mathrm{OUT}+} - V_{\mathrm{OUT}-} &= V_{\mathrm{REF}-} + i(V_{\mathrm{LSB}} + i\Delta V_{\mathrm{LSB}})/S \\ &\quad -(V_{\mathrm{REF}-} + (2^N - i)(V_{\mathrm{LSB}} + (2^N - i)\Delta V_{\mathrm{LSB}})/S) \\ &= (2iV_{\mathrm{LSB}} + 2i2^N \Delta V_{\mathrm{LSB}} - 2^N V_{\mathrm{LSB}} - 2^{2N} \Delta V_{\mathrm{LSB}})/S. \end{aligned}$$

There are only linear signal components (proportional to i) left in the input signal. Note that the common mode signal $V_{\mathrm{OUT}+} + V_{\mathrm{OUT}-}$ contains the second-order distortion term. So the common mode rejection of the succeeding stages is important.

7.2.4 R-2R Ladders

Figure 7.11 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

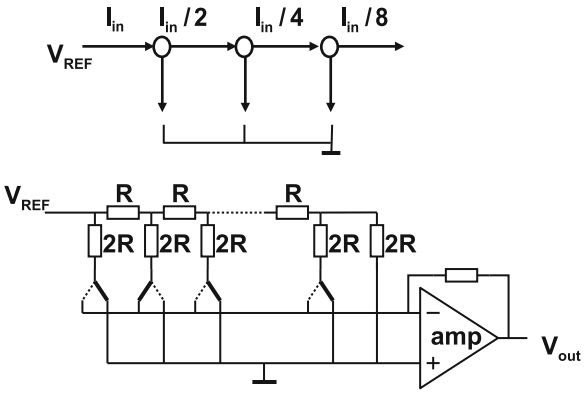


Fig. 7.12 R-2R digital-analog converter with digital encoding

Fig. 7.11 (right-hand side) the branches by resistors R and thereby constructing the $R-2R$ structure.[3] Figure 7.12 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 generates 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. 7.12 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 a R value.

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 (for a DNL<1) to 7 bit. The application is limited to low-resolution low-cost applications such as offset correction.

Example 7.3. In an R-2R ladder each resistor has an uncorrelated standard deviation of 1% of its value. If a digital-to-analog converter is constructed with these resistors, how many bits of resolution can be designed without running into monotonicity problems?

[3] Explanation from Colin Lyden (ADI).

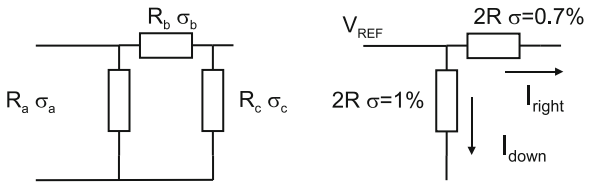


Fig. 7.13 *Left*: basic section of R-2R ladder, *right*: current splitting at the first node

Solution. Examine a basic section of the R-2R ladder in Fig. 7.13. With Eq. 2.20 this yields for the variance of $(R_b + R_c)$: $\sigma_b^2 + \sigma_c^2$. The equation for the parallel resistors then yields

$$\sigma_{\text{tot}}^2 = \left(\frac{R_b + R_c}{R_a + R_b + R_c}\right)^4 \sigma_a^2 + \left(\frac{R_a}{R_a + R_b + R_c}\right)^4 (\sigma_b^2 + \sigma_c^2).$$

An R-2R ladder is designed with $R_a = 2R, R_b = R_c = R$ and $R_{\text{tot}} = R$, which reduces the above equation to $\sigma_{\text{tot}}^2 = (\sigma_a^2 + \sigma_b^2 + \sigma_c^2)/16$. The observation that replacing R_c in the basic schematic by the same three-resistor scheme allows to expand the basic schematic towards an R-2R ladder in an iterative way. Using $1 + r + r^2 + \cdots = 1/(1-r)$, gives: $\sigma_{\text{tot}}^2 = (\sigma_a^2 + \sigma_b^2)/15$. With $\sigma_a = 0.02R$ and $\sigma_b = 0.01R$, $\sigma_{\text{tot}} = 0.01R/\sqrt{3}$.

The influence of the additional sections on the splitting process of the first stage is marginal due to the division by 16.

This allows to reduce the current splitting problem to the current splitting in the first stage, Fig. 7.13 (right) shows two equivalent input resistors, each nominal $2R$. Resistor R_{down} has a specified standard deviation of $\sigma_{R_{\text{down}}} = 0.02R$ and $R_{\text{right}} = (R_b + R_{\text{tot}})$ with $\sigma_{R_{\text{right}}} = R\sqrt{0.01^2 + 0.00577^2} = 0.0115R$. Analyzing the current splitting of $I_{\text{tot}} = I_{\text{right}} + I_{\text{down}}$:

$$\begin{aligned}
\sigma_{I_{\text{right}} - I_{\text{down}}}^2 &= \sigma_{I_{\text{right}}}^2 + \sigma_{I_{\text{down}}}^2 = \left(\frac{dI_{\text{right}}}{dR_{\text{right}}}\right)^2 \sigma_{R\text{right}}^2 + \left(\frac{dI_{\text{down}}}{dR_{\text{down}}}\right)^2 \sigma_{R_{\text{down}}}^2 \\
&= \left(\frac{d}{dR_{\text{right}}}\left(\frac{V_{\text{REF}}}{R_{\text{right}}}\right)\right)^2 \sigma_{R_{\text{right}}}^2 + \left(\frac{d}{dR_{\text{down}}}\left(\frac{V_{\text{REF}}}{R_{\text{down}}}\right)\right)^2 \sigma_{R_{\text{down}}}^2 \\
&= \left(-\frac{V_{\text{REF}}}{R_{\text{right}}^2}\right)^2 \sigma_{R_{\text{right}}}^2 + \left(-\frac{V_{\text{REF}}}{R_{\text{down}}^2}\right)^2 \sigma_{R_{\text{down}}}^2 \\
&= \frac{I_{\text{tot}}}{4R}\left(\sigma_{R_{\text{right}}}^2 + \sigma_{R_{\text{down}}}^2\right) = \frac{I_{\text{tot}}}{4}\sqrt{0.0115^2 + 0.02^2} = 0.0058 I_{\text{tot}}.
\end{aligned}$$

One σ equals 0.58% of the total current. Monotonicity means that the maximum error in $I_{\text{right}} - I_{\text{down}}$ is not exceeding an LSB: $2^{-N}(I_{\text{right}} + I_{\text{down}})$. For $N = 6$, an LSB equals 1.6%. So for a 6-bit R-2R DAC there is a 2.8σ probability or 0.5% chance that non-monotonicity will occur.

7.3 Digital-to-Analog Conversion in the Current Domain

Figure 7.14 shows the block diagram of unary- and binary-coded digital-to-analog converters in the current domain.

The buffer in Fig. 7.14 provides a low-impedance load for the current sources, avoiding modulation of the currents due to their finite output impedance. Major 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 a choice for a buffered output is made, a capacitor array or resistor ladder solution is preferred.

7.3.1 Current Steering Digital-to-Analog Converter

The block diagram in Fig. 7.15 (upper) shows a digital-to-analog converter based on current sources without a bandwidth-limiting buffer. The current sources are directly feeding the load impedance. The upper circuit shows a 3-bit unary array supplemented by a four-bit binary array. The currents sources are switched between the output rail and the power supply. A current that is not contributing to the output current, 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

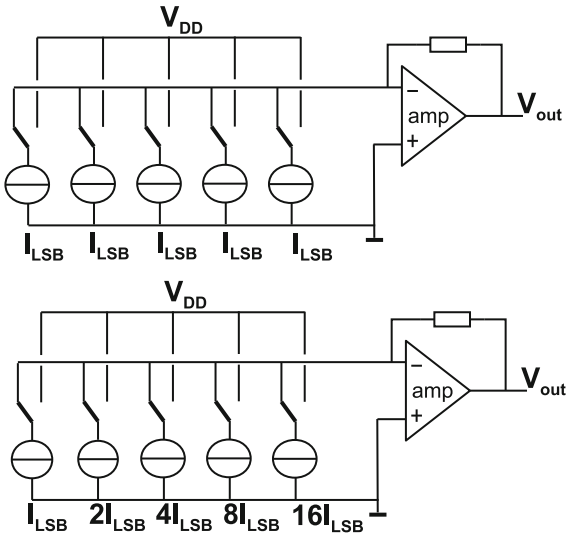


Fig. 7.14 Digital–analog converter with current sources: unary and binary forms

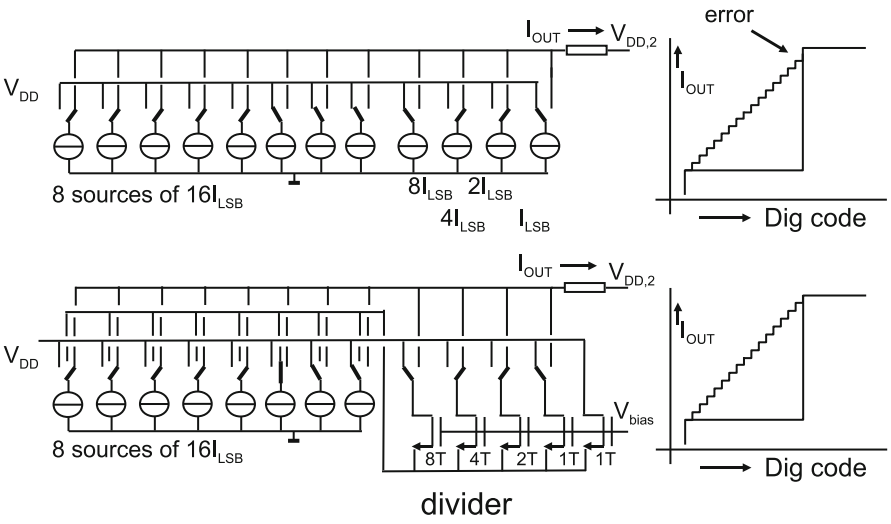


Fig. 7.15 *Upper*: A unary array with current sources is complemented with a binary array. *Lower*: the binary array is replaced by a current divider

lead to linear and non-linear distortion therefore unused currents must be drained in a power rail. Consequently these converters always consume the maximum current.

The speed of this arrangement is only limited by the pole of the output impedance. This topology is suited for delivering a high-performance time-continuous signals into 50 to 75 Ω loads. This architecture is mostly called "current-steering digital-to-analog converter."

These converters combine unary current sources for the MSB values, thereby ensuring good DNL for the major transitions, with binary-coded current sources for the lower bits. The total area is considerably smaller than for a completely unary implementation. A 10-bit converter can be built from 64 unary current sources and 4 binary current sources requiring considerably less area than 1,024 sources in a full unary implementation. The area reduction is of the order of $2^{N_{\text{binary}}}$ where N_{binary} represents the resolution of the binary section. The INL specification depends on the matching and gradient of the unary current sources.

The choice between unary and binary block sizes depends on technology and required DNL performance [135–137]. In resolution-limited 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 (see Eq. 2.110) the total amount of gate area for a certain resolution for both architectures is comparable. Secondary arguments such as wiring overhead will decide.

The design of the unary and binary currents sources requires some care. The global variations in the currents as caused by process, voltage, and temperature apply to all current sources and with careful design (see, e.g., Table 11.7) deviations can be minimized. However, uncorrelated effects and gradients can cause errors at

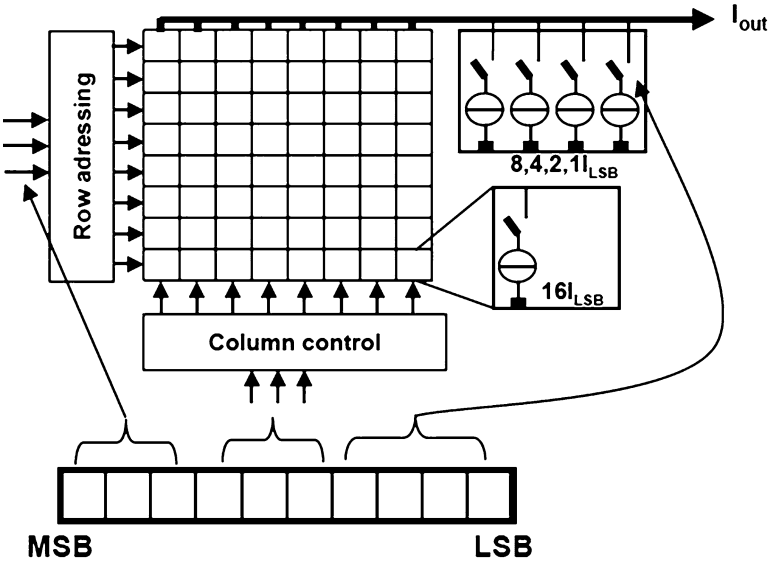


Fig. 7.16 Ten-bit digital–analog converter: the 6 MSBs are implemented as 64 unary current sources in a matrix configuration, while the four LSBs are designed in a binary series

all transitions, but are largest at the transitions between the maximum range of the binary array and the step of the unary array.

A second implementation of a segmented unary-binary architecture is shown in Fig. 7.15 (lower). In this circuit there is a third rail for the unary current sources. The required unary currents are switched to the output node as in the upper plot. The next current source is not connected to the power supply but feeds a binary divider array. Although there are still DNL errors caused inside the binary divider, there is only a minor transition error at the transitions between the binary and unary arrays. The disadvantage of this architecture is in the synchronization between the switching of the binary and unary current sources causing timing-related errors at high frequencies.

7.3.2 Matrix Decoding

A popular arrangement for the digital decoding of the input word into current source control signals is a column-row addressing scheme resembling a random-access memory, as indicated in Fig. 7.16. A straightforward selection, e.g., like someone reading this text, from left to right and top to bottom, will emphasize gradient effects. The current sources that form the unary part of the digital-to-analog converter are preferably selected in a way that cancels gradients [132, 138]. These gradients can occur due to technological deviations such as doping or oxide

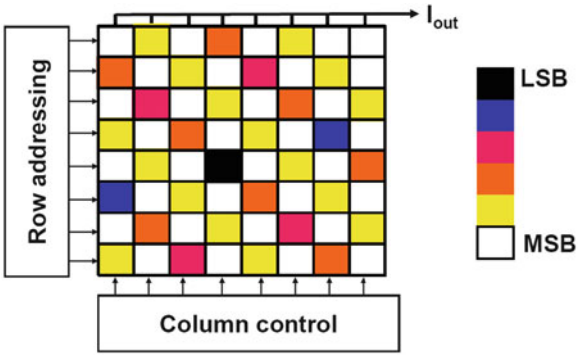


Fig. 7.17 Placement of unary-weighted current sources in a 64-element array. Similar colored squares represent current sources that form together one-bit level [138]

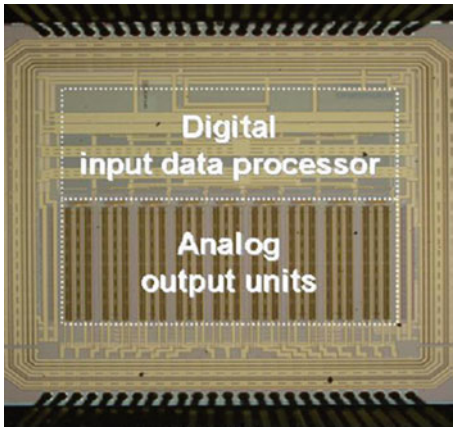


Fig. 7.18 Die photograph of a 16-bit current-steering digital-to-analog converter in 180-nm CMOS (courtesy: J. Briaire)

thickness variations, power supply drop due to voltage drop over current-carrying lines, clock timing gradients, and temperature gradients. In practical design the technological gradients are rather limited in magnitude. Especially the voltage, time [139], and temperature gradients can become rather severe.

Figure 7.17 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 [140]. 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 [141]. Figure 7.18 shows a die photograph of a current-steering digital-to-analog converter based on current source sorting. An algorithm first measures the values of the current sources and then arranges them in groups that optimize the overall performance. For example, the $2^{N_{\text{unary}}}$ unary currents are selected by combining $2^{N_{\text{unary}}+1}$ current

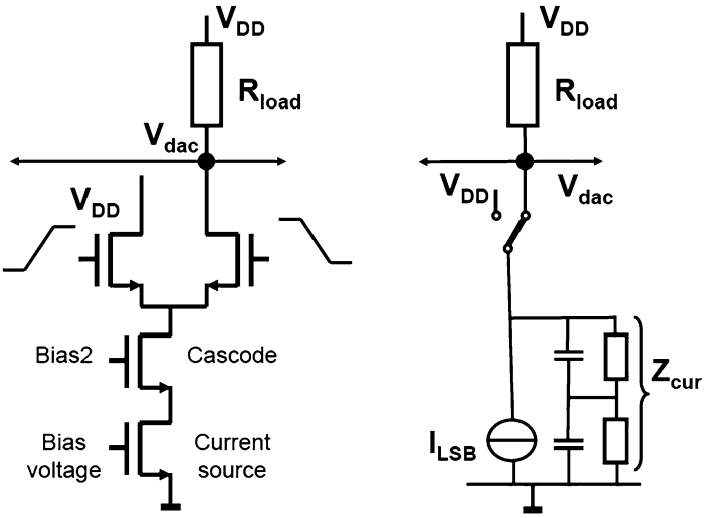


Fig. 7.19 The basic digital-analog converter current cell and its small-signal equivalent circuit

sources in a way that one high-current source is matched by one low-current source. A considerable reduction of inequality can be achieved at the cost of some preprocessing.

More points to consider with matrices of current sources are the following:

- 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 causing mismatch (Sect. 11.3).
- The decoding network consumes power and increases cross talk.
- Close to the final switches the selection signals must be retimed by the main clock signal. Delays between the switching-on and off of various current sources create unequal $\int i\mathrm{d}t$ contributions to the output signal. Proper distribution of the clock signal with equal traveling distances for each cell (10 µm wire causes 0.1 ps time difference) is needed.

7.3.3 Current Cell

Figure 7.19 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).

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

$$V_{\mathrm{DD}} - V_{\mathrm{dac}} = \alpha 2^N I_{\mathrm{LSB}} R_{\mathrm{load}}. \tag{7.12}$$

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. 7.19, 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

$$I_{\mathrm{err}} = \alpha 2^N V_{\mathrm{dac}} / Z_{\mathrm{cur}}. \tag{7.13}$$

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:

$$V_{\mathrm{DD}} - V_{\mathrm{dac}} = R_{\mathrm{load}} \left(\alpha 2^N I_{\mathrm{LSB}} + \frac{\alpha 2^N V_{\mathrm{dac}}}{Z_{\mathrm{cur}}} \right). \tag{7.14}$$

Substitution of a full-swing output signal $\alpha = 0.5 + 0.5 \sin(\omega t)$ results after some manipulation in

$$\mathrm{HD2} = \frac{\text{second-order component}}{\text{first-order component}} = \frac{2^N R_{\mathrm{load}}}{4 Z_{\mathrm{cur}}}. \tag{7.15}$$

This signal ratio can be expressed in dBs via $20^{10}\log(\mathrm{HD2})$. The second-order distortion is directly related to the ratio of the output impedance of the current sources to the load impedance.

A differential output scheme, where the complementary current also drives a similar load, largely compensates this error.

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.

The switch transistors are optimized for fast switching and are controlled by differential pulses. Although the switches are in series with a current source, the voltage drop over the switch can affect the performance, so their on-resistance is made low. As all resistances of the current switches are in parallel, the value of $R_{\mathrm{switch}}/2^N$ must be very low compared to R_{load}.

Yet, 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

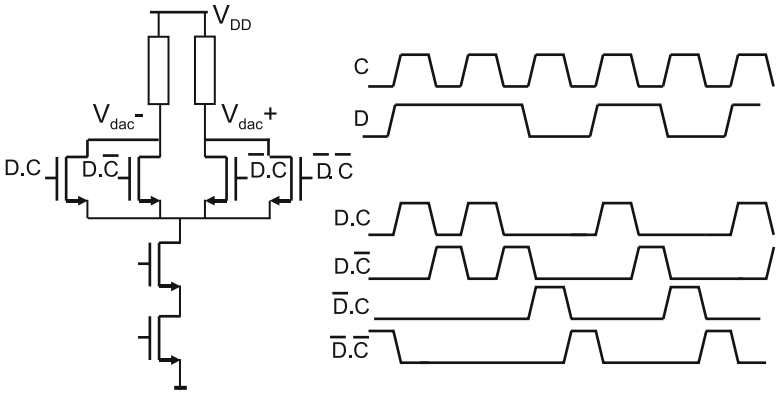


Fig. 7.20 A current cell with quadruple switches keeps the number of transitions equal for every data sequence [142]

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 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. In Fig. 7.20 a switching method is shown that will generate an equal number of transitions for any sequence of data bits [142] extended in [143].

Additional problems with the switches, clock, and signal lines are the timing-related errors.

- Clock jitter over the entire converter has a similar effect in the reconstruction of the signal as jitter in sampling. A signal frequency-dependent noise contribution will appear in the output spectrum (compare Sect. 7.5).
- Wire length differences in clock lines (skew) or in current summing lines can create fixed timing errors (10 µm equals 0.1 ps). As these errors can be correlated for parts of the current sources, both a noise-like behavior as well as distortion can result.
- Mismatch in transistor parameters can create random variation in switching moments; see Fig. 11.32.

Example 7.4. Show that a differential configuration of a current-steering digital-to-analog converter suppresses the second-order distortion.

Solution. If the constant V_{DD} term in Eq. 7.14 is ignored, the signal terms can be evaluated as:

$$V_{dac} = -R_{load}\alpha 2^N I_{LSB}\frac{1}{1-\dfrac{\alpha 2^N R_{load}}{Z_{cur}}} \approx -R_{load}\alpha 2^N I_{LSB}\left(1+\frac{\alpha 2^N R_{load}}{Z_{cur}}\right)$$
$$= C_1\alpha + C_2\alpha^2, \tag{7.16}$$

where the approximation $1/(1-a) \approx (1+a) \quad |a| \ll 1$ is used. If the complementary current $(1-\alpha)2^N I_{LSB}$ is fed in a second resistor of value R_{load} the resulting voltage $V_{dac,inv}$ is easily found by replacing α in the equations by $(1-\alpha)$. The differential voltage is then:

$$V_{dac} - V_{dac,inv} = C_1\alpha + C_2\alpha^2 - (C_1(1-\alpha) + C_2(1-\alpha)^2) = (C_1+C_2)(2\alpha - 1) \tag{7.17}$$

which is linear with the signal term α.

7.3.4 Performance Limits

Current-steering converters without speed-limiting buffer reach high operating frequencies. The speed constraint at the output of these converters is limited by the 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 performance. The penalty for these advantages is power. 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 as there is no signal-dependent current flowing through.

Figure 7.21 compares data from various publications [136, 139, 141, 144, 145] 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. Another metric of comparison is a figure of merit see Eq. 12.6.

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 power penalty is accepted.

Example 7.5. A unary current matrix with 1,024 current sources produces a maximum current of 20 mA in a 50 Ω load. The current sources are built in 0.18 µm CMOS with an output impedance of 100 kΩ. A single-transistor cascode stage is

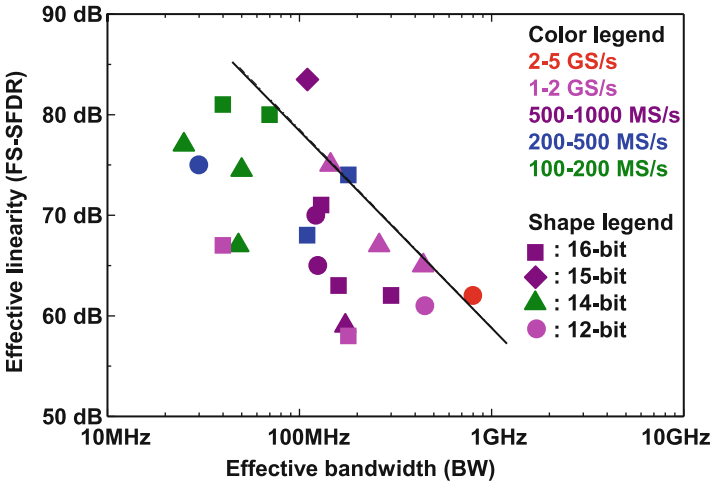


Fig. 7.21 Spurious-free dynamic range versus bandwidth for current-steering digital-to-analog converters published from 2000 to 2009 (courtesy: J. Briaire)

used. What is the total output impedance of one current source if the gate cascade transistor measures 1µm/0.18µm? 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?

Solution. A minimum gate-length transistor has according to Table 2.21 a static feedback factor of $\lambda = 0.05$. So this cascode stage increases the output impedance of the current source to $Z_{\text{cur}} = r_0(1+1/\lambda) = 100\,\text{k}\Omega(1+20) = 2.1\,\text{M}\Omega$; see Eq. 2.196. The dominant distortion component is the second-order term:

$$\text{HD2} = \frac{2^N R_{\text{load}}}{4Z_{\text{cur}}} = 6.1 \times 10^{-3}$$

or 44.3 dB.

In case of $N = 10$, $R_{\text{load}} = 50\,\Omega$, 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 which can be achieved by choosing longer transistors or adding another cascode stage.

If the distortion increases by 3 dB, the impedance $|Z_{\text{cur}}|$ must be 3 dB lower. As Z_{cur} is formed by the resistive output impedance of the cascode stage and the loading capacitor of 100 fF, $|Z_{\text{cur}}|$ will drop 3 dB at the frequency where the capacitive impedance equals the resistive impedance: $r_0(1+1/\lambda) = 1/(2\pi fC)$. This results in $f_{3\,\text{db}} = 0.76\,\text{MHz}$.

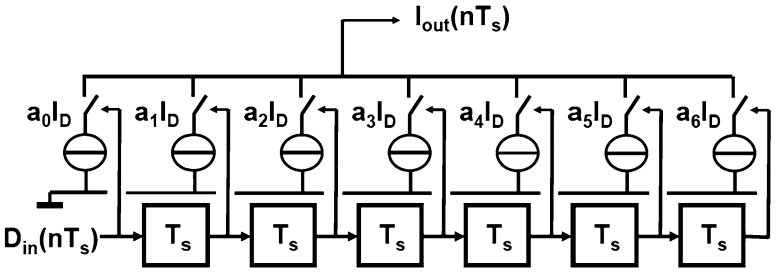


Fig. 7.22 A semi-digital filter and digital-analog converter [146]

7.3.5 Semi-digital Filter/Converters

After the conversion from the digital to the analog domain with a digital-to-analog converter, the higher alias bands contain a lot of energy, despite some filtering by the zero-order hold function. In many applications the high-frequency energy must be removed before the signal can be used.[4] Su and Wooley [146] have proposed a digital-to-analog structure that converts the digital signal back into its analog form and at the same time filters the result. The digital signal is first converted in a high-frequency low-resolution data stream see Sects. 9.2 and 9.3. The semi-digital filter in Fig. 7.22 shifts the signal through a digital shift register where every output of the shift-register sections controls a weighted current source. The outputs of these current sources are summed.

The most simple implementation uses a one-bit digital signal that will switch on or off the weighted current sources. With the help of Sect. 3.2.1 the result is written as:

$$I_{\text{out}}(nT_s) = \sum_{k=0}^{k=K-1} a_k D_{\text{in}}((n-k)T_s) I_D. \tag{7.18}$$

The coefficients are chosen using similar constraints as for a normal FIR filter. This structure is very useful in the reconstruction of delta-modulated signals [147].

The advantage of this structure is the suppression of the alias components in the current domain. Voltages will only appear after the current summation. So higher-order signal components can be effectively suppressed in the current domain and do not generate distortion when they appear as voltages over nonlinear components.

It is important to realize that errors in the coefficients in first order will affect the filter characteristics and not the overall linearity. An error in the current sources that implement a unary digital-to-analog converter will cause a linearity error at that level with harmonic distortion. In a semi-digital converter this error modifies

[4] Think of all the energy your tweeter loudspeakers would have to consume.

the weighting coefficient by Δa_i and will result in an additional term of $1 + \Delta a_i z^{-i}$ which is a linear time-shifted signal contribution. The filtering will be less perfect, but no harmonic distortion will be generated.

7.4 Digital-to-Analog 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 [148]; see Fig. 7.23. 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 [149, Fig. 9]. A unary-decoding scheme with equal capacitors and switches is used to implement the conversion.

The configuration in Fig. 7.23 does not suffer from parasitics that are attached to the top plates of the capacitors as the virtual ground of the opamp stabilizes this node.

A stray-insensitive switching topology uses a standard switched-capacitor technique to move charges. The parasitic capacitances connected to the switched capacitors in Fig. 7.24 are either charged from voltage sources or do not see any charge change because they are connected to the (virtual) ground. In Fig. 7.24 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,

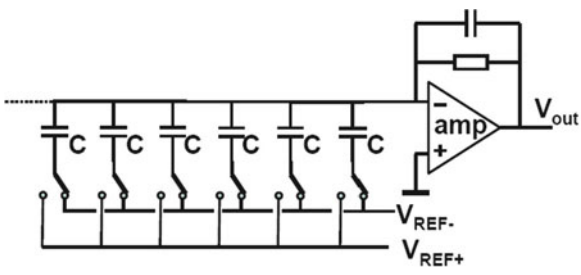


Fig. 7.23 Digital-analog converter based on unary capacitors

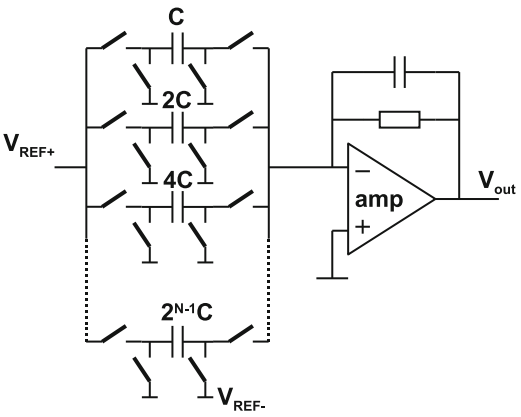


Fig. 7.24 Binary-weighted digital-to-analog converter in a stray-insensitive switched capacitor configuration

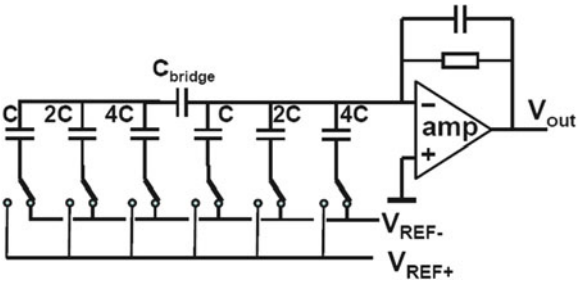


Fig. 7.25 Digital–analog converter with bridging capacitor

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. 7.25. 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. The capacitance of an LSB cell equals C, the LSB+1 cell equals $2C$, and the largest cell uses $2^{k-1}C$. The total capacitance on the LSB side is $(2^k - 1)C$ If an amount of iC on the LSB side is switched from negative reference to positive reference while the remaining capacitance $(2^k - 1 - i)C$ stays connected to the negative reference, the voltage change on the left-hand side of C_{bridge} is

$$\frac{iC(V_{\text{ref}+} - V_{\text{ref}-})}{C(2^k - 1) + C_{\text{bridge}}}. \tag{7.19}$$

This charge injection in the integration capacitor is compared to the desired charge injection in the summation node of the opamp:

$$\frac{C_{\text{bridge}}}{C_{\text{bridge}} + (2^k - 1)C} iC(V_{\text{ref}+} - V_{\text{ref}-}) = \frac{i}{2^k} C(V_{\text{ref}+} - V_{\text{ref}-}). \tag{7.20}$$

From this equation $C_{\text{bridge}} = C$ is found.

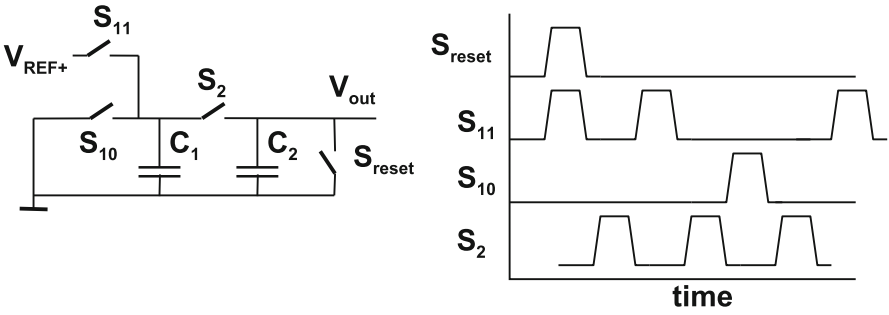


Fig. 7.26 A charge redistribution converter [148]

Note that parasitic capacitances connected to the left side of the bridge capacitor will affect the performance.

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 at 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; see also Sect. 7.5 and Fig. 9.32.

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. 7.26. 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 decides whether C_1 is charged to the positive (via S_{11}) or negative (via S_{10}) reference. Then switch S_2 connects the capacitors in parallel, redistributing the charge from the reference over both capacitors, thereby halving the value. 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.

$$V_{\text{out}}(i) = b_i \frac{C_1}{C_1 + C_2} V_{\text{ref}} + \frac{C_2}{C_1 + C_2} V_{\text{out}}(i-1), \tag{7.21}$$

where $b_i \in \{0, 1\}$, with $i = 0 \ldots (N-1)$. 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:

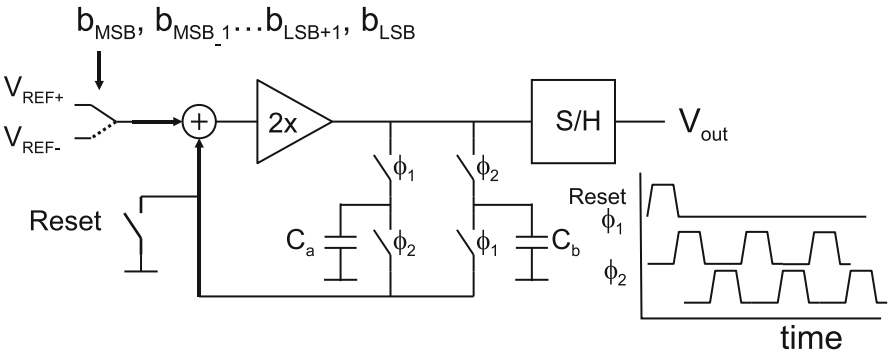


Fig. 7.27 An alternative implementation of an algorithmic digital-to-analog converter

$$V_{\text{out}}(N) = \sum_{i=0}^{i=N-1} \frac{b_i 2^i}{2^N} V_{\text{ref}}, \tag{7.22}$$

where b_0 is the LSB and b_{N-1} is the MSB. 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:

$$\delta V_{\text{out}}(N) = \sum_{i=0}^{i=N-1} \frac{b_i 2^i \delta^{N-1-i}}{2^N} V_{\text{ref}} < \delta V_{\text{ref}}. \tag{7.23}$$

Assuming that the error at the MSB transition must remain under the value of an LSB, sets a limit for the capacitor deviation:

$$\delta = \frac{C_2 - C_1}{C_1 + C_2} < 2^{-N}. \tag{7.24}$$

An error of 0.1% in the capacitors limits the achievable resolution to $N = -^2\log(0.001) \approx 10$ bits.

These switched-capacitor digital-to-analog converters are mostly applied in larger system chips, where they perform low-power conversions.

Example 7.6. Figure 7.27 shows a variant on the design of Fig. 7.26. The incoming bits are added to the stored result which is then amplified by 2 and stored for the next cycle in the capacitors. Now the operation starts with the MSB that is fed into the algorithmic digital-to-analog converter, processed and multiplied by 2 and followed by the MSB-1. Discuss the merits of this modification.

Solution. The result of this operation bears great resemblance to Eq. 7.22: $V_{\text{out}}(N) = \sum_{i=N}^{i=1} b_{i-1} 2^i V_{\text{ref}}$, where b_0 is the LSB and b_{N-1} is the MSB. The denominator term 2^N is missing. In order to keep the overall result of the addition within operation margins, V_{ref} must now equal the small value of V_{LSB}. Any error on this quality (noise, etc.) will be amplified by the loop and result in poor signal performance. In a preferred implementation the multiplication is with a factor 0.5

and the sequence starts with the LSB. Due to the successive divisions, any error that occurred in the first cycles with respect to the large V_{ref} is further reduced.

Example 7.7. Charge-domain digital-to-analog converters can suffer from capacitor mismatch and from KT/C noise. At what signal level is the contribution of both effects in terms of energy equal for a 1 pF capacitor?

Solution. A 1 pF capacitor will show a mismatch with respect to another capacitor. According to Eq. 2.91: $\sigma_{\Delta C}/C = A_C/\sqrt{C}$. With the value in Table 2.19 and $C = 1,000$ fF, the relative mismatch is found as: 1.6×10^{-4} or 0.16 fF. This same signal-to-spurious ratio can be expected for a processed signal A_{rms}.

In Fig. 3.10 the kT/C noise equals 65 μV_{rms}. Both contributions are equal if $1.6 \times 10^{-4} \times A_{\text{rms}} = 65 \times 10^{-6} V_{\text{rms}}$, or $A_{\text{rms}} = 0.4 V_{\text{rms}}$. This value corresponds to e.g., a sine wave of 0.56 V amplitude or 1.12 $V_{\text{peak–peak}}$. It is clear that both sources of unwanted energy play a different role in a design. Note that, while kT/C noise acts as a noise floor from 0 to $f_s/2$, the effects of capacitor mismatch can take many forms: from spurious single-tone component to shaped noise after data-weighted averaging.

7.5 Digital-to-Analog 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. 7.28. One of the first techniques to digitize information was based on the succession of pulses [105] and was originally called pulse code modulation (PCM). This specific implementation is today referred to as pulse density modulation (PDM).

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

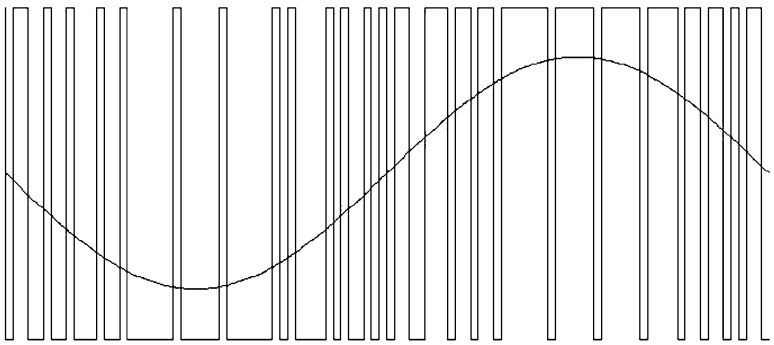


Fig. 7.28 The low-frequency content of a pulse sequence contains the signal

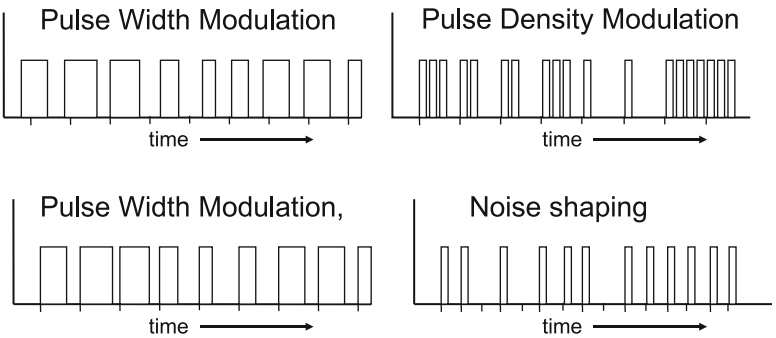


Fig. 7.29 Several forms of time-domain digital-to-analog conversion

conversion quality. The linearity problems due to component inaccuracies are circumvented. Pulse-width modulation (PWM) or PDM 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 7.29 shows four differently coded sequences. On the left side are the PWM 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 pulse. Sigma–delta modulation and noise shaping are forms of PDM.

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:

$$\begin{aligned} \text{Energy in PWM signal} \quad &\propto (+/-A)^2 \quad = A^2 \\ \text{Energy in maximum sine wave} &\propto \int (A\sin(\omega t))^2 \mathrm{d}t = A^2/2. \end{aligned}$$

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 nonlinear elements, down-mixing of harmonics of the fundamental signal can occur.

[5]Compare this result to the power in the harmonics of a block wave in Eq. 2.6.

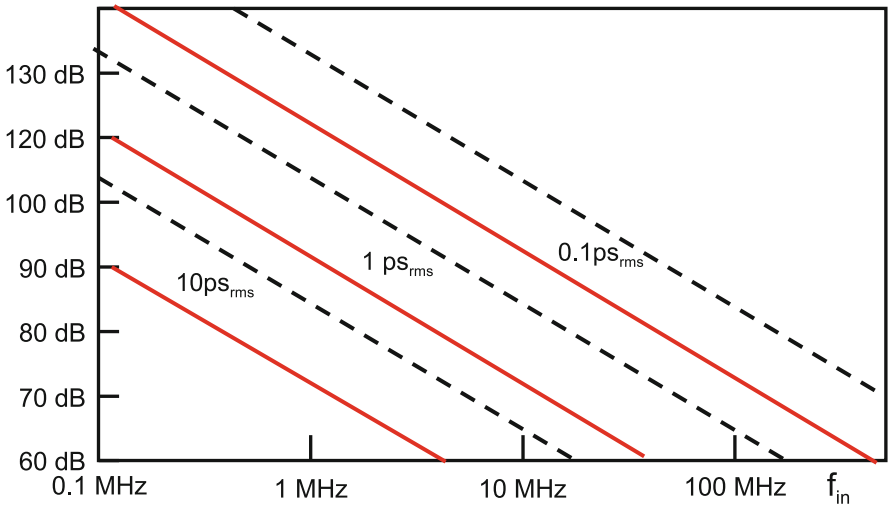


Fig. 7.30 Comparison of jitter in normal sampling (*dotted line*) and in one-bit signals where a moderate oversampling ratio of 5 is used

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. 3.1.7). 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

$$\text{SNR} = \frac{\frac{1}{T_a}\int_{t=0}^{T_a}(A\sin(\omega t))^2\mathrm{d}t}{\frac{1}{T_a}\int_{t=0}^{T_a}(2A\sigma_{\text{jit}}\alpha f_s)^2\mathrm{d}t} = \frac{1}{8(\alpha f_s\sigma_{\text{jit}})^2} = \left(\frac{1}{4\sqrt{2}\text{OSR}\alpha f_{\text{sig}}\sigma_{\text{jit}}}\right)^2 \tag{7.25}$$

with $\text{OSR} = f_s/2f_{\text{sig}} \gg 1$ as the oversampling ratio. Comparison with the signal-to-noise ratio of sampled signals in Eq. 3.26

$$\text{SNR} = \left(\frac{1}{2\pi f_{\text{sig}}\sigma_{\text{jit}}}\right)^2 \tag{7.26}$$

[6] Assuming only one edge is jittering.

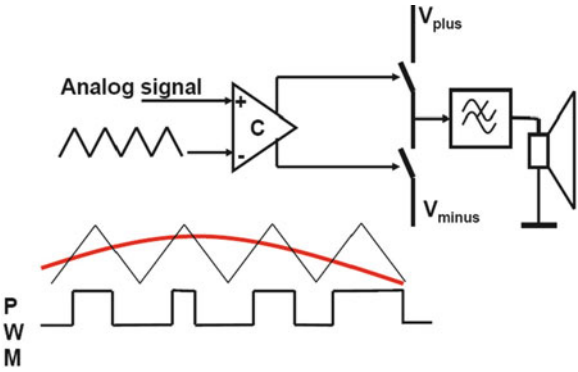


Fig. 7.31 Class-D amplifier with PWM

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. 7.30.

7.5.1 Class-D Amplifiers

A well-known application of pulse-width modulation is in audio class-D amplifiers as shown in Fig. 7.31. 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., [150].

Example 7.8. What efficiency can be expected from a PWM class-D output stage for 4 Ω load impedance with 0.2 Ω resistance per switch, 20 nF switch gate capacitance, and 0.5 MHz clock frequency?

Solution. The major loss mechanisms in class-D are IR-drops, R_{on} of the switch, and switching losses of both switches: $CV_{DD}^2 f$. For the total expected efficiency this means

$$\eta = \frac{P_{load}}{P_{load} + P_{res} + 2P_{cap}} = \frac{I^2 R_{load}}{I^2 R_{load} + I^2 R_{on} + 2CV_{DD}^2 f}$$

substituting $V_{DD} = I(R_{load} + R_{on})$ allows to eliminate the current, and an efficiency of 88% is found.

7.6 Accuracy

7.6.1 Accuracy in Resistors Strings

The accuracy in which a quantity can be subdivided limits the obtainable performance of a digital-to-analog converter. In Fig. 7.32 a ladder of $M = 256$ equally designed resistors is connected between a reference voltage and ground. Ideally the voltage at the mth position is:

$$V(m) = \frac{m}{M} V_{\text{ref}} = \frac{mR}{mR + (M-m)R} V_{\text{ref}} = \frac{R_1}{R_1 + R_2} V_{\text{ref}}. \tag{7.27}$$

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:

$$E(R_1(m)) = E(mR) = mE(R) = mR, \quad \sigma_{R_1}^2 = m\sigma_R^2. \tag{7.28}$$

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

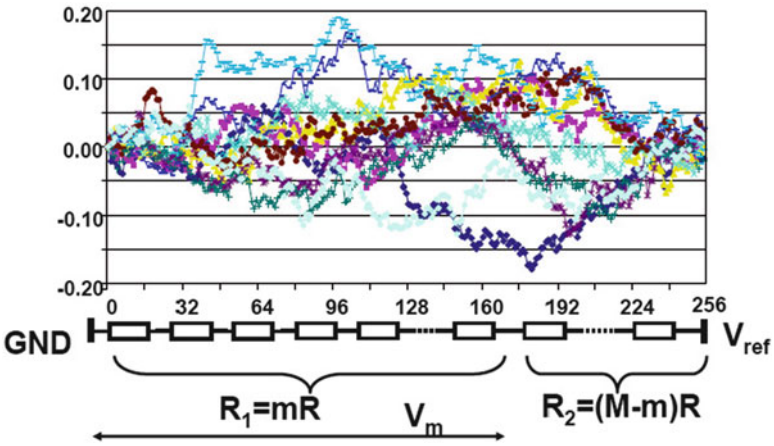


Fig. 7.32 A resistor string is connected between a reference voltage and ground. The simulation shows 10 instances of the effect of $\sigma_R/R = 1\%$ mismatch. In this 256-device ladder the maximum deviation is $\pm 0.2R$

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 (Eq. 2.18):

$$\begin{aligned}\sigma_V^2(m) &= \left(\frac{\partial V(m)}{\partial R_1}\right)^2 \sigma_{R_1}^2 + \left(\frac{\partial V(m)}{\partial R_2}\right)^2 \sigma_{R_2}^2 \\ &= \left(\frac{R_2}{(R_1+R_2)^2}\right)^2 \sigma_{R_1}^2 V_{\text{ref}}^2 + \left(\frac{-R_1}{(R_1+R_2)^2}\right)^2 \sigma_{R_2}^2 V_{\text{ref}}^2 \\ &= \frac{m(M-m)}{M^3} \frac{\sigma_R^2}{R^2} V_{\text{ref}}^2.\end{aligned}$$

Compare to Eq. 7.5. The maximum value of the variance occurs at $m = M/2$ for which the variance is found as

$$\sigma_V^2(m = M/2) = \frac{1}{4M} \frac{\sigma_R^2}{R^2} V_{\text{ref}}^2 = \frac{M}{4} \frac{\sigma_R^2}{R^2} V_{\text{LSB}}^2, \tag{7.29}$$

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 Eq. 7.29 can serve to estimate the INL. With the help of Table 2.11 the sigma margin can be chosen.

In the example of Fig. 7.32 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 determined by the single resistor variance and the number of resistors. The DNL is determined by the largest deviating resistors amongst M random-varying resistors with each a relative spread of 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.

Example 7.9. If the resistors in a string show a relative spread of 1%, what is the probability that one resistor in a string of 256 exceeds a deviation of 4%?

Solution. The probability p that one resistor with value R_a and $\sigma_R/R = 1\%$ deviates more than 4% from the average value R is:

$$p = P\left(\left|\frac{R_a - R}{R}\right| > \frac{4\sigma_R}{R}\right) = P(|x| > 4\sigma) \tag{7.30}$$

[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.

with Table 2.11 $p = 6.3 \times 10^{-5}$. The yield per resistor is $(1-p)$. The yield for a resistor string with $M = 256$ resistors is $(1-p)^M = 98.4\%$.

7.6.2 Accuracy in Current Source Arrays

In a similar manner the effects of random variation in a unary current source array or a capacitor array can be calculated. The variance in currents 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:

$$I(m) = \frac{mI}{mI + (M-m)I} I_{\text{tot}} \tag{7.31}$$

the mathematical description is identical to the resistor string problem and yields:

$$\sigma_{MI/2}^2 = \frac{1}{4M} \frac{\sigma_I^2}{I^2} I_{\text{tot}}^2. \tag{7.32}$$

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 7.33 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 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. Data-weighted averaging techniques circumvent the occurrence of harmonic distortion due to random errors; see Sect. 7.7.3.

Digital-to-analog converters based on a binary architecture are analyzed after some modifications 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:

$$I_{N-1} = 2^{N-1} I_{\text{LSB}}, \quad I_{N-2} = 2^{N-2} I_{\text{LSB}}, \quad \ldots, \quad I_1 = 2^1 I_{\text{LSB}}, \quad I_0 = I_{\text{LSB}}. \tag{7.33}$$

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}:

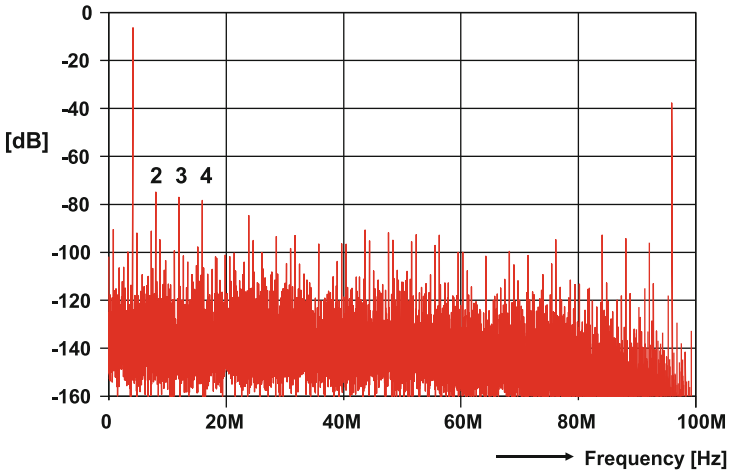


Fig. 7.33 Spectrum of a digital-to-analog converter with random mismatch. An array of 256 current sources is simulated with a normally distributed mismatch of $\sigma_I/I = 1\%$. The signal frequency is 3.97 MHz at 100 Ms/s sample rate; the quantization error is removed. The second-, third-, and fourth- order distortion components are labeled

$$\Delta I = I_{N-1} - I_{N-2} - \cdots - I_1 - I_0 = 2^{N-1} I_{\mathrm{LSB}} - 2^{N-2} I_{\mathrm{LSB}} \cdots - 2^1 I_{\mathrm{LSB}} - 2^0 I_{\mathrm{LSB}} = I_{\mathrm{LSB}}. \tag{7.34}$$

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. 7.14 (lower). The current branch of weight k in the array is described by $I_k = 2^k I_{\mathrm{LSB}}$. Let each basic current source I_{LSB} suffer from a variation mechanism characterized by a normal distribution with mean value $I_{\mathrm{LSB}m}$ and a variance σ_I^2. Then each branch will be characterized by a mean $2^k I_{\mathrm{LSB}m}$ and a variance $\sigma_{I_k}^2 = 2^k \sigma_I^2$ see Eq. 2.19. With the same equation the variance for the current step on the MSB transition is:

$$\sigma_{\Delta I}^2 = (2^{N-1})\sigma_I^2 + (2^{N-2} + \ldots + 2^1 + 2^0)\sigma_I^2 = (2^N - 1)\sigma_I^2. \tag{7.35}$$

  Monotonicity requires that the value of ΔI remains positive. If a 3σ probability (99.7%) is considered an acceptable yield loss, then $3 \times \sqrt{(2^N - 1)}\sigma_I < I_{\mathrm{LSB}}$. For an 8-bit converter this requirement results in $\sigma_I < 0.02 I_{\mathrm{LSB}}$.
- In an R-2R ladder as in Fig. 7.12, 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 $\sigma^2_{R_{1,2}}$. With the help of Eq. 2.18

$$
\begin{aligned}
I_{\mathrm{MSB}} &= \frac{R_2}{R_1+R_2}2^N I_{\mathrm{LSB}} \\
\frac{\sigma^2_{I\mathrm{MSB}}}{(2^N I_{\mathrm{LSB}})^2} &= \left(\frac{\partial I_{\mathrm{MSB}}}{\partial R_1}\right)^2 \sigma^2_{R_1} + \left(\frac{\partial I_{\mathrm{MSB}}}{\partial R_2}\right)^2 \sigma^2_{R_2} = \frac{R_1^2\sigma^2_{R_2}+R_2^2\sigma^2_{R_1}}{(R_1+R_2)^4} \\
&\approx \frac{2\sigma^2_{R_1}}{4(R_1)^2}.
\end{aligned}
\tag{7.36}
$$

Requiring for monotonicity that $3\times\sigma_{I_{\mathrm{MSB}}} < I_{\mathrm{LSB}}$, results for 8-bit resolution in $\sigma_{R_1}/R_1 = 0.002$. 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.

Digital-to-analog converter designs can be optimized based on the above analysis [137].

Example 7.10. Groups of 1, 2, 4, ... unit current sources are combined to form a binary architecture digital-to-analog converter. Each unit current source shows a random mismatch of $1\sigma = 1\%$ of the LSB current. How many bits can be designed when a DNL of 0.5 LSB must be reached for 99.7% (or -3 to $+3\sigma$) of the produced devices.

Solution. The absolute mismatch of a single LSB current source is $\sigma_I = 0.01 I_{\mathrm{LSB}}$. Putting n current sources parallel increases the average current to nI_{LSB} and the mismatch to $\sigma_{nI} = \sqrt{n}\sigma_I = \sqrt{n}0.01 I_{\mathrm{LSB}}$. The worst-case transition in an N-bit binary digital-to-analog converter is at the switching point between the MSB current $2^{N-1}I_{\mathrm{LSB}}$ with $\sigma_{I_{\mathrm{MSB}}} = \sqrt{2^{N-1}}\sigma_I$ and the LSB to MSB-1 currents: $(2^{N-1}-1)I_{\mathrm{LSB}}$ with $\sigma_{I\mathrm{lower-bits}} = \sqrt{2^{N-1}-1}\sigma_I$. The nominal difference between the groups of current sources is exactly $1I_{\mathrm{LSB}}$. The variance of the difference between these groups is

$$\sigma^2_{\mathrm{diff}} = (2^{N-1})\sigma_I^2 + (2^{N-1}-1)\sigma_I^2 = (2^N-1)\sigma_I^2$$

When $3\sigma_{\mathrm{diff}} < 0.5 I_{\mathrm{LSB}}$ the result is $N = 8$.

Example 7.11. A binary current digital-to-analog converter is build with weighted current sources, each with a random mismatch of 1% of the current, independent of the current value. How many bits can be designed as a binary section when a DNL of 0.5 LSB must be reached for 99.7% (or -3 to $+3\sigma$) of the produced devices?

Solution. The absolute mismatch of the ith current source ($i = 0,\ldots,N-1$) is now $0.01 I_{\mathrm{LSB}} \times 2^i$. The worst-case transition in an N-bit binary digital-to-analog converter is at the switching point between the MSB current $2^{N-1}I_{\mathrm{LSB}}$ with $\sigma_{I_{\mathrm{MSB}}} = 0.01\,2^{N-1} I_{\mathrm{LSB}}$ and the LSB to MSB-1 currents: $\sigma_{I\mathrm{lower-bits}} =$

$0.01 I_{\mathrm{LSB}} \sqrt{2^{2(N-2)} + \cdots + 2^8 + 2^4 + 2^2 + 2^0}$. The nominal difference between the groups of current sources is $1 I_{\mathrm{LSB}}$. The variance of the difference between these groups is

$$\sigma_{\mathrm{diff}}^2 = 2^{2(N-1)} (0.01 I_{\mathrm{LSB}})^2 + (2^{2(N-2)} + \cdots + 2^8 + 2^4 + 2^2 + 2^0)(0.01 I_{\mathrm{LSB}})^2$$

When $3\sigma_{\mathrm{diff}} < 0.5 I_{\mathrm{LSB}}$ the result is $N = 3$. The 1% relative spread in the MSB current clearly dominates. If a process shows large sensitivity to relative spreads (e.g., random mask variations in the gate length) it is better to design the converter with small current sources.

7.7 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. 7.34. 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 11.7 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.

In Fig. 7.35 the errors $\varepsilon_{1,2,3,i}$ affect the different paths through the analog-to-digital or digital-to-analog converter. Three classes of mitigation of additive errors

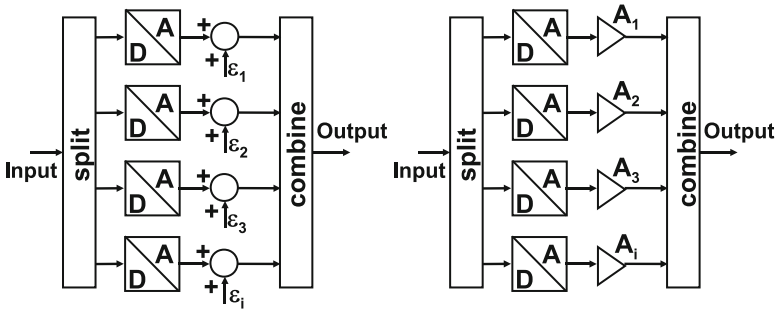


Fig. 7.34 Additive and multiplicative errors in conversion

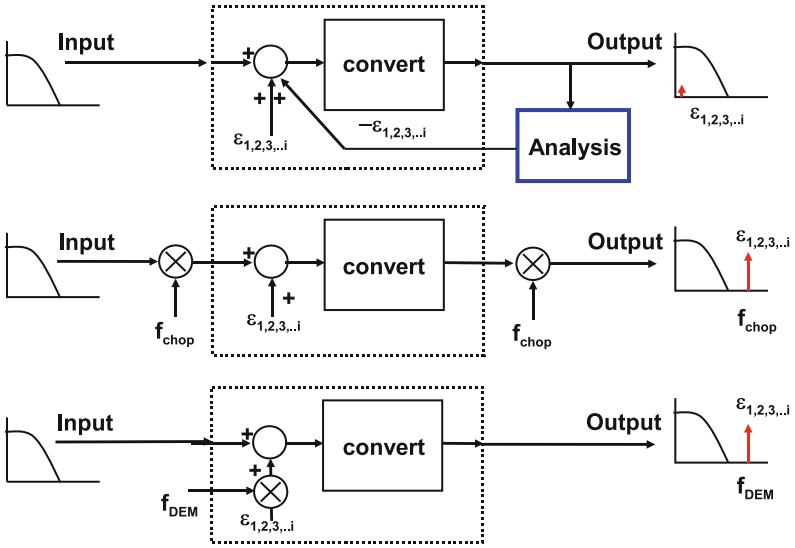


Fig. 7.35 Three methods to mitigate additive errors. *Top*: calibration, *middle*: chopping, *bottom*: dynamic element matching

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. For example, 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 signal path. An example of the calibration of a matrix of current sources is given in Sect. 7.7.1.

The second method to mitigate additive errors was 50 years ago already in use for precision instrumentation. The chopping technique [151], Fig. 7.35 (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 converter [152], 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; see also Sect. 3.1.3.

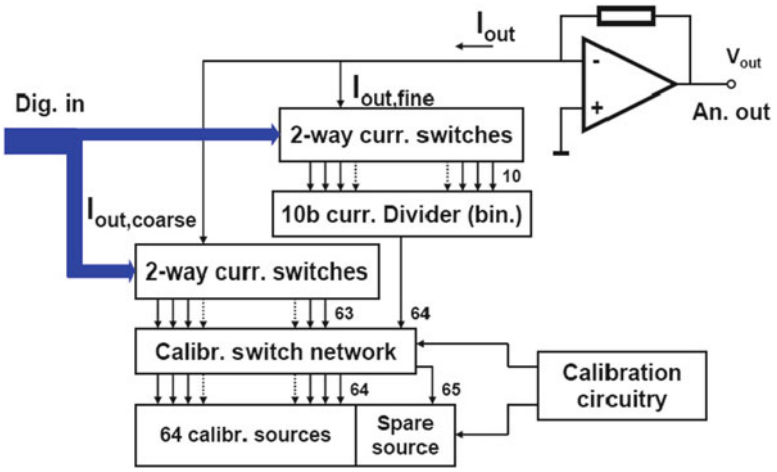


Fig. 7.36 An example of a digital-to-analog converter applying current calibration [133]

The last principle 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. 7.7.2 and 7.7.3.

The discussion in this section is limited to mitigation of additive errors. The impact and avoidance of multiplicative and timing errors are discussed in Sect. 8.8.

7.7.1 Current Calibration

With the change from bipolar devices to CMOS techniques new forms of calibration became possible. Figure 7.36 shows a unary current array with one spare current source. The unary row of current sources is extended by one, from 64 to 65. This allows to calibrate every clock cycle one of the current sources of the unary array. In Fig. 7.37 two NMOS transistor form the calibrated current source: M_1 supplies the majority of current as a traditional current source. M_{1a} is used to calibrate and to mitigate the inherent mismatch in these sources. During the calibration the spare current source is connected as a sort of track-and-hold circuit by closing the switch between gate and drain. Feeding this arrangement with a reference current will force the gate voltage of M_{1a} to settle at the level needed to balance this current. After opening the switches the gate capacitor will hold a voltage that makes that the total current from M_1 and M_{1a} is a copy of the reference current. This current source is ready for use. Another 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

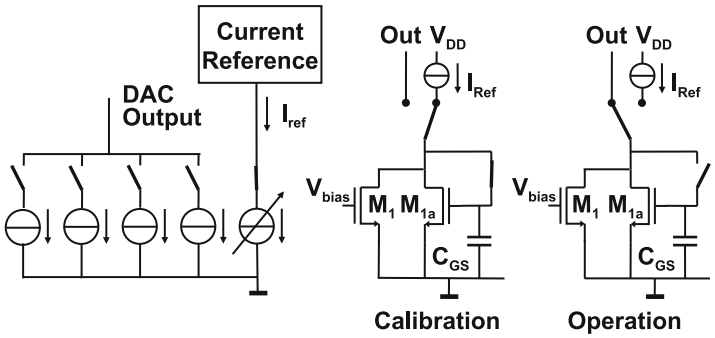


Fig. 7.37 The current calibration technique [133]

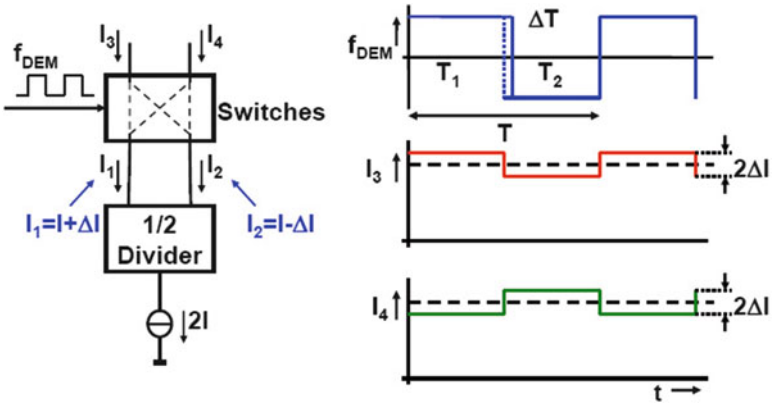


Fig. 7.38 Dynamic element matching [129]

in the digital-to-analog converter of Fig. 7.36 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 most of the current is generated by the standard current source M_1 and only a few percent is calibrated via M_{1a}.

7.7.2 Dynamic Element Matching

An often used technique is swapping components, currents, or voltages. In Fig. 7.38 the two currents I_1 and I_2 have been created by dividing a main current; however,

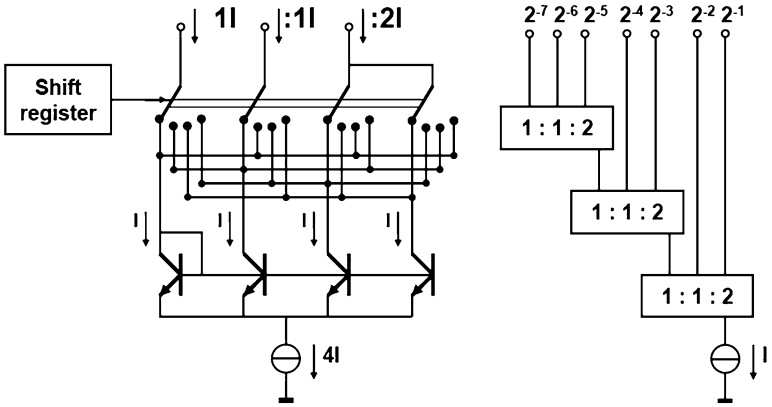


Fig. 7.39 An example of a digital-to-analog converters applying dynamic element matching [132]

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.

$$\frac{I_3 - I_4}{I_3 + I_4} = \frac{I_1 - I_2}{I_1 + I_2} \times \frac{t_1 - t_2}{t_1 + t_2}. \tag{7.37}$$

This technique is known as "dynamic element matching" [129] and is applied in many types of converters. Figure 7.39 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 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.

Example 7.12. What happens to an external input signal near to the chopping frequency? And what happens to an input signal close to the DEM frequency?

Solution. Figure 7.40 shows the chopping mechanism in comparison to the dynamic element matching. Chopping moves the signal frequency to a higher frequency before the error is introduced. After that the error and the up-modulated signal frequency are modulated again, causing the signal frequency to appear at its original position and the error at the chopping frequency.

The dynamic element mechanism directly up-modulates the error, before it is added into the signal chain.

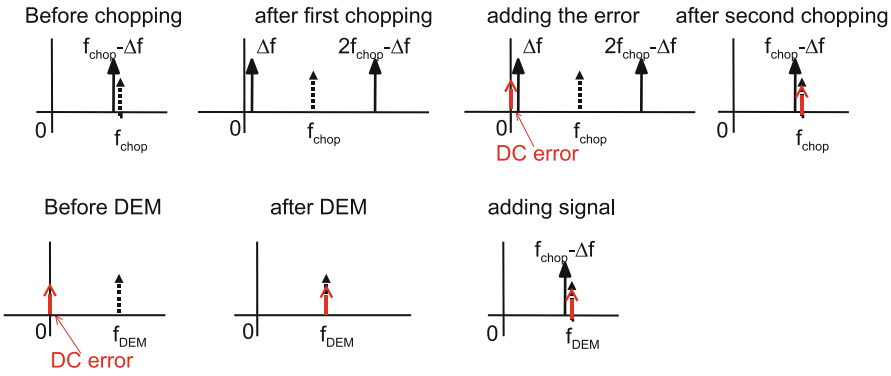


Fig. 7.40 Comparison of chopping (*above*) and dynamic element matching (*below*)

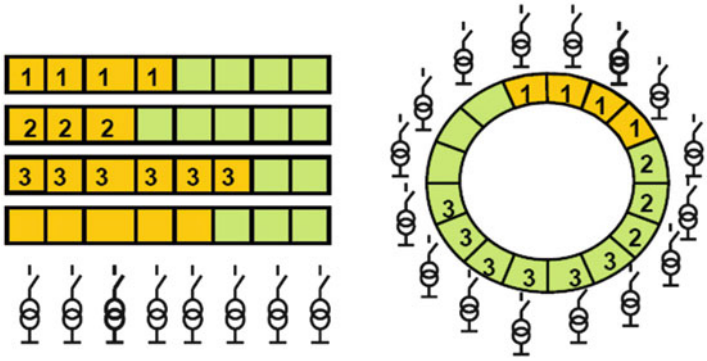


Fig. 7.41 Data weighted averaging [154, 155]

7.7.3 Data-Weighted Averaging

The elements in the current source array in Fig. 7.14 or the resistor string are always used in the same order as depicted in Fig. 7.41 (left). The signal in this digital-to-analog converter is being built up starting with the leftmost 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. 7.33.

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 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. For example, 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 [153]. 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 [154–157].[8] The elements are now functionally arranged in a circle; see Fig. 7.41 (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:

$$m_x = E(x_i) = \frac{1}{2^N} \sum_{i=0}^{i=2^N-1} x_i. \tag{7.38}$$

A sample at time $t = nT$ contains k elements and its output value will be:

$$y(nT) = \sum_{i=0}^{i=k-1} x_i = \sum_{i=0}^{i=k-1} m_x + \sum_{i=0}^{i=k-1} (x_i - m_x). \tag{7.39}$$

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 nonzero mean value; however, in practice, the result will be a minor change in gain.

The error signal contains spectral components. If there would be only one deviating element x_p with an error ε_p, Fig. 7.42, 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. For example, 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 midrange 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:

$$P_p \propto f_p \varepsilon_p^2. \tag{7.40}$$

[8] Two sources are available as the originator of DWA: Michael Story [156] and Maloberti [157].

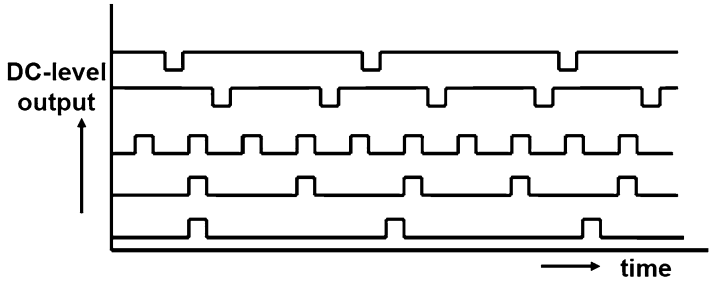


Fig. 7.42 A single mismatching current source will appear at a low frequency if the average signal value is close to the extremes of the range

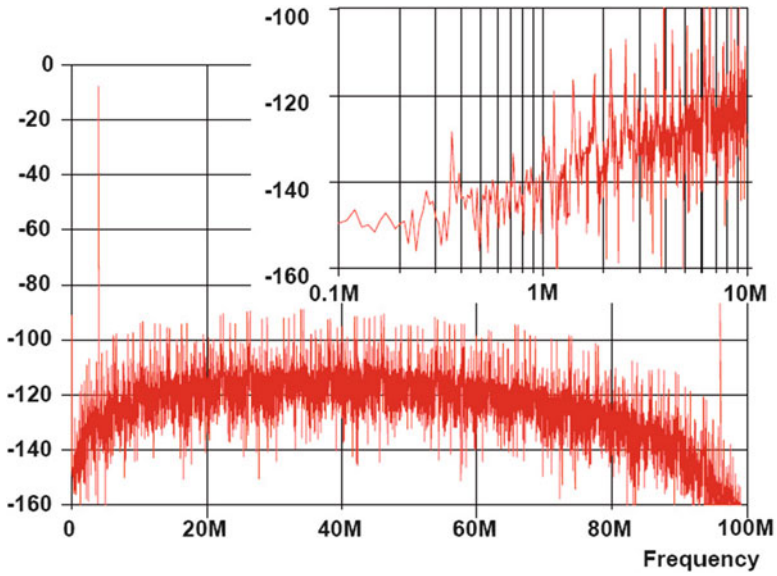


Fig. 7.43 Spectrum of a data-weighted averaging digital-to-analog converter. 256 current sources are simulated with a normally distributed mismatch of $\sigma_I/I = 1\%$ (no quantization error). The signal frequency is 3.97 MHz at 100 Ms/s. In the *inset* a small fraction of the range is visible. The spectrum contains a lot of tones

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 [154].

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. 7.43.

A formal proof requires the introduction of some stochastic mathematics [154]. The following reasoning summarizes the line of thought without pretending to be a solid mathematical proof. The second term of Eq. 7.39 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

$$\sum_{i=k}^{i=2^N-1} (x_i - m_x) = \sum_{i=0}^{i=2^N-1} (x_i - m_x) - \sum_{i=0}^{i=k-1} (x_i - m_x) = -\sum_{i=0}^{i=k-1} (x_i - m_x). \quad (7.41)$$

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

$$E(z) = \sum_{i=k}^{i=2^N-1} (x_i - m_x)(1 - z^{-\alpha}). \quad (7.42)$$

With a crude approximation the frequency behavior of the error is:

$$E(\omega) \propto \sum_{i=k}^{i=2^N-1} (x_i - m_x)(2\sin(\omega T_s/2)). \quad (7.43)$$

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 random errors. Figure 7.43 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 inset. 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 bandwidth 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 suppression at arbitrary points in the spectrum [147]. Miller and Petrie [158] shows a high-speed implementation of the digital decoding.

Unfortunately the 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] As a result often the DWA algorithm is extended with additional randomizing algorithms like the addition of dither; see Sect. 5.3.3 and [159].

[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.

7.8 Implementation Examples

7.8.1 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] allow 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 [160] with a matrix organization [161].

Figure 7.44 shows the ladder structure: the coarse ladder consists of two ladders, each with 16 large-area resistors of 250 Ω which are connected antiparallel 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 1,024 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

[10] Diffused resistors are a preferred alternative in more advanced processes.

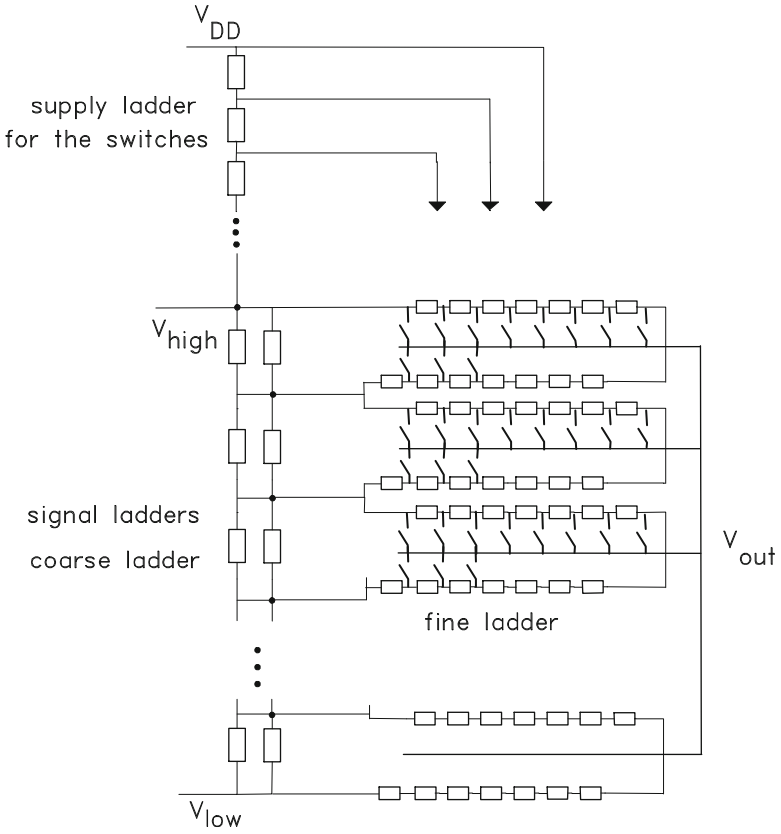


Fig. 7.44 Resistor network for a video digital-to-analog converter [162]

distortion of high-frequency output signals. This source of varying impedance has been eliminated by means of a resistive output rail. The inset in Fig. 7.45 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 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

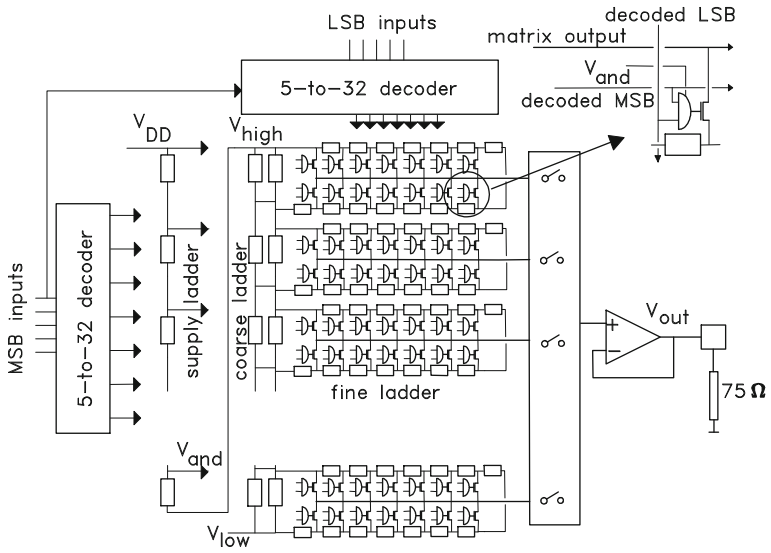


Fig. 7.45 Block diagram of the digital-to-analog converter [162]

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 with Miller compensation. The Miller feedback current flows via the source of the cascode transistor to the high-ohmic node [68]. This approach avoids a stop resistor in series with the Miller capacitor. The stop resistor is not desired because the transconductance of PMOS driver is not very constant.

The p-channel input stage operates on input voltages ranging from 0 to 2.2 V; see Fig. 7.46. 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_{series} is on the order of 25–75 Ω; it keeps up a feedback path even at frequencies at which the bondpad capacitance shorts the circuit output. It also serves as a line termination. The swing of the output load resistor is 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 (0.1–1.1 V on 75 Ω), the average

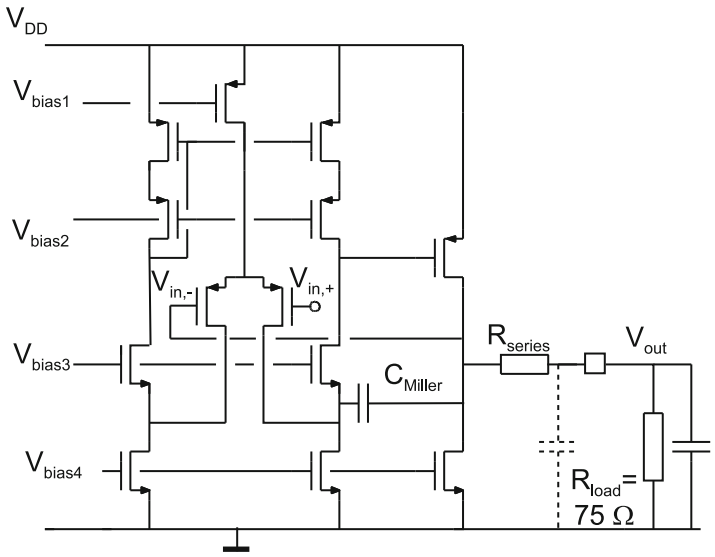


Fig. 7.46 Folded-cascode opamp circuit with Miller compensation used for the buffer

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. 7.8.3 summarizes the performance that is more extensively reported for a 1 µm technology in [162]. Figure 7.47 shows a photograph of a triple 10-bit converter.

7.8.2 Current-Domain Digital-to-Analog Conversion

High-speed CMOS digital-to-analog converters designed with a current cell matrix allow fast and accurate settling [130, 163].

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 7.16 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 7.48 shows the chip photograph of this design.

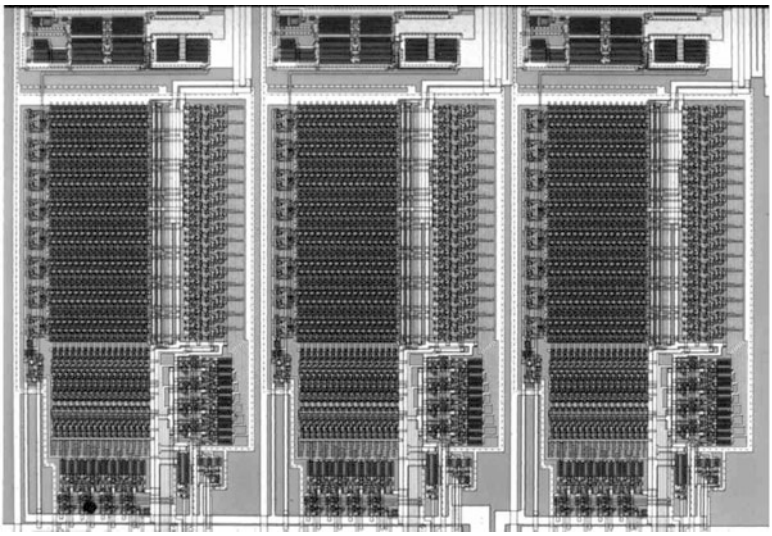


Fig. 7.47 This 10-bit triple digital-analog converter based on voltage dividers is used for applications in the video domain for the basic video colors: red, green, and blue

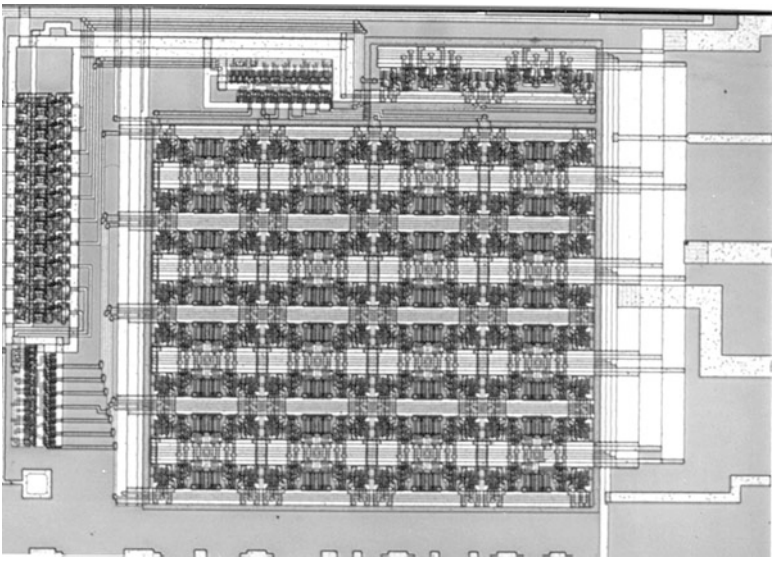


Fig. 7.48 Chip photograph of a 10-bit current digital-to-analog converter

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	1.0 µm CMOS
DC resolution	10 bit	10 bit
Sample frequency	>100 MHz	>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 (−1 dB)	20 MHz	>20 MHz
Minimum power supply (THD< −40 dB)	3 V	3 V
Output in 75 Ω	1 V	1 V
Output at minimum code	100 mV	<1 mV
Average current (50 MHz, 75 Ω)	10 mA	15 mA
Average 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

7.8.3 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 Millercompensation 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.

7.8.4 Algorithmic Charge-Based Digital-to-Analog Converter

Algorithmic principles [49, 164, 165] 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:

$$V_{\text{out}} = V_{\text{LSB}}(a_0 2^0 + a_1 2^1 + \cdots + a_p 2^p + \cdots + a_{N-1} 2^{N-1}) = V_{\text{LSB}}(W_{\text{LSB}} + W_{\text{MSB}} 2^p), \tag{7.44}$$

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 nonnegative 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. 7.49 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: $V_{\text{out}} = V_{\text{LSB}} \times W_{\text{LSB}}$. Now the output of section A is raised to $2^p V_{\text{LSB}}$ and W_{MSB} transfers take place to section B, resulting in $V_{\text{out}} = (W_{\text{LSB}} + 2^p W_{\text{MSB}}) V_{\text{LSB}}$. This implementation leads however to a signal-dependent offset at the output. The offset of section B $V_{\text{off},B}$ is transferred $(W_{\text{MSB}} + W_{\text{LSB}})$ times into the output signal:

$$V_{\text{out}} = (W_{\text{LSB}} + 2^p W_{\text{MSB}}) V_{\text{LSB}} + (W_{\text{MSB}} + W_{\text{LSB}}) V_{\text{off},B} \tag{7.45}$$

thereby creating a nonlinear dependence and signal distortion.

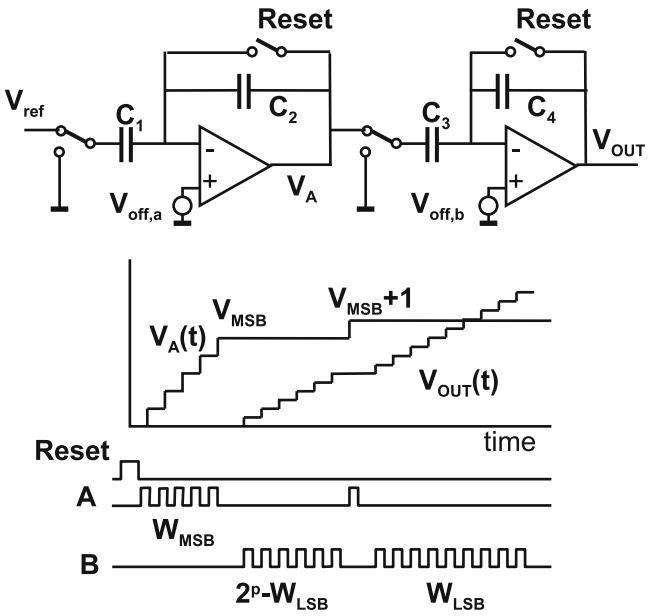


Fig. 7.49 Basic switched-capacitor network with the timing diagram for converting "1011010" [164]

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 Eq. 7.44 in:

$$V_{\text{out}} = V_{\text{ref}} \frac{C_1 C_3}{C_2 C_4} (W_{\text{MSB}} (2^p - W_{\text{LSB}}) + (W_{\text{MSB}} + 1) W_{\text{LSB}}). \tag{7.46}$$

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:

$$\begin{aligned} M_1 &= W_{\text{MSB}}, \\ M_2 &= 2^p - W_{\text{LSB}}, \\ M_3 &= 1, \\ M_4 &= W_{\text{LSB}} \end{aligned}$$

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. 7.49 the transfers for $N = 7$ and $p = 4$ have been indicated. Equation 7.46 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

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

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 Eq. 7.46 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, Eq. 7.46 is used to form the value represented by the 10 most significant bits then Eq. 7.46 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:

$$V_{\mathrm{out}} \approx V_{\mathrm{LSB}}(W_{\mathrm{LSB}} + 2^p W_{\mathrm{MSB}}) \frac{C_1 C_3}{C_2 C_4} \left(1 - \frac{W_{\mathrm{MSB}} C_1}{2 A_{\mathrm{DC}} C_2}\right). \tag{7.47}$$

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.

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.

Exercises

7.1. If the 10-bit digital-to-analog converter of Fig. 7.44 is constructed from a 4-bit coarse ladder with 16 resistors of 300 Ω and 16 6-bit fine ladder sections with 64 resistors of 50 Ω each, calculate the maximum resistance in this ladder to the grounded reference terminals.

7.2. In a resistor string digital-to-analog converter all resistors are 10% larger. What is the change in INL? What other changes must be expected?

7.3. In a resistor string digital-to-analog converter all even-numbered resistors are 2% larger, while the odd numbered are on spec. What is the change in INL and in THD? What other changes must be expected?

7.4. Show that a differential read-out of a ladder with a second-order gradient as in Fig. 7.10 (left) shows third-order distortion.

7.5. In a resistor string digital-to-analog converter all resistors are randomly 1% larger or smaller. What is the change in INL, THD, and SNR? What other changes must be expected?

7.6. The divider in the lower schematic of Fig. 7.15 has a redundant transistor "1T," What are the consequences if this transistor is removed?

7.7. Compare the binary sections of Fig. 7.15 (upper and lower). Make an estimate of the DNL difference for an equal number of bits in the binary sections.

7.8. A 5-bit fine-resistor string is directly via switches connected to a 5-bit coarse resistor string. As a consequence current from the coarse string will run via the fine-resistor string. The resistors of the coarse string are 100 Ω each. What should be the value of the fine resistors if the maximum DNL due to the resistive loading of the coarse string must be less than 0.5 LSB?

7.9. Connect in the previous example a current source to both ends of the fine ladder. Does this resolve the DNL problem? Is this solution free of other DNL problems?

7.10. An 8-bit unary current-steering digital-to-analog converter is driving a 1 kΩ load. Each current source has a parallel impedance of 5 MΩ and 0.2 pF. Sketch the resulting distortion behavior over frequency. Now the unused current is fed

in a second 1 kΩ load allowing differential operation. There is a mismatch of 1% between both load resistors. Sketch the distortion.

7.11. An 8-bit digital-to-analog converter based on a resistor string, suffers from a linear gradient of 1% over the entire structure. What is the INL, DNL, and THD?

7.12. Replace in an R-2R ladder the resistors by MOS transistors. Under what circumstances can this MOS-2MOS ladder be used as a digital-to-analog converter? Estimate in a 0.18 µm technology the performance for a 0.5 V reference voltage.

7.13. In a current-steering digital-to-analog converter the current sources suffer from 5% random mismatch. What DNL can be achieved for a 12-bit unary architecture? How much resolution can be implemented in binary format if a DNL$\leq$ 0.5 LSB must be achieved for 99.7% of the samples?

7.14. The current sources of a current-steering digital-to-analog converter must achieve a random mismatch of $\sigma_I/I < 1\%$. What are the W, L values for a CMOS65 transistor if $V_{GS} - V_T < 0.3$ V? Use the technological data from Table 2.20.

7.15. Due to mismatches the INL pattern of an 8-bit binary architecture digital-to-analog converter shows steps of 1 LSB at 1/8, 2/8,...,7/8 of the scale. Calculate the distortion.

7.16. How much distortion is caused by a gradient of 2%.

7.17. Due to unpredictable wiring patterns, the top-plate connections of the capacitors in Fig. 7.23 and 7.25 can experience a 1% additional parasitic capacitance to ground. What will be the consequence for the achievable resolution?

7.18. A spurious level of -80 dB is required for a digital-to-analog converter. What is the maximum gradient that can be tolerated and what is the maximum random mismatch?

7.19. A digital-to-analog converter uses an output buffer. What is the maximum frequency sine wave of 1 $V_{peak-peak}$ that can be delivered to a 3 pF capacitor if the output current is limited to 100 µA? What changes if a 10 kΩ resistor is connected parallel to the capacitor?

7.20. A data-weighted averaging algorithm is used to eliminate conversion errors in a bandwidth located around $f_s/8$. Construct a sample sequence that will reduce the errors in that bandwidth; use [147].

7.21. Modify Example 7.4 by expanding the approximation to $1/(1-a) \approx (1+a+a^2) \quad |a| \ll 1$. Show that an odd distortion component remains.

Chapter 8
Analog-to-Digital Conversion

Abstract Several classifications exist of Nyquist-rate analog-to-digital converters. In this chapter the converters are subdivided in parallel search, sequential search, and linear search. Each of these architectures requires a comparator. Therefore this building block is extensively analyzed in all its aspects. The section is concluded with a comparator catalog.

The full-flash converter is the conversion solution for the highest speed range. Moreover it is a building block for more complex converters. Variants, such as gain stage, interpolation, and folding are analyzed and described.

The sub-ranging methods and pipeline converters are the solutions for the medium speed range demands. Next to a treatise on the various aspects of the architecture an analysis is made of the error sources, calibration techniques, and design issues.

In the next sections the successive approximation and linear topologies are discussed. These topologies are slower but receive today more attention as a massive parallelism allows them to compete with the performance of the pipeline converter. The issues associated with multiplexing are analyzed.

Finally some less prominent ideas for conversion are briefly highlighted.

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 quantization. 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 preprocessed 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:

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_8,

© Springer Science+Business Media, LLC 2013

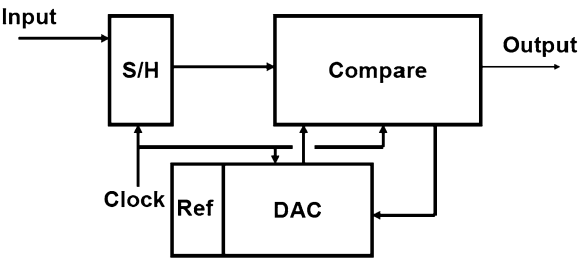


Fig. 8.1 Components in analog-to-digital conversion

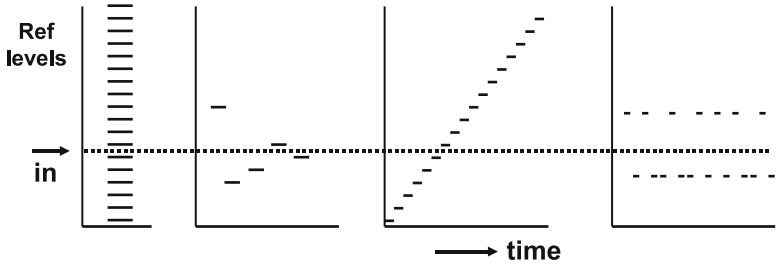


Fig. 8.2 Classification of analog to digital conversion principles. In the vertical dimension the available levels are depicted. *Left*: in direct conversion all levels are provided. The sequential search algorithm chooses the next level based on the information of the previous comparisons. In linear search each level is available at consecutive clock periods. Finally the oversampled conversion switches between a few levels

- 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. Applications 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, pipeline conversion, and multistep conversion. The combination of high accuracy and rather high speed makes these converters suitable for many industrial and communication applications.

- 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.

8.1 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.

[1]This bandwidth can start at DC, but also far above f_s using sub-sampling in Sect. 12.1.4.

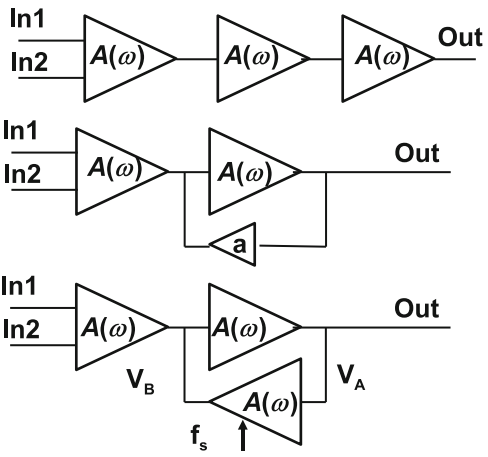


Fig. 8.3 Three basic comparator designs: straightforward amplification, amplifier with hysteresis, and a latched or regenerative amplifier

- 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 $\tau = A/\omega_{\mathrm{UGBW}}$. The response of a cascade of M comparators is

$$H(\omega) = \left(\frac{A}{1+s\tau}\right)^M = \left(\frac{\omega_{\mathrm{UGBW}}}{s+\omega_{\mathrm{UGBW}}}\right)^M .$$

An excitation with a step function results in

$$V_{\mathrm{out}}(t) = A^M V_{\mathrm{in}}(t=0)\left[1 - \mathrm{e}^{-t/\tau}\sum_{i=1}^{M}\frac{(t/\tau)^{i-1}}{(i-1)!}\right]. \tag{8.1}$$

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 straightforward 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:

$$\begin{aligned} V_A(s) &= v_B(t=0)\frac{-A(1+s\tau)}{s((1+s\tau)^2 - A^2)} \\ &= v_B(t=0)\left(\frac{-A/(1-A^2)}{s} + \frac{A\tau/(2(1+A))}{1+s\tau+A} + \frac{A\tau/(2(1-A))}{1+s\tau-A}\right). \end{aligned}$$

The inverse Laplace transform gives

$$v_A(t) = v_B(t=0)\left(\frac{-A}{(1-A^2)} + \frac{A}{2(1+A)}\mathrm{e}^{-(A+1)t/\tau} + \frac{A}{2(1-A)}\mathrm{e}^{(A-1)t/\tau}\right). \quad (8.2)$$

After a few time constants have elapsed only the middle term in brackets is relevant:

$$v_A(t) \approx \frac{-v_B(t=0)}{2}\mathrm{e}^{-(A+1)t/\tau} \approx \frac{-v_B(t=0)}{2}\mathrm{e}^{\omega_{\mathrm{UGBW}}t}. \quad (8.3)$$

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 amplifiers with a gain A 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.

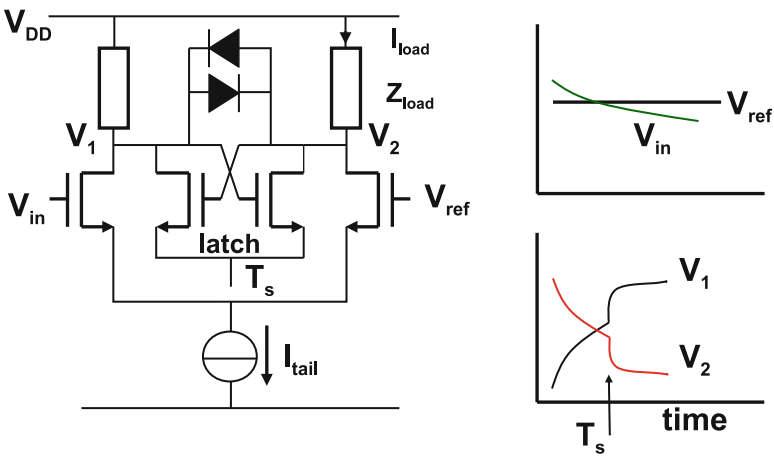


Fig. 8.4 A simple regenerative comparator circuit

8.1.1 Dynamic Behavior of the Comparator

Most comparators use an input differential pair. Here 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. The 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: $g_{m,\text{in}} \times Z_{\text{load}}$. 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 very high values. Too much gain will cause saturation, slewing, and unnecessary kickback. The achievable speed performance is limited by the unity-gain bandwidth; 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 small-signal mode the input transconductance and the capacitive load of the nodes V_1 and V_2 determine the bandwidth:

$$\omega = \frac{g_{m,\text{input}}}{C_{\text{load}}}. \tag{8.4}$$

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 input terminals will become stronger enhancing kickback.

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:

$$\begin{aligned} \text{slew-rate} = \frac{\mathrm{d}V}{\mathrm{d}t} &= \frac{I_{\text{tail}}}{C_{\text{load}}} \\ C_{\text{load}} \frac{\mathrm{d}V_1(t)}{\mathrm{d}t} &\leq I_{\text{tail}} \\ 2\pi f_{\text{in}} V_{1,\text{max}} C_{\text{load}} &\leq I_{\text{tail}}, \end{aligned} \tag{8.5}$$

  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 (nonlinear 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

[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.

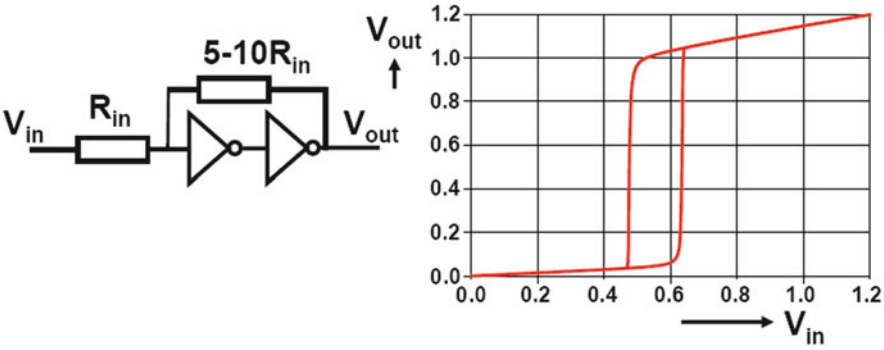


Fig. 8.5 Hysteresis created by positive feedback in a two-stage buffer

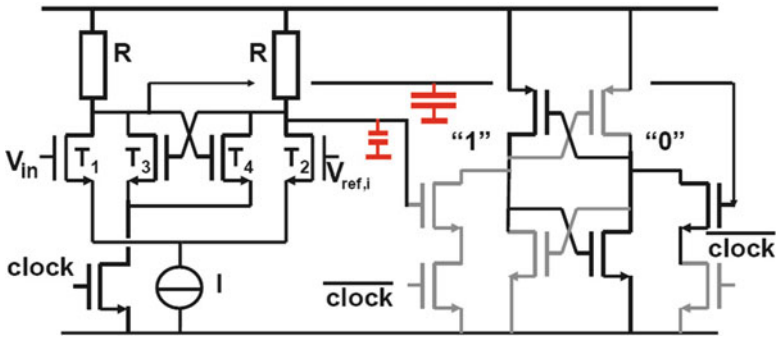


Fig. 8.6 Hysteresis created by the signal-dependent loading of the input latch by the second latch. The *gray*-colored transistors are not conductive and form only a small load capacitor

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 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.

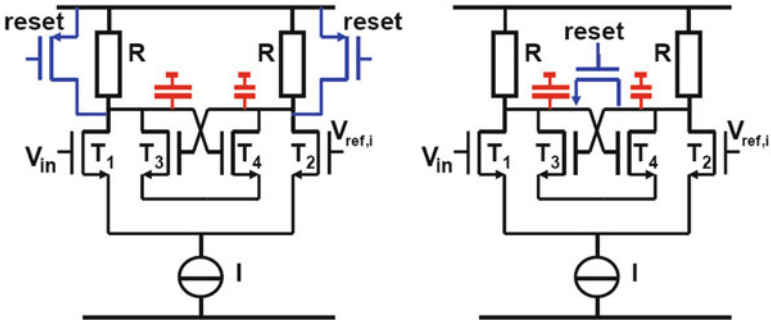


Fig. 8.7 Two methods to clear the state in comparator circuit

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 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 intact 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 nonlinearities. Also in multiplexed analog converters the random offset returns as a spurious tone. The input-referred random offset σ_{Vin} must be designed to a value a factor 5–10 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 for conversion errors on system level.

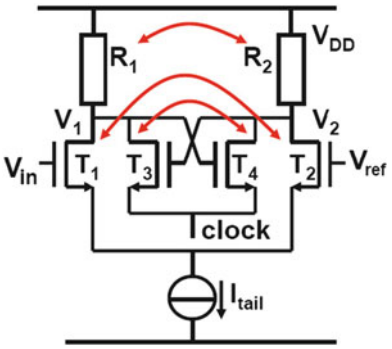


Fig. 8.8 Mismatch sources in a comparator

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 contributions can be referred back to an equivalent input-referred random mismatch σ_{Vin}. With the help of Eq. 2.18, the input-referred random offset is described as

$$\sigma^2_{\mathrm{Vin}} = \left(\frac{\partial V_{\mathrm{in}}}{\partial V_{\mathrm{T},12}}\right)^2 \sigma^2_{V_{\mathrm{T},12}} + \left(\frac{\partial V_{\mathrm{in}}}{\partial V_{\mathrm{T},34}}\right)^2 \sigma^2_{V_{\mathrm{T},34}} + \left(\frac{\partial V_{\mathrm{in}}}{\partial R_{1,2}}\right)^2 \sigma^2_R$$

$$\sigma^2_{\mathrm{Vin}} = \sigma^2_{V_{\mathrm{T},12}} + \frac{g^2_{m,34}}{g^2_{m,12}} \sigma^2_{V_{\mathrm{T},34}} + \frac{I^2_{\mathrm{load}}}{g^2_{m,12}} \frac{\sigma^2_R}{R^2_{1,2}}. \tag{8.6}$$

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 nonzero 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

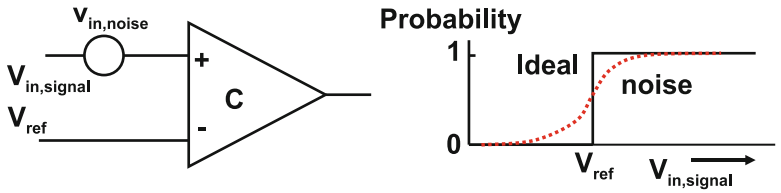


Fig. 8.9 Thermal noise will create a probabilistic behavior of the comparator decision (*dotted line*)

in a similar manner as for mismatch. If the input transconductance dominates the noise, the effective value of the input-referred noise is

$$v_{\text{in,noise}} = \sqrt{4kT\frac{\text{BW}}{g_m}} \tag{8.7}$$

and with a bandwidth of 10 GHz and an input transconductance of 1 mA/V the value for the input-referred rms noise is $v_{\text{in,noise}} = 400\,\mu\text{V}$. 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 the time a comparator or latch needs to form a digital signal depends on the initial overdrive voltage. In other words, the input differential voltage determines the delay of the comparator.[3]

For very small overdrive voltages, there is a fraction α of the LSB size where the comparator has insufficient voltage difference on its nodes to reach a decision in the limited time T_s of a clock pulse. Now the comparator will not generate a clear "zero" or "one" output level. Without unambiguous signals the succeeding logical circuitry can generate large errors, 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 develop exponentially in time with a time constant τ; see Eq. 8.3. This time constant

[3]This voltage-delay relation can be exploited: measuring the delay is used to quantify the input voltage and create more resolution.

is determined by the internal node capacitance and the transconductance of the latch transistors. Now α can be approximated by

$$\alpha \approx \frac{V_{\text{latch}}}{V_{\text{LSB}}} \mathrm{e}^{-T_s/\tau} \quad \Rightarrow \quad \text{BER} \approx 2^N \mathrm{e}^{-T_s/\tau}, \tag{8.8}$$

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 assumes 5 fF total capacitance for 1 µm gate width, with 5 µA/V transconductance for the same gate width. A time period of $T_s = 20$ 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 1 bit from the preceding code.

From a fundamental point of view the BER cannot 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.

Example 8.1. Calculate the bit error rate of a latch in 0.18 µm CMOS with 0.5/0.18 µm NMOS transistors and 100 µA total current. The latch is used in a 0.5 Gs/s application.

Solution. The transconductance of the transistors is $g_m = \sqrt{2I(W/L)\beta_\square} = 0.3$ mA/V. The gate capacitance is $0.5 \times 0.18 \times 8.3$ fF = 0.75 fF. The succeeding stage will load latch with double or triple this amount. Some wiring and diffusion capacitance will result in a total load of around 10 fF, leading to a time constant of 35 ps. The decision period is a 40% fraction of the 2 ns sample rate. So the bit error rate is estimated at $\approx 2^N \mathrm{e}^{-T_s/\tau} = 2^N \times 1.2 \times 10^{-10}$. For a 7-bit resolution converter this would lead to BER = 1.5×10^{-8}.

8.1.5 Kickback

The comparator is a nonlinear 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.

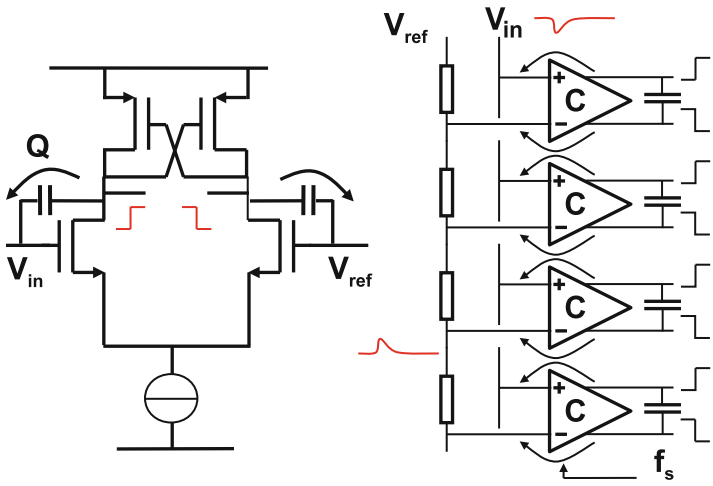


Fig. 8.10 Kick-back in a comparator circuit

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 "kickback." In some comparator designs these transitions can be as large as the power supply voltage. Figure 8.10(right) shows a full-flash circuit. In this architecture the kickback 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 kickback.

Reducing the kickback signal requires to decrease the coupling capacitances by reducing the dimensions of the input-related circuit elements. Inside the comparator the swing of the signals must be sufficient to drive the next stages but can be kept as low as a few tenths of a volt by swing limiting measures. The additional advantage is that the transistors in the comparator remain in inversion, thereby avoiding time loss due to recharging. A rigorous method applies an additional preamplifier. Unfortunately these 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.

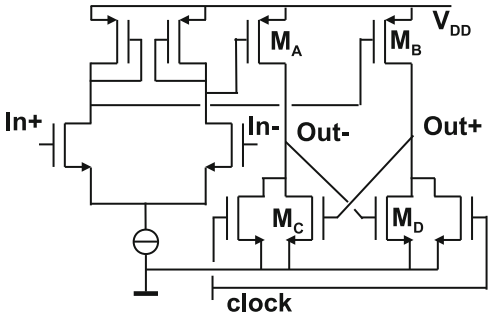


Fig. 8.11 A comparator in CMOS based on an NMOS enhancement/depletion design [166]

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 [166] 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 [149, 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 preamplifier from the latch. The problem with an intermediate node is that it creates an additional pole and slows down the design. 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 kickback 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 also consists of a preamplifier (left) and a latch [167]. 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.

The comparator in Fig. 8.13 is designed for a folding analog-to-digital converter [168]. It consists of an input stage that feeds its differential current in a cascode stage. This construction avoids an intermediate pole that would slow down the

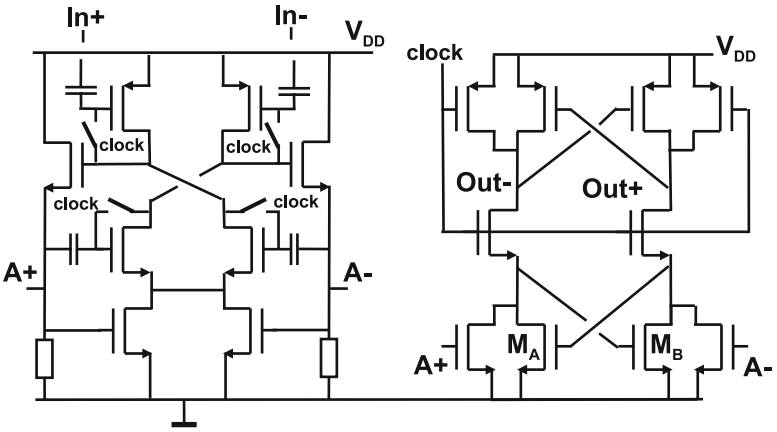


Fig. 8.12 A comparator based on the work in [167]

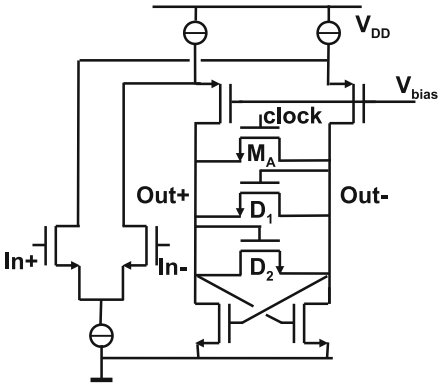


Fig. 8.13 A comparator designed for a folding analog-to-digital converter [168]

circuit. On the other hand a reasonable shielding for kickback 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.

Figure 8.14 uses three phases to reach a decision [169]. When $\overline{clock}$ is high the comparator is in reset mode. The signal is amplified over the resistance of the reset transistor. After $\overline{clock}$ goes low, during a short period, the lower latch can amplify the signal. This amplification allows to suppress the mismatch of the upper latch section. After *clock* goes high, this amplified signal is boosted to digital levels. Some sensitivity to the nonoverlapping clock phases is present, as well as a considerable kickback.

The comparator in Fig. 8.15 is based on former bipolar designs [170]. The comparator routes a constant current through either the amplification side (clock is

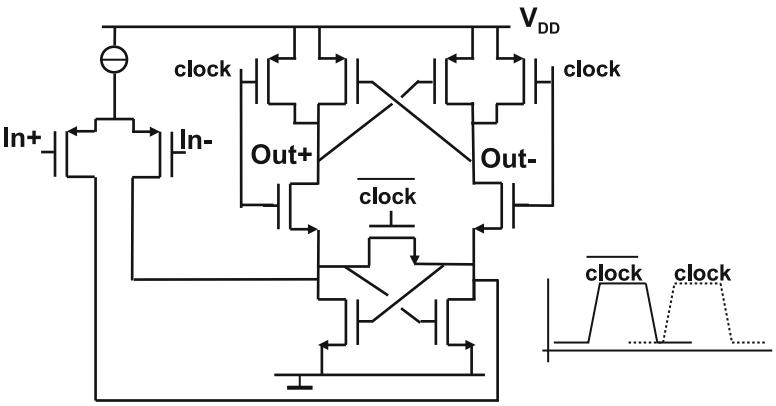


Fig. 8.14 A comparator for high speed operation [169]

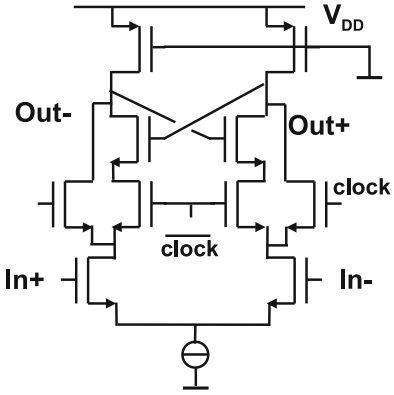


Fig. 8.15 A comparator designed for a folding analog-to-digital converter [170]

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 kickback 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. 8.16). The middle PMOS in the preamplifier serves as a load for the input pair. The sources of the input pair are switched to the positive power supply at the end of the amplification, which produces kickback noise via the coupling of the gate. The latch (right) is also used in the "StrongArm" design [172] 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.

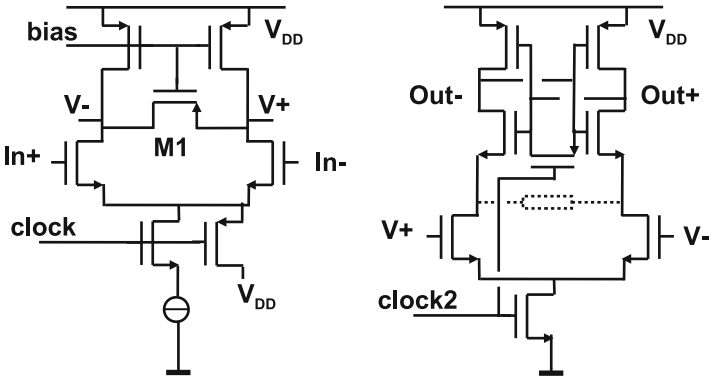


Fig. 8.16 A comparator for high-speed operation [171]. The latch (*right*) is based on a design for the "StrongArm" processor [172]

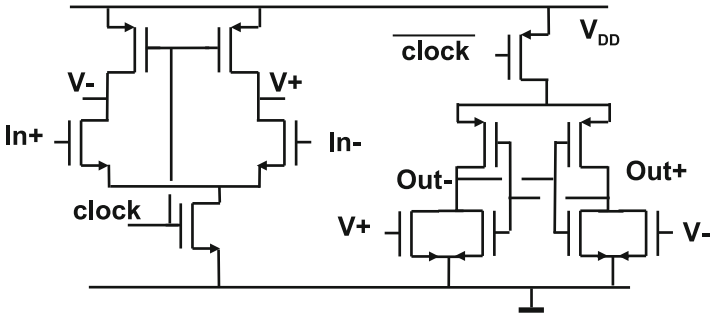


Fig. 8.17 A double-tail comparator combines accuracy with speed [174]

After full settling the latch is consuming no current. The 1-sigma input-referred random offset in [171] 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 [173] where a good energy efficiency is achieved.

The comparator in Fig. 8.17 [174] is a further development of Fig. 8.16. The amplification is now carried out in a pre-stage, thereby reducing the amount of stacked transistors and the kickback noise. The design does not require any DC current, which makes it very suited for achieving low-power operation.

A more advanced comparator is shown in Fig. 8.18. The design is differential, thereby eliminating the PSRR problems. The random input offset of the comparator is estimated at 36 mV 1-sigma [99]. Still some time is needed for the bias setting in this 0.18 µm CMOS design. Interleaving of two input stages reduces in this converter this speed penalty.

Example 8.2. A differential NMOST input pair ($W/L = 50/2$) is loaded with a PMOST current mirror ($W/L = 36/1\,\mu\text{m}$) in a process with $A_{V_{T},N} =$

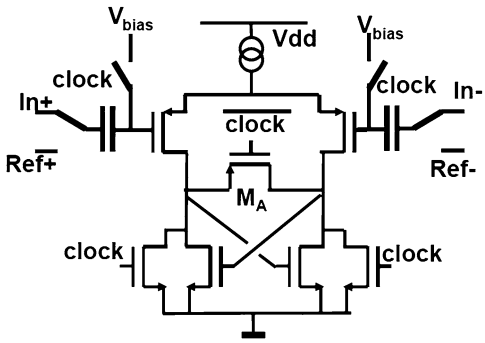


Fig. 8.18 A comparator design for an interleaved pipeline converter [99]

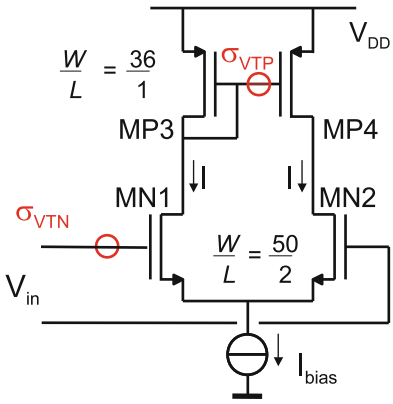


Fig. 8.19 The circuit schematic

$A_{V_{T,P}}$=6 mV µm, and the ratio between the beta square (for $W/L = 1$) is $N/P = 3$. Calculate the input-referred mismatch.

Solution. In Fig. 8.19 the schematic of this circuit is drawn. With the help of Eq. 2.110 the standard deviations for the NMOS and PMOS threshold voltages are found: $\sigma_{V_{TN}} = 6/\sqrt{50 \times 2} = 0.6\,\text{mV}$, and $\sigma_{V_{TP}} = 1\,\text{mV}$. The PMOS transistor mismatch must be referred back to the input. This threshold mismatch translates into a PMOS current mismatch via $g_{m,P}\sigma_{V_{TP}}$. The effect at the input terminals is the input-referred mismatch voltage. This voltage must generate in the NMOS transistors a current that compensates the mismatch current generated by the PMOS devices: $g_{m,N}\sigma_{V_{\text{in}}P} = g_{m,P}\sigma_{V_{TP}}$. So the total input-referred mismatch is composed of the NMOS and PMOS contributions:

$$\sigma_{\text{Vintot}} = \sqrt{\frac{g_{m,P}^2}{g_{m,N}^2}\sigma_{V_{TP}}^2 + \sigma_{V_{TN}}^2} = 0.92\,\text{mV}.$$

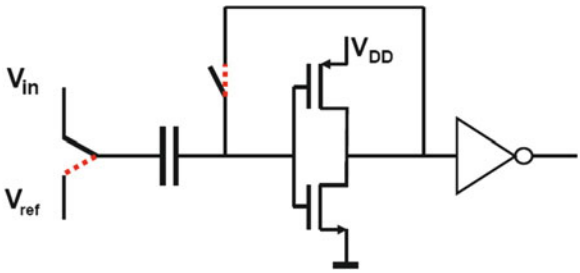


Fig. 8.20 An auto-zero comparator as used in [175]

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 "auto-zero" [175] or cancel [176] this offset. The basic schematic of Fig. 8.20 is similar to the track-and-hold circuit in Fig. 4.20. 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 the switches toggle the change on the input of the comparator equals $V_{\text{in}} - V_{\text{ref}}$. 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 [175] 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. 8.20 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 cancellation 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. 8.20, 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

$$V_{\text{out, error}} = e(z)(1 - z^{-0.5}). \tag{8.9}$$

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.

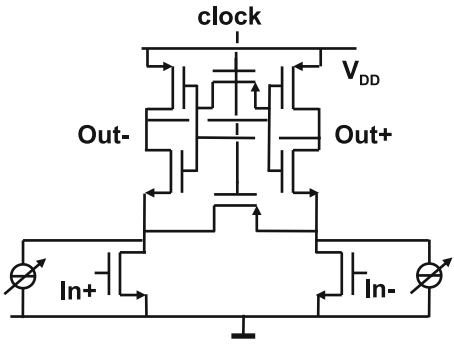


Fig. 8.21 The current sources are controlled to auto-zero this comparator [177]

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. 7.7. Another class of auto-zero mechanisms uses adjustable voltages or currents to reduce the input-referred offset. Figure 8.21 shows the basic idea implemented with two controlled current sources [177]. The necessary corrections are carried out by means of an algorithm, either at start-up or during operation. If the chip contains non-volatile memory also programming after production is possible.

8.1.8 Track-and-Hold Plus 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. Fukushima et al. [175] and Kusumoto et al. [178] 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. 8.20. 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 [179], 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 8.22 shows the comparator that was used in several analog-to-digital converter designs (see Sects. 8.9.1–8.9.3).

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

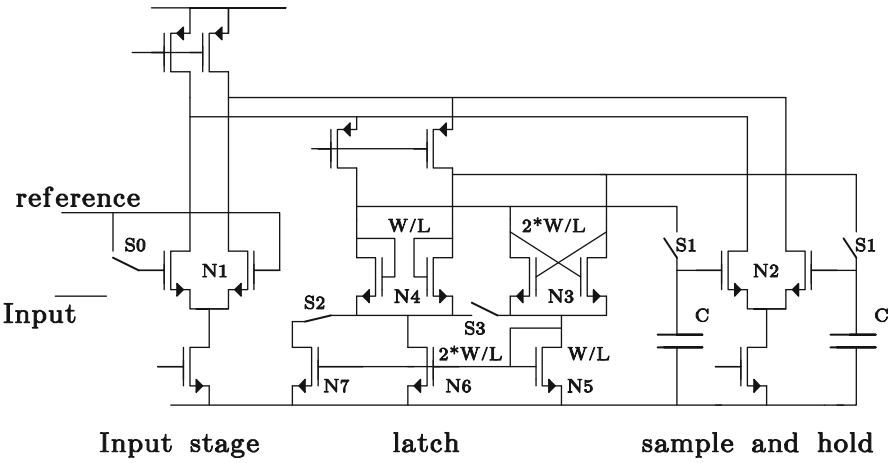


Fig. 8.22 The basic comparator schematic. Switches are shown in the position for the amplification phase (Design: J.v.Rens)

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 $V_{\text{in}} - V_{\text{ref}} + V_{\text{off}}$ 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 $V_{\text{in}} - V_{\text{ref}} + V_{\text{off}}$, 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 eight 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.

8.2 Full-Flash Converters

Full-flash[4] 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 [180]. Another 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 $2^N - 1$ comparators; see Fig. 8.23. A decoder combines the comparator decisions to a digital output word.

The input signal is compared to all reference levels simultaneously. 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 digital 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 converters; however, their complexity and their power consumption grow with the number of comparator levels 2^N. Also the area and the input capacitance grow exponentially with N.

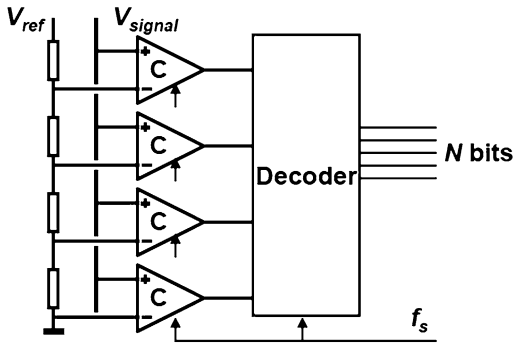


Fig. 8.23 A "full-flash" analog-digital converter

[4]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 subranging architecture.

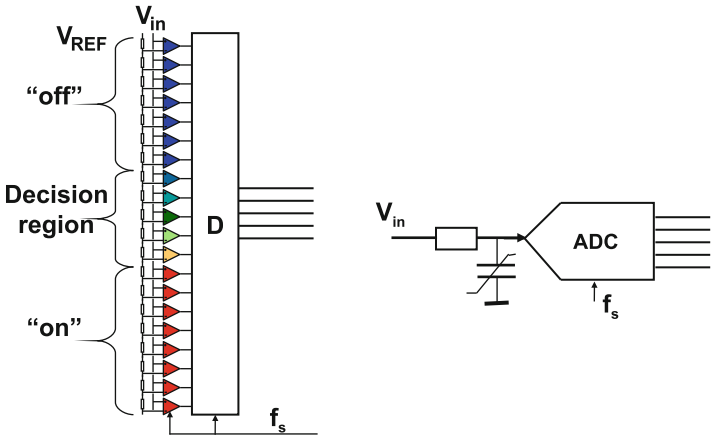


Fig. 8.24 A "full-flash" converter presents a non-linear input impedance

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-level single-ended 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. 8.24, 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 for single-sided operation 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 nonlinear capacitor, the capacitive current is found using Eq. 2.56:

$$I_C = C(V_C(t)) \frac{\mathrm{d}V_C(t)}{\mathrm{d}t} + V_C(t) \frac{\mathrm{d}C(V_C(t))}{\mathrm{d}t} = C_0 \frac{\mathrm{d}V_C(t)}{\mathrm{d}t} + \frac{2V_C(t)\Delta C}{V_{\text{ref}}} \frac{\mathrm{d}V_C(t)}{\mathrm{d}t}. \tag{8.11}$$

A correct analysis requires to equate this current to the current through the resistor, resulting in a nonlinear differential equation. The substitution $V_C(t) = 0.5V_{\text{ref}} + 0.5V_{\text{ref}} \sin(\omega t)$ assumes that the nonlinear term is relatively small and will not substantially change $V_C(t)$. Evaluation of the last part of the equation leads to a

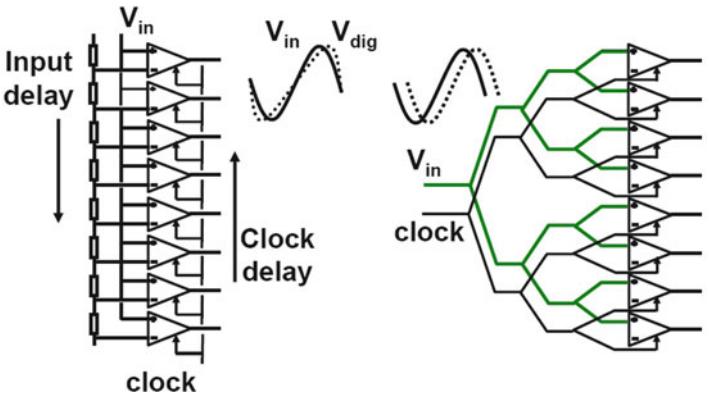


Fig. 8.25 A simple distribution strategy for input signal and clock can easily lead to performance loss at high frequencies (*left*). A tree-like lay-out is necessary to avoid delay differences

second harmonic current, which is multiplied with the input resistor R and compared to the first harmonic term in $V_C(t)$:

$$\mathrm{HD2} = \frac{\omega R \Delta C}{2}. \tag{8.12}$$

This estimate for HD2 is proportional to the signal frequency, the input impedance, and the amount of capacitive variation. For $f = 1\,\mathrm{GHz}$, $R = 50\,\Omega$, and a capacitive variation of 100 fF a distortion level of $-37\,\mathrm{dB}$ results. Changing the comparator into a full-differential design with two input terminals and two reference voltage terminals will solve this problem at the expense of more current and area.

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. 3.1.7. 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. 8.25(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:

$$\begin{aligned} V_{\mathrm{in}}(t + \Delta T(t)) &= V_a \sin(\omega(t + \Delta T \sin(\omega t))) \approx V_a \sin(\omega t) + V_a \omega \Delta T \sin(\omega t)\cos(\omega t) \\ &= V_a \sin(\omega t) + 0.5 V_a \omega \Delta T (1 - \cos(2\omega t)). \end{aligned} \tag{8.13}$$

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 treelike structure as in Fig. 8.25(right) is applied in high-performance converters.

Example 8.3. A group of comparators is on regular intervals connected to the input signal line of total length 640 µm and width 1 µm. Every comparator has an input capacitance of 0.1 pF. The clock arrives synchronously at all comparators. Estimate the expected distortion for a 100 MHz signal.

Solution. The total resistance of the wire is $(640/1) \times 0.1\,\Omega = 64\,\Omega$. The total capacitance is 6.4 pF. In Sect. 7.2.1 an analysis was made for a double terminated structure with distributed resistance and capacitance. The structure in the present example can be considered a half of that problem. So the time constant is easily found by considering that a double structure behaves with $\tau R_{\text{tot}} C_{\text{tot}}/8$ leading to $\tau = 0.2$ ns. The relative magnitude of the second harmonic component is estimated with Eq. 8.13 as 6%.

8.2.1 Ladder Implementation

In Sect. 7.2.2 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. See also example 7.1.

$$\Delta V_{\text{middle}} = R_{\text{tap}} I_{\text{tap}} 2^{2N}/8. \tag{8.16}$$

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 kickback effects [181].

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 bits of resolution is required, some precautions have to be taken to mitigate the gradients. An antiparallel 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 kickback, 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.

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 $\sigma_{\Delta V_{\text{be}}} = 0.3$ mV. A pair of NMOS transistors with small gate lengths show random offsets in the order of $\sigma_{\Delta V_{\text{T}}} = 2$–$6$ 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 8.26 shows the distribution of the actual DNL of 1,000 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 nonlinear function. The comparator random offset was in this example chosen to be $\sigma = 0.15 V_{\text{LSB}}$, corresponding to $\sigma = 0.586 \times 10^{-3} V_{\text{ref}}$.

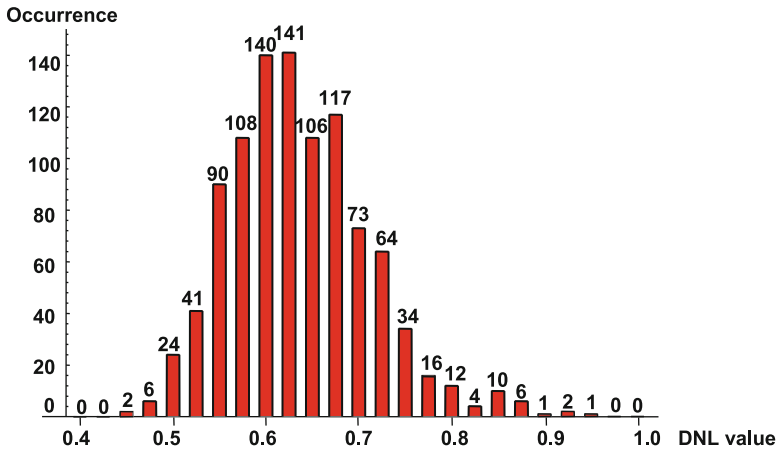


Fig. 8.26 Typical histogram of DNL values for an 8-bit full-flash architecture, with $\sigma = 0.15V_{\text{LSB}}$. The mean DNL in this simulation is 0.64 LSB

There are almost no trials with an actual DNL$<$ 0.5 LSB, although all trials result in monotonicity. For 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:

$$\begin{aligned} \text{Yield} = (1-p)^{2^N-2} < 0.5 \qquad & \rightarrow p > 0.9973 \\ \text{from Table 2.11} \quad p = 0.9973 \leftrightarrow \alpha = 3.0 \qquad & \rightarrow \frac{\text{DNL} \times \text{V}_{\text{LSB}}}{\sigma\sqrt{2}} = 3.0 \\ \text{with } \sigma = 0.15V_{\text{LSB}} \qquad & \rightarrow \text{DNL} = 0.64. \end{aligned} \tag{8.18}$$

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

$$\begin{aligned} &\text{Yield} = (1-p)^{(2^N-2)} \quad \text{with} \\ &p = P\left(\frac{V_{j+1} - V_j - V_{\text{LSB}}}{\sigma\sqrt{2}} < \frac{-V_{\text{LSB}}}{\sigma\sqrt{2}}\right). \end{aligned} \tag{8.19}$$

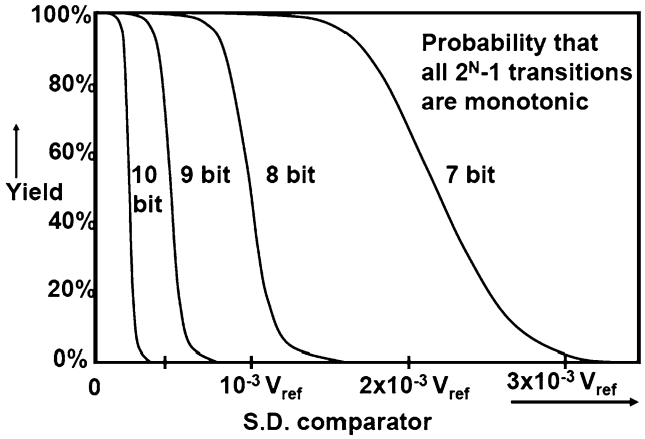


Fig. 8.27 Yield on monotonicity versus the standard deviation of the comparator random offset

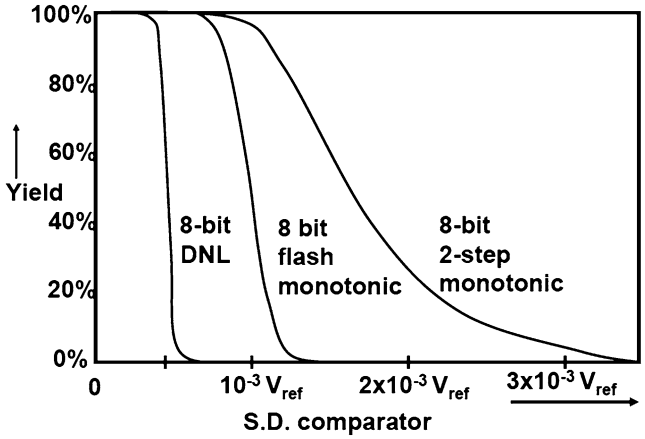


Fig. 8.28 Yield on DNL<0.5 LSB, on monotonicity for 8-bit full-flash, and on monotonicity for an 8-bit 2-step subrange architecture

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 mV. The probability function is normalized to a standard normal curve $N(0,1)$.

Figure 8.27 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 [182]. For higher accuracies trimming, higher input voltages or other forms of offset correction are needed. The 8-bit converter from Fig. 8.26 with $\sigma = 0.586 \times 10^{-3} V_{\text{ref}}$ is indeed monotonic with close to 100% yield.

At first sight 1-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. 8.28 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 95–99% 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 1 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.

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 1 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 1 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 three 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

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 $N(0,\sigma)$	$S_N \approx 3\cdots4$	$S_{N+1} \approx S_N + 0.3$
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	$WL = A_{V_T}^2/\sigma_N^2$	$WL = A_{V_T}^2/\sigma_{N+1}^2$
10	Capacitance of all gates	$3\times2^N A_{V_T}^2 C_{\text{ox}}/\sigma_N^2 = 18S_N^2 2^{3N} A_{V_T}^2 C_{\text{ox}}/V_{\text{ref}}^2$	$3\times2^{N+1} A_{V_T}^2 C_{\text{ox}}/\sigma_{N+1}^2 = 18S_{N+1}^2 2^{3(N+1)} A_{V_T}^2 C_{\text{ox}}/V_{\text{ref}}^2$

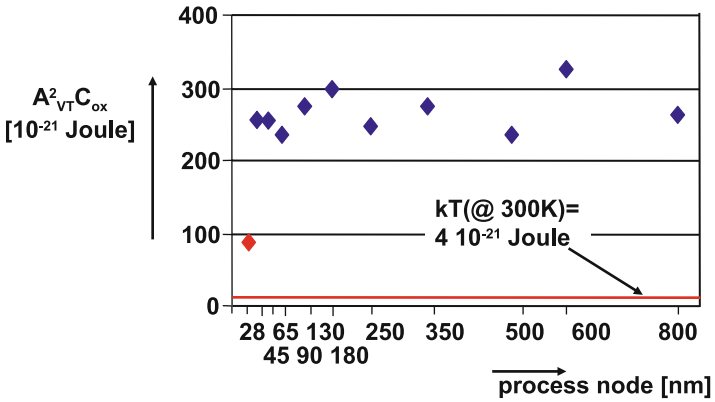


Fig. 8.29 The technology factor $A_{V_T}^2 C_{\text{ox}}$ versus process generation

$$\frac{C_{\text{in},N+1}}{C_{\text{in},N}} = \frac{8S_{N+1}^2}{S_N^2} \approx 10 \quad \text{for } N = 5,\ldots,8. \tag{8.21}$$

The last lines in Table 8.1 show that the technology factor $A_{V_T}^2 C_{\text{ox}}$ with dimension of energy, joule, is determining the outcome. This factor can be plotted versus the process generation, in Fig. 8.29. It is clear that the energy associated with threshold mismatch hardly changes over time. Given the reducing power supply (from 3.3 V down to 1.1 V) this implies that the power efficiency of straight-forward full-flash architectures will not scale and perhaps even deteriorate. The first indications on the performance in 28-nm high-k factor metal-gate processes are promising.

Table 8.2 Comparison of 7- and 8-bit full-flash analog-to-digital converters

		7-bit	8-bit
1	Input range	1 V	1 V
2	LSB size	7.8 mV	3.9 mV
3	Number of comparators	127	255
4	Probability on monotonicity per comparator pair for 95% ADC yield	0.99960	0.99980
5	Required σ's in $N(0,\sigma)$	3.35	3.55
6	Allowed input-referred random error	2.33 mV	1.10 mV
7	Number of MOS pairs in comparator	3	3
8	Random error per input pair	1.34 mV	0.63 mV
9	Area per MOS $A_{V_T} = 3.5\,\text{mV}\,\mu\text{m}$	6.8 μm^2	30.7 μm^2
9	Capacitance per MOS $C_{ox} = 12.6\,\text{fF}/\mu\text{m}^2$	86 fF	387 fF
11	Capacitance of all input gates	11 pF	99 pF

Table 8.3 The calculated yield for 6, 7, and 8 bit

N	V_{LSB} (mV)	x	$P(x)$	$P^{2^N}(x)$
6	15.6	7.1	$1-10^{-11}$	0.999999
7	7.8	3.55	0.0002	0.975
8	3.9	1.77	0.0384	4×10^{-5}

Example 8.4. Compare the input capacitance of a 7-bit and an 8-bit full-flash converter with 1-V input range in a 65-nm process.

Solution. Table 8.2 shows the effect of increasing the resolution with 1 bit in a full-flash analog-to-digital converter.

Example 8.5. In a simple flash converter the size of the transistor input pair of the comparators is 20/5 µm in a process with $A_{V_T} = 15\,\text{mV}\,\mu\text{m}$. The expected input signal is $1\,\text{V}_{\text{peak}-\text{peak}}$. What resolution limit (if monotonicity is required with 99% yield) do you expect?

Solution. Monotonicity means that in case of an increasing input voltage, the comparator connected to a lower reference voltage switches before the comparator connected to a V_{LSB} higher reference voltage. The standard deviation of the input pair and trip level is calculated as: $\sigma_{\text{trip},i} = A_{V_T}/\sqrt{WL} = 1.5\,\text{mV}$, ignoring other contributions. The nominal difference between two trip levels is $V_{LSB} = 2\text{V}/2^N$. As both trip levels suffer from uncertainty, the standard deviation of the difference equals $\sigma_{i+1,i} = \sqrt{\sigma^2_{\text{trip},i+1} + \sigma^2_{\text{trip},i}} = 1.5\sqrt{2}\,\text{mV}$. The probability $P(x)$ that the difference between two trip levels stays within V_{LSB} corresponds to the value of a normal $N(0,1)$ distribution for the quantity $x > \sigma_{i+1,i}/V_{LSB}$.

Monotonicity in a converter requires that all 2^N differences are within one LSB. So this probability requires to multiply all 2^N individual probabilities. Table 8.3 shows that a 7-bit converter is at the edge of acceptable yield.

The faster way to this result uses Fig. 8.27.

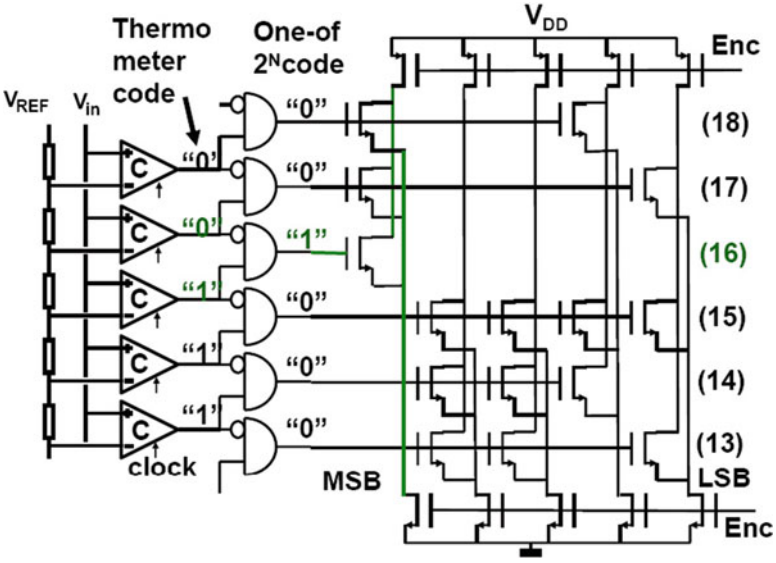


Fig. 8.30 A wired NOR based decoder scheme

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 8.30 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 metastability 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 8.31 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

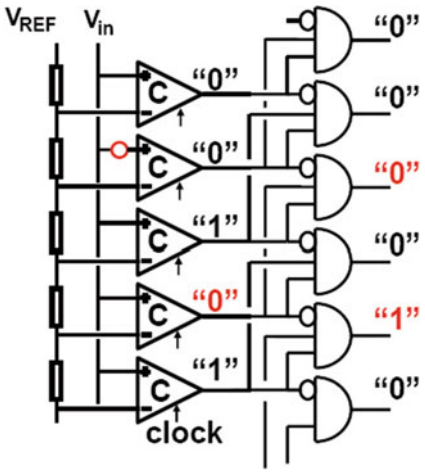


Fig. 8.31 A wired NOR-based decoder scheme with bubble correction

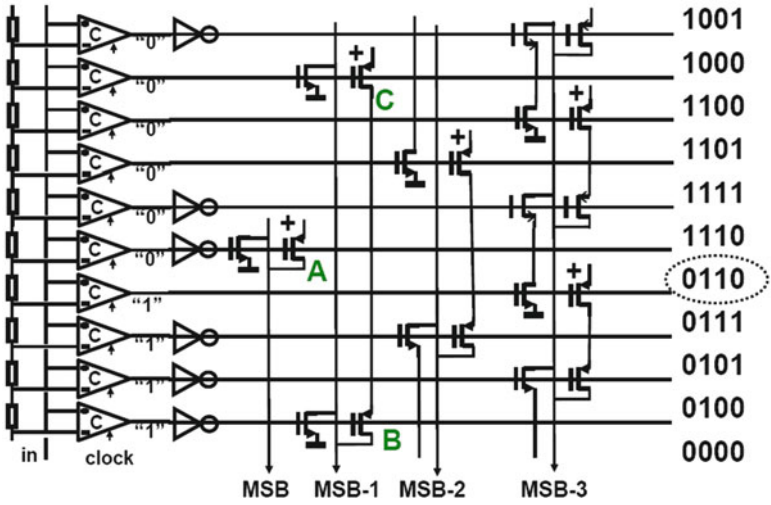


Fig. 8.32 A wired gray-decoder scheme

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 metastability. A meta-stable comparator output in Fig. 8.30 will affect two decode gates and their associated decode lines. If the metastable condition continues, opposing signals may appear in the two decode gates and a major decoding error will happen.

The MSB line in the Gray-decoding scheme of Fig. 8.32 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.

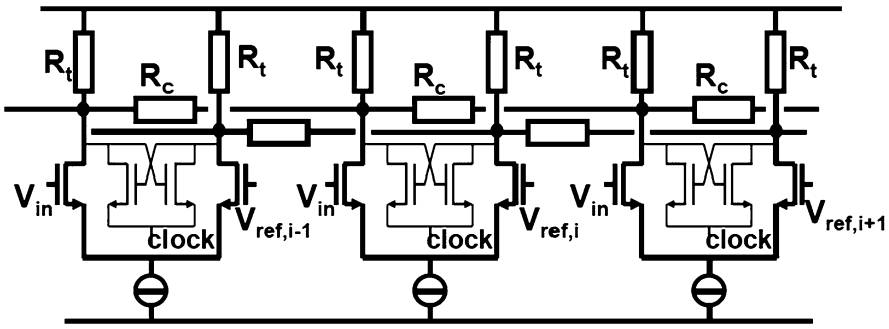


Fig. 8.33 The averaging scheme after [183]

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. 8.32 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.

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 1990s 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 [183] 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 8.33 shows the topology. The input stages of the comparators generate a differential current that forms a differential voltage over the resistors R_t. These currents are combined through coupling resistors R_c.

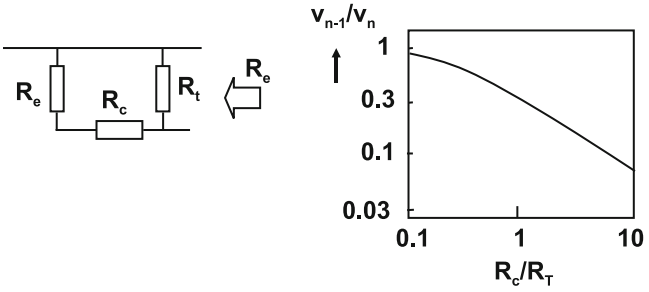


Fig. 8.34 Starting point for the analysis is the equivalent resistor network on the *left*. Equation 8.23 is shown to the *right* [184]

Figure 8.34(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 [183, 184]

$$\frac{R_e}{R_t} = -\frac{1}{2}\alpha + \frac{1}{2}\sqrt{\alpha^2 + 4\alpha}, \tag{8.22}$$

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 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. 8.34(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 [185] used an averaging over five 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 [184] 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.

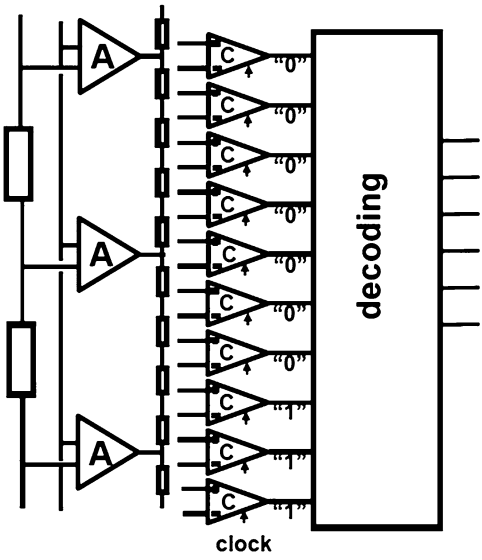


Fig. 8.35 A gain stage is applied to increase the voltage swing on the comparators

A second form of improving the ratio between the signal and random mismatch is shown in Fig. 8.35. 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.

Figure 8.36 shows a full-flash converter placed in a system-on-chip design. Frequently averaging and interpolation are combined into the same design [184, 186].

8.2.5 Frequency-Dependent Mismatch

So far, no frequency dependencies of the input gain stages have been considered. Figure 8.37 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 $C = WLC_{\text{ox}}$ and $\sigma_{Vo} = A_{V_\text{T}}\sqrt{N_T/WL}$. A_{V_T} 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:

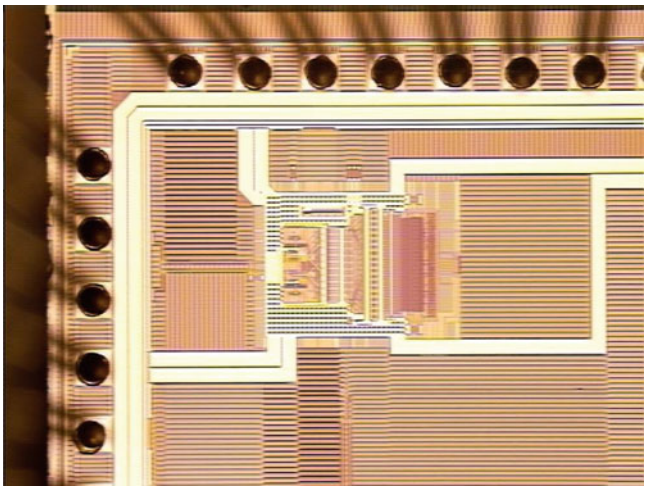


Fig. 8.36 A picture of a 6-bit full-flash ADC converter [184]

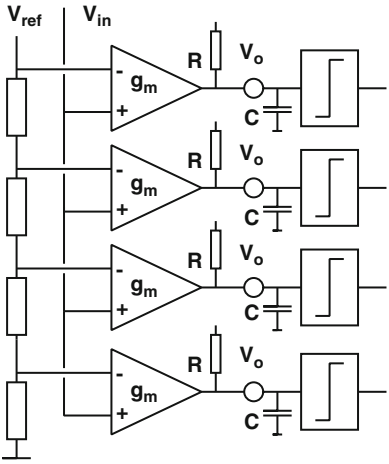


Fig. 8.37 Random offset reduction by means of a gain stage

$$\sigma = \frac{\sigma_{Vo}(1+j\omega RC)}{g_m R} \approx \frac{A_{V_T}\omega C_{ox}\sqrt{N_T WL}}{g_m} \qquad \omega RC > 1. \tag{8.24}$$

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_{V_T}, W, L), and power (g_m, number of comparators). For 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.

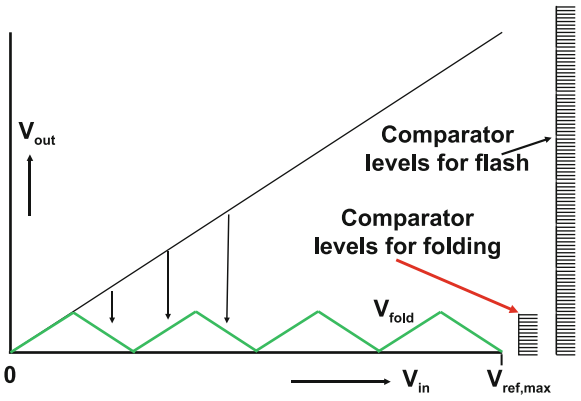


Fig. 8.38 A folding analog-to-digital converter folds the input signal into a smaller signal range

8.2.6 *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. 2.44.
- 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 reduce 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.

8.2.7 *Folding Converter*

A variant of a full-flash converter is the folding analog-to-digital converter [2, 168, 187, 188] in which a preprocessing stage "folds" the input signal. Figure 8.38 shows

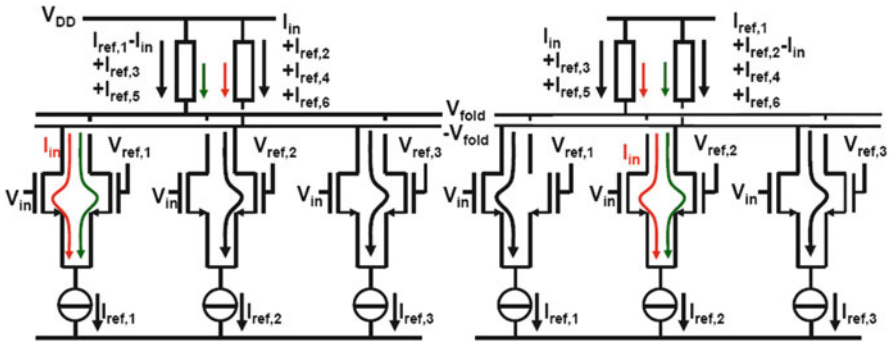


Fig. 8.39 The basic folding circuit: the input pairs are designed to reverse the signal current in the top resistors for each folding section

the general idea: the input signal is folded into (in this example) eight sections. The resulting folded range is applied to the succeeding analog-to-digital converter 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 partitioning reduces the total number of comparators. If an 8-bit flash converter is split in eight 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 developed in bipolar technology [187, 189]. Later on the folding principle was applied in CMOS analog-to-digital conversion [168, 170]. 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 $1/16, 3/16, 5/16, \ldots$ of the full-scale reference voltage. In Fig. 8.39 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 transconductance 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 signal 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. 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.

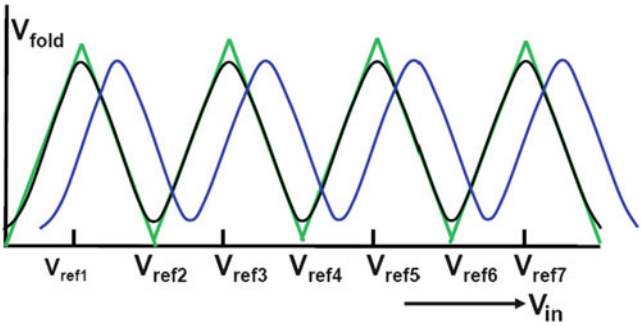


Fig. 8.40 The basic folding circuit suffers from distortions near the switch points. In a practical design two folding circuits with a mutual offset allow an interpolation scheme to determine the switching levels

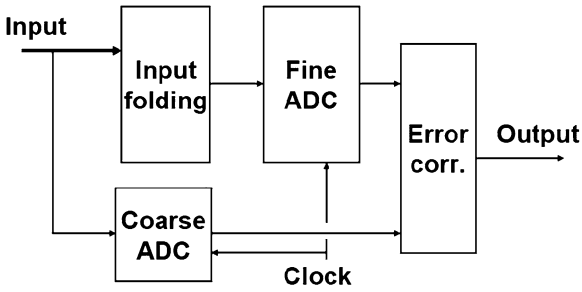


Fig. 8.41 A folding analog-to-digital converter architecture incorporating a sample-and-hold [168]

The crossover point between the stages in the basic concept is a weak point in the architecture of Fig. 8.39. In order to mitigate the problems with the crossover points, mostly a dual folding architecture as in Fig. 8.40 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 crossover 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 crossover of the appropriate combination of (interpolated) folding signals [168, 185]. Proper design of the preprocessing stage where high speed and high yield are combined is the critical issue.

The architecture of Fig. 8.41 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 preprocessing are reduced and that architectures can be used that make multiple use of the input signal. The second 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

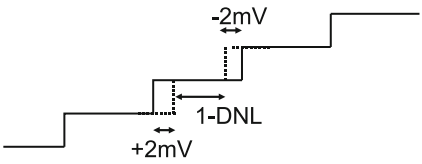


Fig. 8.42 The shift of two trip levels in a worst case situation

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 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 [185] by using an interpolation method at the cost of a lot of power. In bipolar technology a performance level of 10 bits has been reported [190], while in CMOS an 8–9-bit level is state of the art [168, 170, 185]. A form of calibration can further improve the resolution [191].

Example 8.6. In an AD converter with an input range of 1.024 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?

Solution. A DNL error is caused because the trip levels shift due to comparator mismatch (Fig. 8.42). In this case potentially the ith trip level can shift $+2$ mV, while the $i+1$th trip level shifts -2 mV. These two errors together cause a maximum DNL error of 0.5 LSB. Consequently $0.5V_{\mathrm{LSB}} = 2+2$ mV. An LSB size of 8 mV results in a 7-bit converter.

Any converter based on a single comparator can perform better, e.g., a successive approximation converter.

8.2.8 Digital Output Power

The result of an analog-to-digital conversion is a digital word changing its value at the speed of the sample clock. Obviously this interface to the digital processing consumes power as is given by the standard digital power equation. Charging a capacitance C_{load} requires $C_{\mathrm{load}}V_{\mathrm{DD}}^2$ energy from the power supply. Assuming that at every sample pulse half of the N output pins will change state and only half of those will require energy from the power supply yields

$$P_{\mathrm{dig-out}} = N f_s C_{\mathrm{load}} V_{\mathrm{DD}}^2/4.$$

With $f_s = 1$ GHz, $V_{\mathrm{DD}} = 1$ V, $N = 10$ bit, and on-chip capacitive load of $C_{\mathrm{load}} = 1$ pF, the power consumption is 10 mW. Much of this power contains signal-related

components that can cause serious performance degradation if these components influence the analog signals. The digital output power is preferably delivered by a separate power supply.

If an external load must be driven by the digital output, the capacitive load is at least an order of magnitude higher: 10–20 pF. The power needed in the digital output driver climbs to 0.2 W or more. Special precautions must be taken to avoid coupling between these drivers and the rest of the analog-to-digital converter and other analog circuits. Separate power wiring, power bond pads interleaved between the output bits. Also the heat production of the drivers can be of importance. There are various ways to reduce the effect of the digital output section. Examples are: reducing the output load[5] by connecting an extern digital register closely to the analog-to-digital package, low-swing outputs, and small damping resistors in series. More expensive measures implement differential output ports or digitally coded outputs. Low-sample-rate converters can use a parallel-to-series output. All N data bits are shifted out via a single bond pad at a clock rate of $N \times f_s$. In fact, this will not reduce the needed output power, but the much higher frequencies will contain less signal-related energy.

8.3 Subranging 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. Multistep methods allow to achieve higher resolutions. In Fig. 8.43 the signal flow in a two-step converter is shown. The signal is sampled and held at the input. A limited

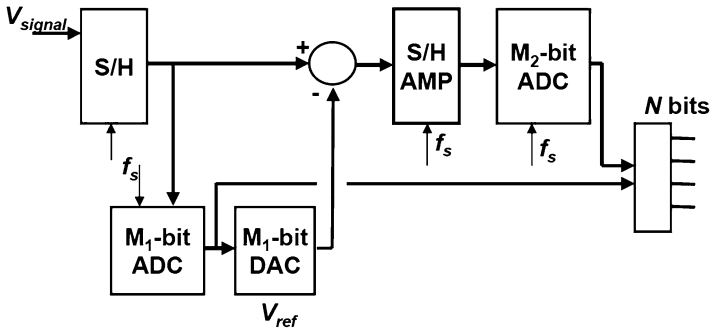


Fig. 8.43 Block diagram of a two-step converter

[5] An often encountered error is a direct connection between the ADC chip and the input port of a laptop. Certainly some spurs related to the internal processing will be visible in the analog-to-digital conversion spectrum.

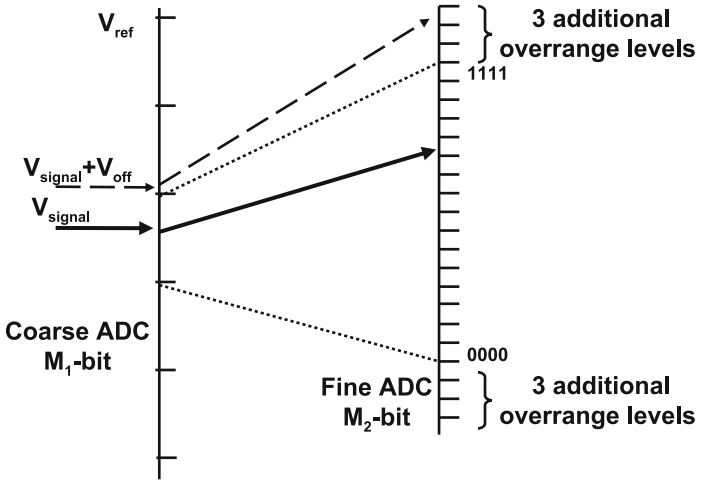


Fig. 8.44 Coarse-fine conversion: on the *left-hand* scale the trip points of the first (*coarse*) converter are indicated. The remaining signal (*bold line*) is amplified and converted by the second (*fine*) converter. The *dashed arrow* indicates a situation where the coarse converter has decided for the wrong range. Over range in the fine converter corrects this decision

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 "subranging 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.

8.3.1 Overrange

A disadvantage of the set-up in Fig. 8.43 with $M_1 + M_2 = N$ is 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. 8.44),

errors from the first converter can be corrected [192]. This "over range" limits the accuracy requirements on the coarse analog-to-digital converter. In some designs the over range 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 pipelined 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 over range 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 [193].

8.3.2 Monkey Switching

The principle of subranging 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. 8.45 an elementary connection topology is depicted. The fine ladder is connected to the selected range of the coarse ladder via two buffers A and B. These buffers can have a gain of 1 or more. If these buffers suffer from offsets the range defined by the coarse ladder will be stretched or shrunk. In Fig. 8.45 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 if the signal increases the sample that is in the adjacent coarse-ladder segment. This sequence causes an INL and a DNL error at the transition points and 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 Fig. 8.46 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."[6] The DNL transitions are strongly reduced. However, the INL

[6]Some similarity exists with the way a monkey climbs a tree.

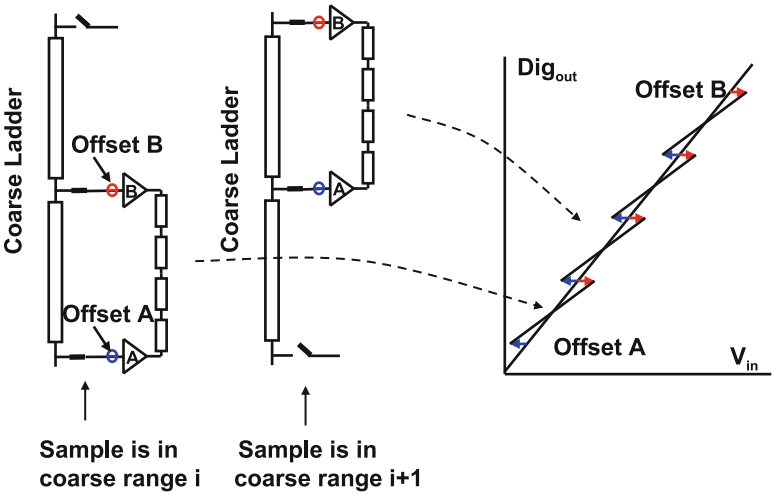


Fig. 8.45 The effect of offset in a subranging scheme. INL and DNL are degraded at the coarse code transitions

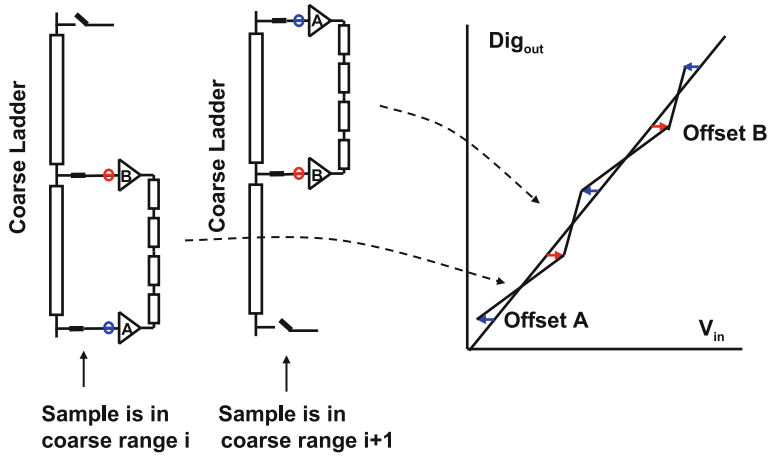


Fig. 8.46 The effect of offset can be reduced by appropriate switching schemes. Only the INL is seriously degraded at the coarse code transitions while the DNL is reduced

errors remain and are visible together with the errors in the fine converter as a repetitive pattern; see Fig. 8.47. For a recent alternative see [194].

The setup of the timing is essential especially if additional time periods are allocated to offset cancellation. Typically the coarse converter is activated after some 10–20% of the hold period. This allows maximum time for the digital-to-analog converter, the subtraction mechanism, and the succeeding amplifier to settle.

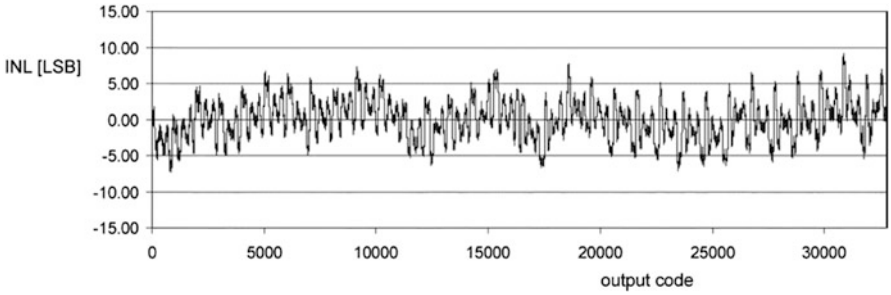


Fig. 8.47 Sub-ranging converters show a repetitive pattern in their transfer curve. This curve is from a 15-bit resolution converter [193]

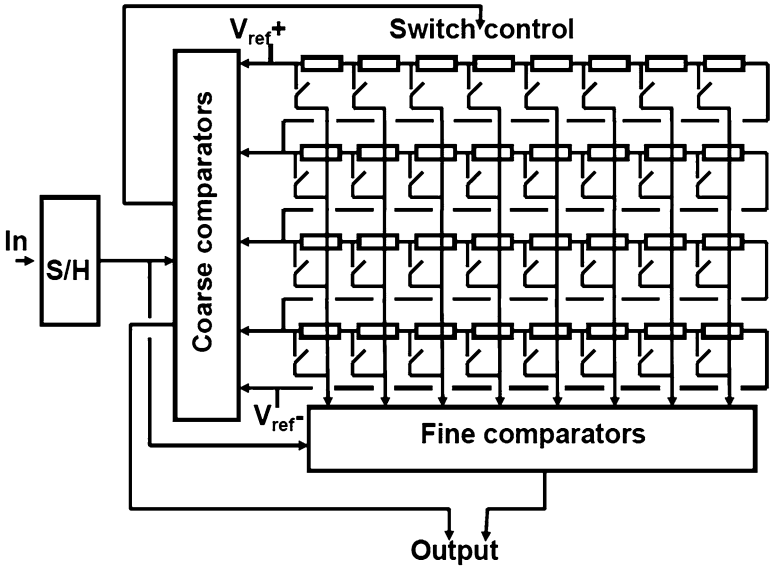


Fig. 8.48 A subranging analog-to-digital converter based on a switched resistor array [160, 175, 178]

An alternative to the standard subrange 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 [160, 175, 178, 179] as shown in Fig. 8.48. The coarse comparators are connected to ladder taps spaced at eight LSB positions apart. The decision of the coarse converter will select a row of switches which is then fed to the fine comparators. Another extension [175] is to use two banks of fine comparators that will alternately digitize the signal, thereby allowing more time for the settling process. In [178] a capacitive interpolator is used to form the intermediate values for the fine conversion.

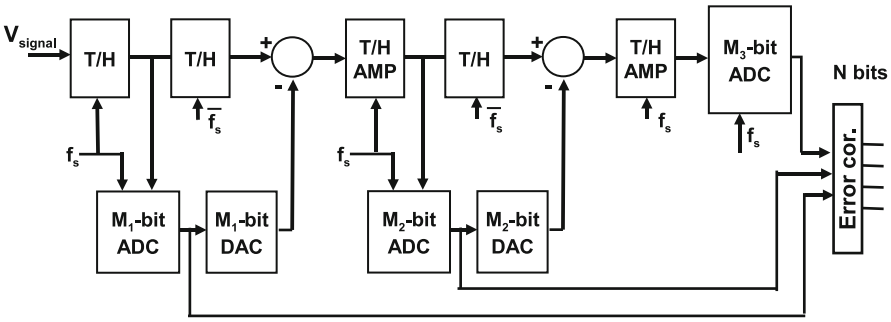


Fig. 8.49 A subranging analog-to-digital converter with three sections. A typical partitioning uses $M_1 = M_2 = 5$ bits and $M_3 = 4$ bits for a 14-bit converter

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 on the ladder under control. In [160] this is realized via a second low-ohmic ladder in parallel to the main ladder similar to Fig. 7.44.

The dominant implementation of stand-alone subrange analog-to-digital converters for industrial applications[7] uses a three-stage approach [195]; see Fig. 8.49. 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 bits and 4–6 bits 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 bits the ENOB ranges from 11.3 to 11.8 bits. Sample rates between 150 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.

Example 8.7. In a 10-bit 1-V subrange ADC the buffers between the 7-bit coarse and the 3-bit fine converter have 3 and -2 mV offsets. Sketch the resulting INL in case the buffers are switched directly or in a "monkey" way. What is in both cases the largest DNL?

Solution. Figure 8.50 shows a part of the transfer curve in case the low-side buffer has a positive offset and the high-side buffer a negative offset. A plateau appears in

[7] 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.

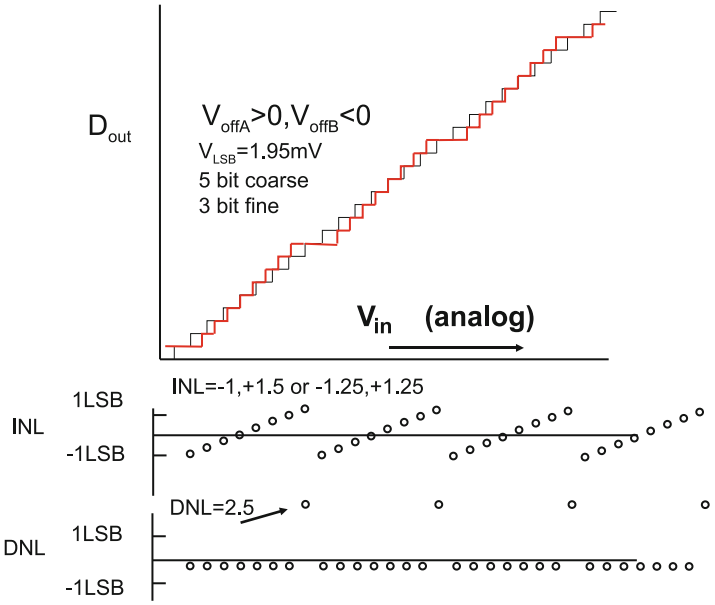


Fig. 8.50 Transfer curve, INL, and DNL for a 3-bit fine section in a two-step analog-to-digital converter, with offsets in the buffers

the overall transfer curve with a DNL error corresponding to the offsets divided by V_{LSB}. If the offset appear the other way around, the transfer will look like Fig. 8.51. In this example two codes will be missing.

Figure 8.52 shows the effect of the monkey-switching scheme. The maximum deviation of the INL has not changed, as it is determined by the offsets, but the local deviations at the ends of the fine range have largely disappeared. The remaining transfer curve consists of 8 codes that are together $|V_{\mathrm{off},A}| + |V_{\mathrm{off},B}|$ too large, followed by 8 codes that are the same amount too small. The DNL error per code is 5 mV/8; related to a 10-bit LSB this would result in DNL = 0.6 LSB.

8.4 1-Bit Pipeline Analog-to-Digital Converters

Subranging avoids the exponential hardware and power increase of flash conversion. Using more bits per subrange stage reduces the number of stages needed and the associated track-and-holds. If the first stage is implemented with M_1 =3–5 bits, the intermediate gain can be 2^{M_1}, thereby reducing the influence of errors in the succeeding stages. When a design is pushed to the limits of the technology, the maximum unity-gain bandwidth is determined by the available power. More intermediate gain will then correspond to slower settling. Reducing the subrange to the extreme results in subranges of one single bit. Therefore a popular variant

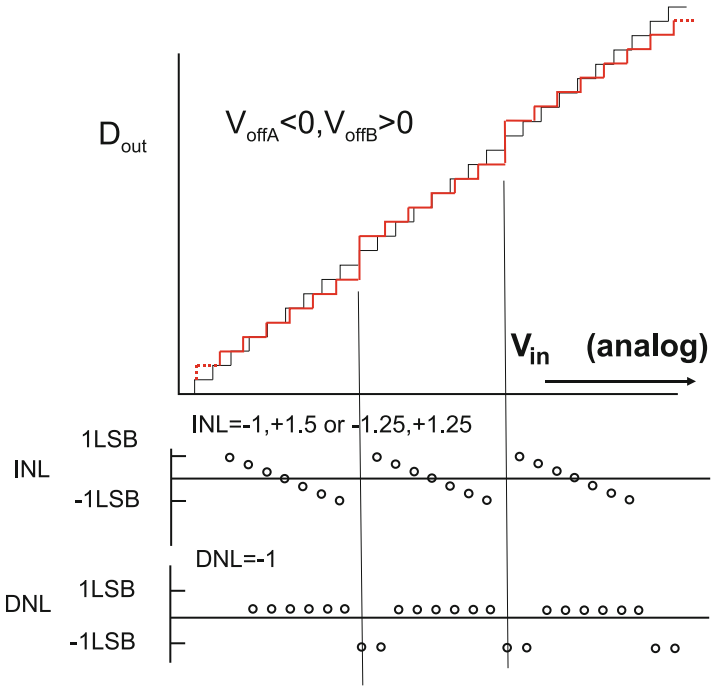


Fig. 8.51 Transfer curve, INL, and DNL for a 3-bit fine section in a two-step analog-to-digital converter, with opposite offsets in the buffers

of the subrange analog-to-digital converter in CMOS technology is the "pipeline" converter (Fig. 8.53) [192, 196]. This converter consists of a pipeline of N more-or-less identical 1-bit stages. For each bit of resolution there is one stage. Each stage comprises a track-and-hold, a comparator connected to a 1-bit digital-to-analog converter, a subtraction mechanism, and a multiplication circuit.

The important advantages are fast settling due to a low interstage gain (2 or less) and a 1-bit digital-to-analog converter. With only two output values, this digital-to-analog converter is by definition perfectly linear.

The operation of the pipeline converter can be viewed from different angles. A pipeline converter can be seen as the extreme form of subranging with 1 bit per subrange.

From another point of view, the 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 ith stage,

$$V_{\text{out},i} = 2 \times V_{\text{in},i} - D_i V_{\text{ref}}, \tag{8.25}$$

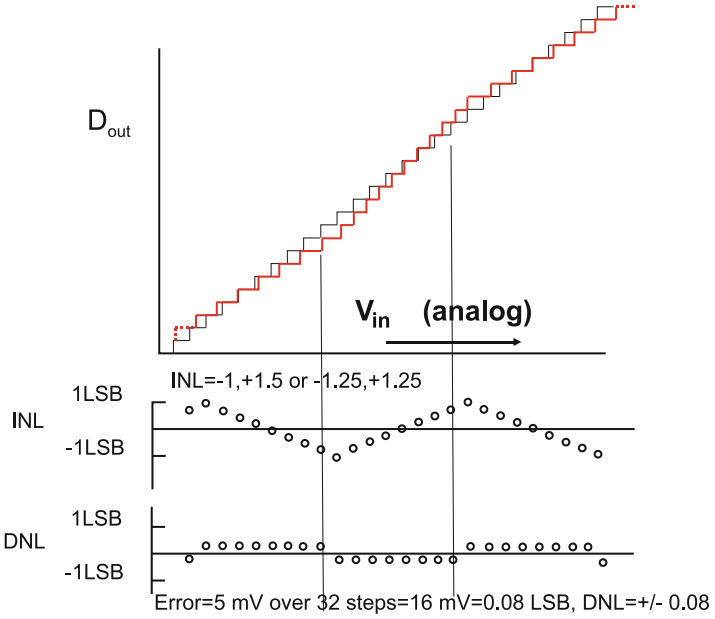


Fig. 8.52 Transfer curve, INL, and DNL for a 3-bit fine section in a two-step analog-to-digital converter. Monkey switching reduces the DNL errors

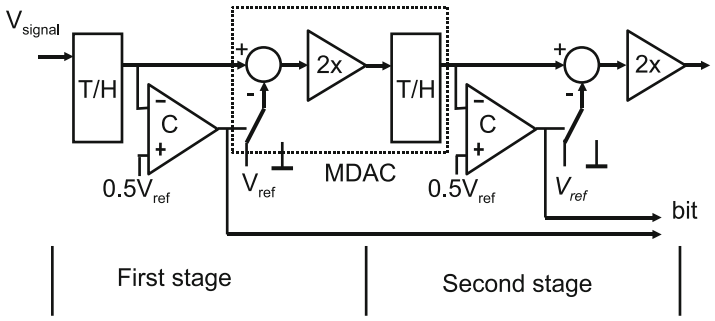


Fig. 8.53 Basic pipe-line analog-to-digital converter

where $D_i = 0, 1$ is the decision bit. This is essentially the kernel of the successive approximation algorithm from Fig. 8.70. 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. 8.54. 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.

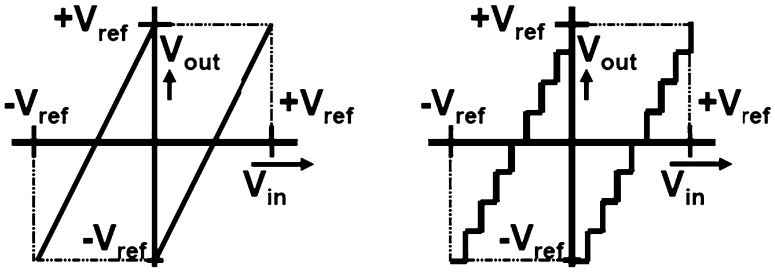


Fig. 8.54 Transfer characteristics for an MDAC in a pipe-line analog-to-digital converter. *Left*: the first stage, *right*: indicating the levels where a succeeding 3-bit flash converter would quantize. The MDAC and the flash form a four-bit converter

The choice to operate the sections at 1-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. Random errors, 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 stages add up as

$$v_{\text{in, random}} = \sqrt{v_{1,\text{random}}^2 + 2^{-2} v_{2,\text{random}}^2 + 2^{-4} v_{3,\text{random}}^2 + \cdots}. \tag{8.26}$$

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 capacitor 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 [197].

With a conversion of 1 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.

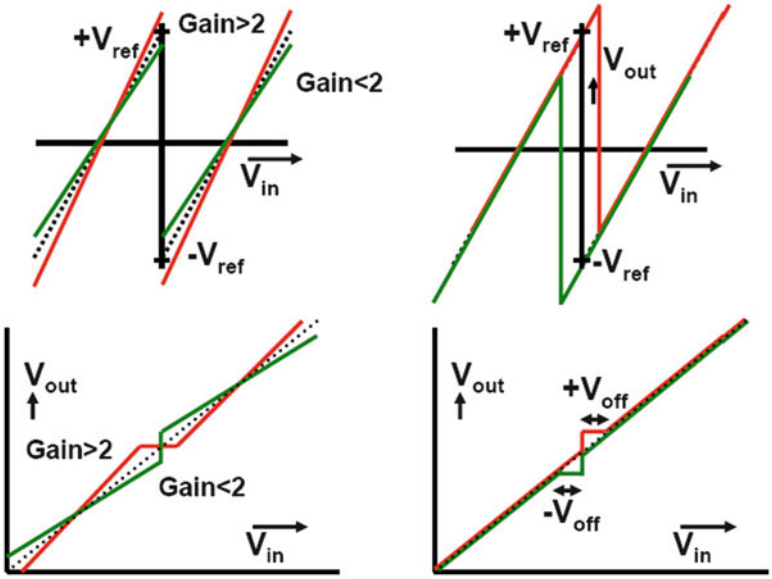


Fig. 8.55 Two errors in a pipeline stage. *Left*: the gain is not equal to 2 and *right*: the comparator suffers from a positive or negative offset. A converter with these errors in the first stage has an overall transfer characteristics as shown in the lower plots

8.4.1 Error Sources in Pipeline Converters

Two important errors in this architecture are gain errors and comparator offset (Fig. 8.55). 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, 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. 8.56. 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

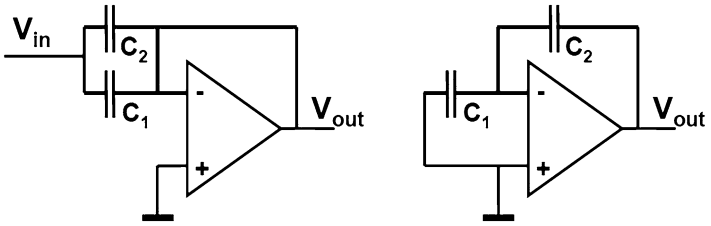


Fig. 8.56 The multiply-by-two operation in a pipeline converter is implemented by sampling the signal on two equal capacitors (*left*). In the multiply phase (*right*), all charge is transferred to C_2 [149]

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:

$$\varepsilon_{\text{DCgain}} = \frac{1}{A_0 C_2/(C_1+C_2)} \ll 2^{-N+i} \quad \rightarrow \quad A_0 > 2^{N-(i-1)} \tag{8.28}$$

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. Alternative ways to reduce the gain are based on reduction of the output swing in Fig. 4.24 [102] and multiple gain steps [198].

Errors in the capacitor ratio can be the result of technological problems or deviations caused by the layout 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

$$\varepsilon_{\text{cap}} = \frac{C_1 - C_2}{C_1 + C_2} \ll 2^{-N+i}. \tag{8.29}$$

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:

$$\varepsilon_{\text{settle}} = \mathrm{e}^{-2\pi f_{\text{UGBW}} T_{\text{settle}}} \ll 2^{-N+1} \quad \rightarrow \quad f_{\text{UGBW}} > \frac{f_s(N-i)}{2\pi\alpha} \tag{8.30}$$

$\alpha < 0.5$ 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 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

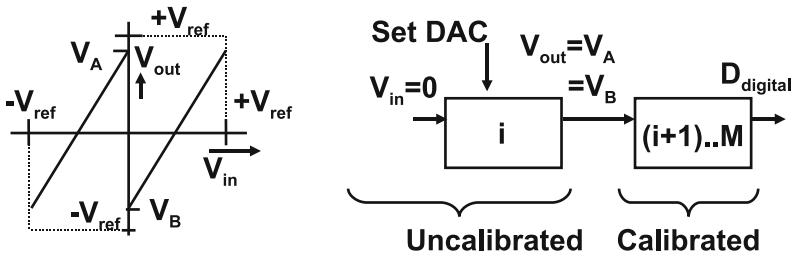


Fig. 8.57 The gain in the stage is deliberately made smaller than 2. The converter is calibrated starting at its back end and working toward the front-end [200]

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 $0.5V_{\mathrm{LSB}}$. 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 [199] 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 [100] 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. 8.56 and, in 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 Reduced Radix Converters with Digital Calibration

In order to circumvent the accuracy problems Karanicolas [200] 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. 8.57. With less than 1-bit resolution per stage, more stages M are needed than the resolution N. This is in fact a conversion with a radix or base less than 2.

In order to reconstruct the input from a given output value the following equations are used:

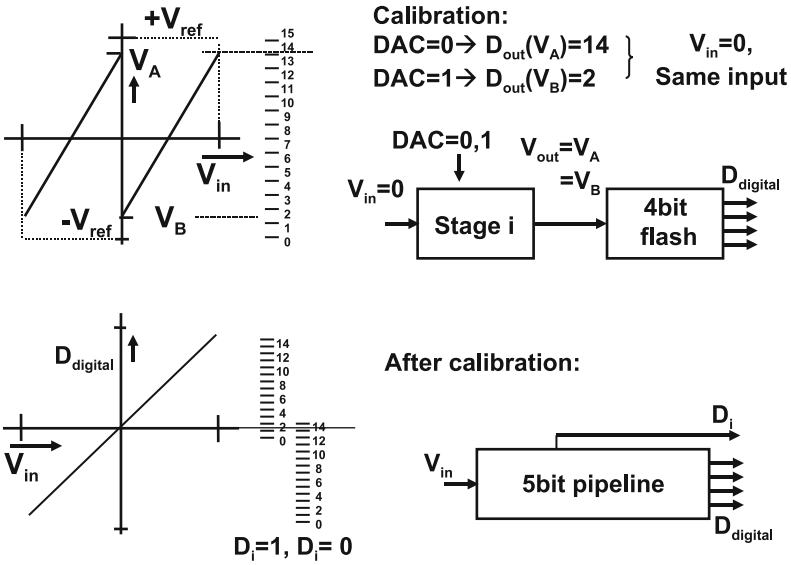


Fig. 8.58 Example of pipeline calibration

$$V_{\text{in}} < 0, \rightarrow V_{\text{in}} = V_{\text{out}} - V_A$$
$$V_{\text{in}} > 0, \rightarrow V_{\text{in}} = V_{\text{out}} - V_B. \tag{8.31}$$

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 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 $+V_{\text{ref}}$. The output will be equal to the value V_A. Similarly V_B is obtained with the digital-to-analog converter set to $-V_{\text{ref}}$. 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 8.58 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 $V_{\text{in}} = 0$). 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. 8.53 requires full and accurate settling in all stages. Digital correction is possible when the gain factor is reduced.

The 1.5-bit pipeline converter allows a more extensive error correction as in each stage two comparators are used; see Fig. 8.59. 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. 8.60 [98, 201–205]. 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 bits:

$$V_{\text{out}} = 2V_{\text{in}} + \begin{pmatrix} +V_{\text{ref}} \\ 0 \\ -V_{\text{ref}} \end{pmatrix}. \tag{8.32}$$

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.

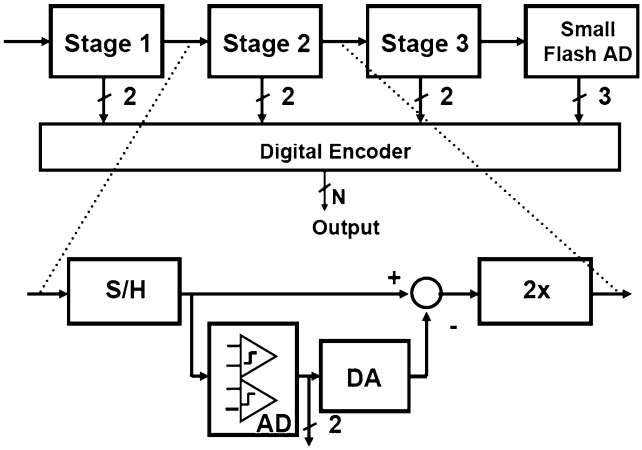


Fig. 8.59 The architecture of a 1.5-bit pipeline analog-to-digital converter is derived from the coarse-fine architecture

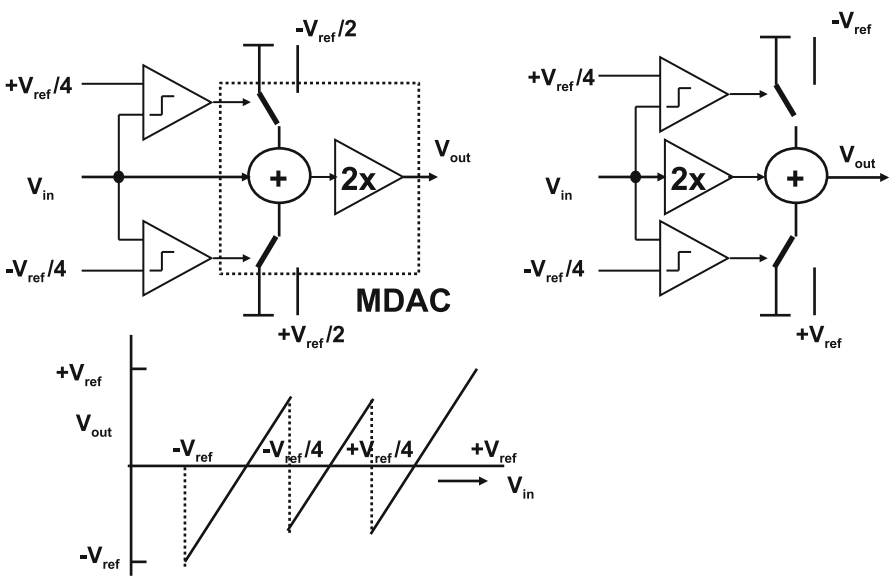


Fig. 8.60 Basic operation of a 1.5-bit pipe-line analog-to-digital converter. The section in the *dotted box* is called a multiplying digital-to-analog converter (MDAC)

The typical transfer curve from Fig. 8.60 shows that negative signals are shifted upwards and positive signals are shifted downwards in the plot with respect to the values around zero.

The section consisting of the subtraction of the reference and the multiplication is called a multiplying digital-to-analog converter (MDAC).

8.5.1 Design of an MDAC 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.

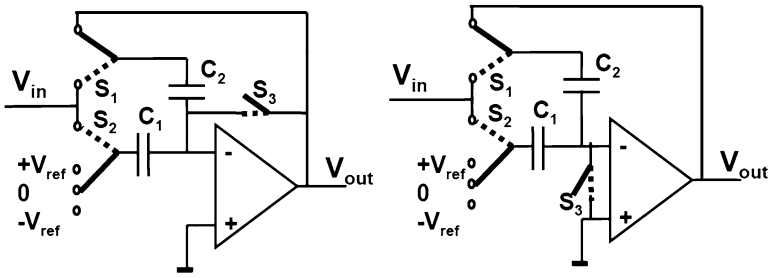


Fig. 8.61 A track-and-hold with multiply-and-subtract stage (MDAC) used in 1.5-bit pipeline converters. Two alternative schemes are shown for creating a ground node in pipeline converters. In the hold phase the switches in the drawn positions; for the track phase the switches are in the dotted positions

A popular implementation is shown in Fig. 8.61(left). This MDAC configuration allows a large-signal swing without any repercussion for the input of 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. 8.61(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 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 required capacitor accuracy is given by Eq. 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. High-resolution converters require the track-and-hold capacitor to fulfill the kT/C noise limit. The input-referred noise of the opamp must be added and during the switching another kT/C addition occurs. Depending on the required specifications and the design of the opamp a design guideline is to use an excess factor of $E_{\text{noise}} \approx 3$:

$$\frac{E_{\text{noise}} kT}{C_1} < \frac{V_{\text{LSB}}^2}{12} = \frac{V_{\text{pp}}^2 2^{-2N}}{12}, \tag{8.34}$$

where V_{pp} represents the peak–peak value of the input signal.

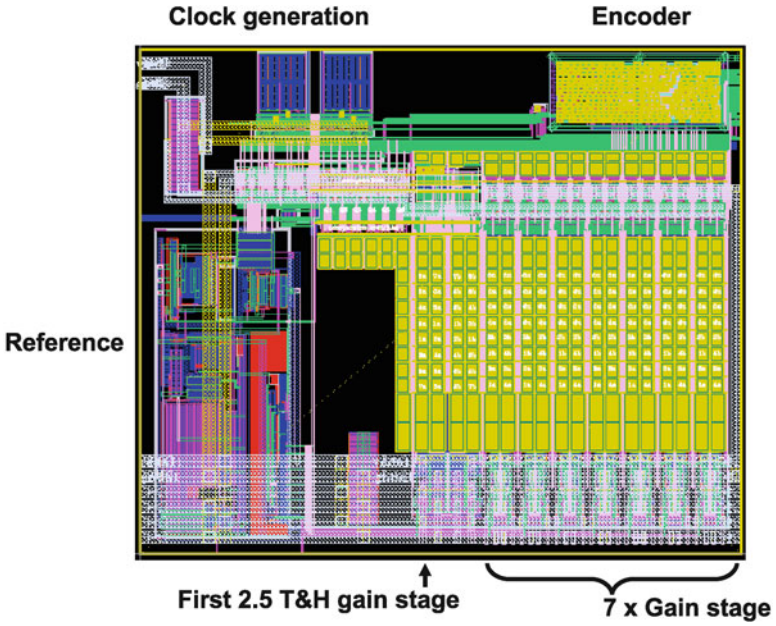


Fig. 8.62 A 10-bit pipeline layout. The first stage contains 2.5 bits and the seven remaining stages 1.5 bits. The depicted area in 90 nm CMOS is $500 \times 600\,\mu m^2$ (courtesy: Govert Geelen NXP [206])

The capacitor choice is a determining factor in the design. The total noise accumulates as in Eq. 8.26 and Fig. 4.21. However, the capacitor value also determines the current setting of the buffer and thereby most of the analog-to-digital converters power.

When all stages resolve 1.5 bits, the contribution of the second and third stages is still significant. Often the same capacitor value for these stages is used and scaling is applied after the third stage. Lower capacitive load means lower currents, and power saving is obvious [197]. The alternative is to use more levels in the first stage, e.g., 6 [206] or 7 [99].

Figure 8.62 shows the layout of a 10-bit pipeline converter. Most of the area is used by the capacitors.

With a large swing of the input signal ($2\ V_{pp}$) 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 bond pad that is connected to the input terminals. Its protection measures must be examined as too much parasitic capacitance will affect the performance. Special RF bond pads are an alternative.

The amplification stage has to fulfill similar demands as in the 1-bit pipeline converter. The topology of the opamp depends on the specifications. Low accuracy and high speed will push toward simple transconductance stages, while high accuracy may involve (folded) cascode amplifiers. The DC gain is set by the overall resolution

Table 8.4 Comparison of 8- and 12-bit 1.5-bit pipeline stages

Parameter	Equation	$N = 8$ bits	$N = 12$ bits
Sample rate	f_s	100 Ms/s	20 Ms/s
LSB size	$V_{\text{LSB}} = 2^{-N} V_{\text{pp}}$	3.9 mV	0.24 mV
Sampling cap	8.34: $E_{\text{noise}} kT/C_1 < V_{\text{LSB}}^2/12$	9.5 fF	2.4 pF
Capacitor matching	8.29: $(C_1 - C_2)/(C_1 + C_2) \ll 2^{-N}$	10^{-3}	10^{-4}
Opamp gain	8.28: $A_0 > 2^N$	256	4096
UGBW ($\alpha = 0.5$)	8.30: $f_{\text{UGBW}} > f_s N/2\pi\alpha$	270 MHz	80 MHz
Slew current	8.35: $I_{\text{slew}} > C_{\text{load}}(\text{d}V/\text{d}t)$	4 µA	400 µA

Inspired by ISSCC workshops

in Eq. 8.28 and the settling in Eq. 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 Eq. 8.30. Moreover the slew-rate requirements must be met:

$$I_{\text{bias}} > I_{\text{slew}} = C_{\text{load}} \frac{\text{d}V}{\text{d}t}, \tag{8.35}$$

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, and the next stages can be designed with relaxed specifications, following the stage-scaling principle.

In this example of Fig. 8.61(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. 8.61(right); see, e.g., [206]. The real ground node for the sampling process removes the input-referred opamp noise. This configuration trades off the offset cancellation of the scheme in Fig. 8.61(left) for reduced noise performance in Fig. 8.61(right).

Example 8.8. Determine the parameters for the first section of a pipeline stage for 8-bit resolution at 100 Ms/s and 12-bit resolution at 20 Ms/s, with a 1 V input range.

Solution. Table 8.4 lists the main parameters. Of course fine-tuning with a simulation is needed to incorporate the parasitic effects.

8.5.2 Redundancy

Figure 8.63 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

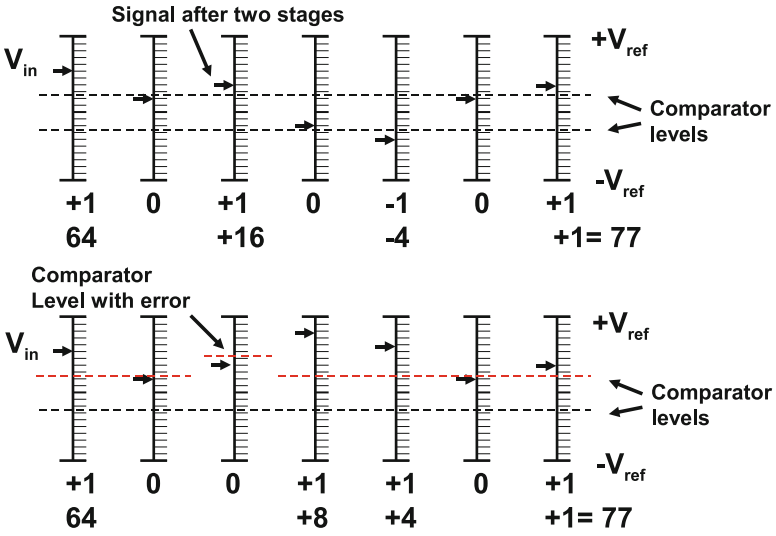


Fig. 8.63 1.5-bit pipeline analog-to-digital converter uses the redundancy of two decisions per stage. The *upper sequence* shows the ideal behavior. In the *lower sequence* one comparator level is shifted, e.g., due to offset. The redundancy corrects the mistaken decisions

$V_{\text{in}} = 0.6V_{\text{ref}}$ clearly exceeds the upper comparator threshold at $0.25V_{\text{ref}}$. Therefore the first coefficient is determined to be $a_{N-1} = 1$. The residue is formed as $2 \times V_{\text{in}} - V_{\text{ref}} = 0.2V_{\text{ref}}$. 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 $77/128 = 0.602V_{\text{ref}}$.

In the lower trace of Fig. 8.63 the comparator level of the third section is moved to $0.45V_{\text{ref}}$. It is easy to see how the redundancy scheme corrects for this error. See also the remarks for the RSD converter in Fig. 8.79. The redundancy eliminates the need for accurate comparators but leaves full accuracy requirements on the switched-capacitor processing stages.

Figure 8.64 shows the layout of a 16-bit pipeline converter.

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 10–14 bits of resolution and 50–500 Ms/s sampling rates. These high-performance converters aim at the same markets as high-end subranging converters. Similar to the high-performance subranging 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 to the

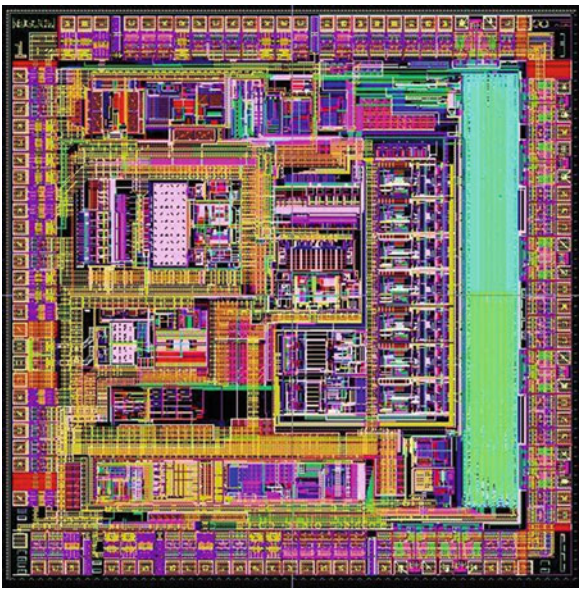


Fig. 8.64 Lay-out of a 16-bit 125 Ms/s pipeline analog-to-digital converter ENOB = 11.6 (courtesy Ph. Gandy, NXP Caen)

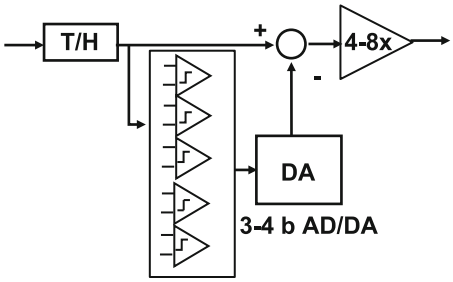


Fig. 8.65 Advanced pipeline concepts use a multi-bit input stage

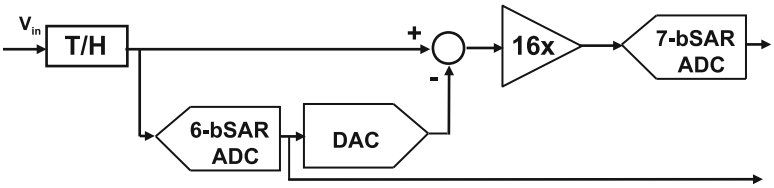


Fig. 8.66 A 6-bit successive approximation analog-to-digital converter strongly reduces the input range [210]

noise and distortion budget, most of these converters use input stages with 2.5–4-bit resolution (Fig. 8.65) [207–209]. The analysis in [202] indicates that this is an optimum range.

A radical variant is proposed in Fig. 8.66 [210]. The input quantizer is a 6-bit successive approximation analog-to-digital converter that reduces the input

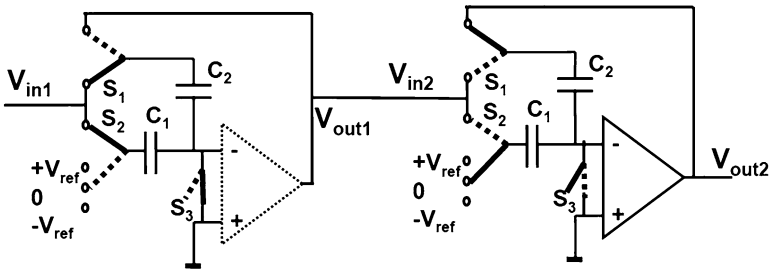


Fig. 8.67 The opamp is shared between two stages in a 1.5-bit pipeline converter. The *left stage* is sampling, and the *right stage* is amplifying. During the next half phase the opamp is moved to the first stage which amplifies, while the second stage samples without an opamp. The switches for reconnecting the opamp have been left out for clarity

range significantly. The linearity requirements of the summation and multiply operation are easier met at the cost of the loss of timing for the successive approximation operation. Compare [211].

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. 8.61(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. 8.67. The opamp now serves two sections of the pipeline converter and reduces the required number of opamps by a factor two [201, 212]. 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 [213] propose to remove the dedicated front-end track-and-hold function. The pipeline stage as in Fig. 8.61 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 [204, 205] 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 in [214, 215]. A comparator switches on and off two current sources in

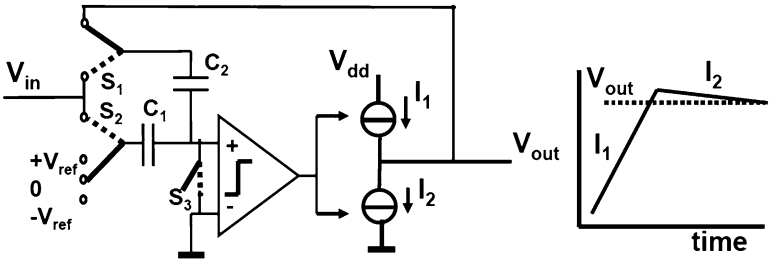


Fig. 8.68 The opamp is removed and replaced by a comparator and two current sources

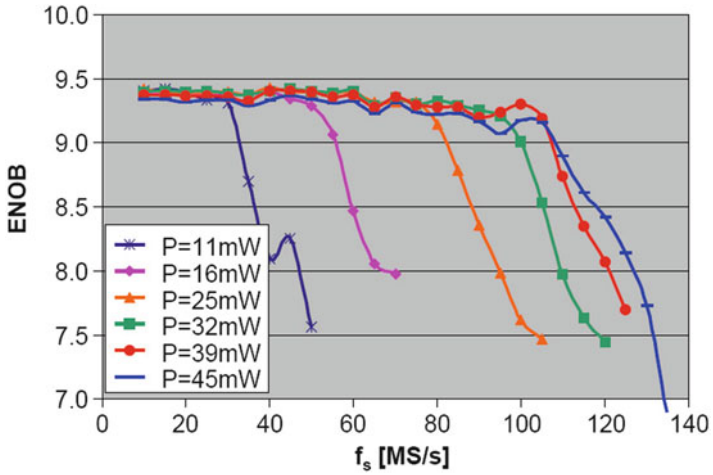


Fig. 8.69 Measured effective number of bits versus sample rate at different power levels. The signal frequency is each time just below $f_s/2$ (courtesy: Govert Geelen NXP [206])

Fig. 8.68. 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.

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. 8.69 [206]. 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 [216].

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 8.70 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 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 8.71 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 output value approximates the input value. Another implementation reduces the input value by subsequent subtraction.

The approximation algorithm in Figs. 8.71 and 8.72(left) requires that the comparator operates over the entire input range to full specification. The reverse and complementary forms Fig. 8.72(middle and left) either decrease the signal value to a zero level or complement to reach the full reference. The comparator has a much

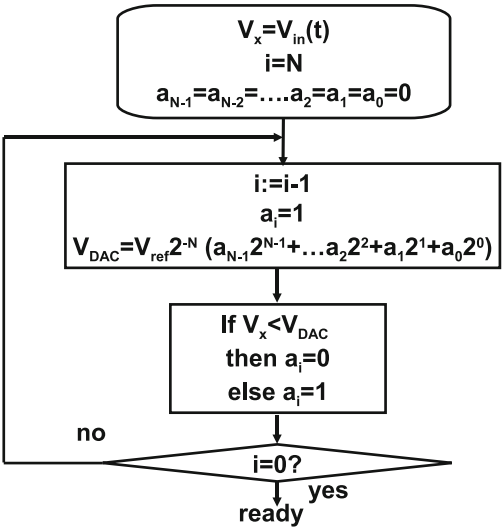


Fig. 8.70 A flow diagram for successive approximation

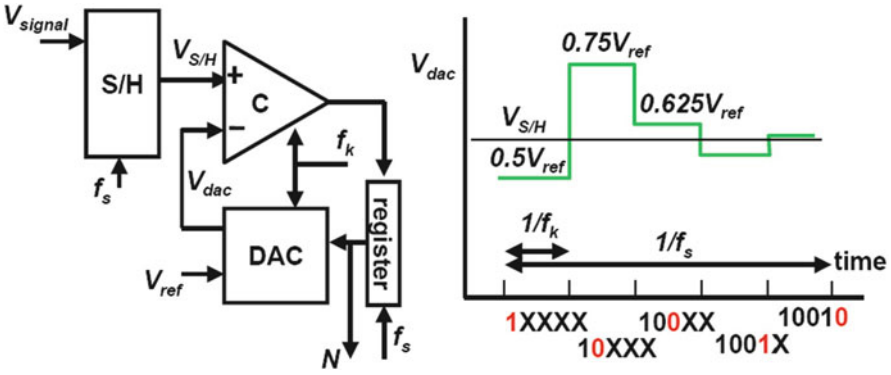


Fig. 8.71 A successive approximation analog-to-digital converter with approximation sequence

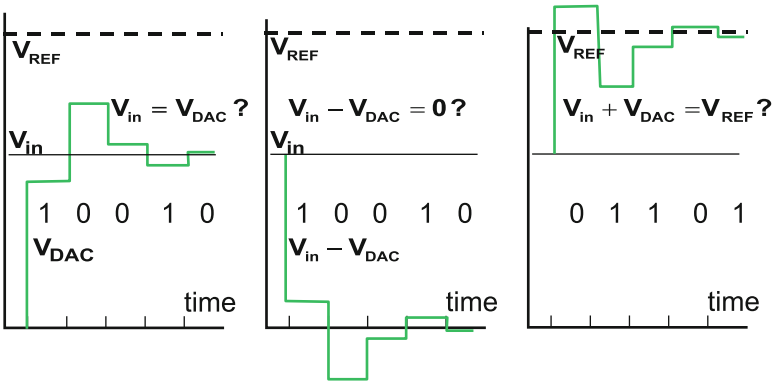


Fig. 8.72 Three forms of successive approximation analog-to-digital converter algorithms: normal, reverse, and complementary

easier task. As can be seen from the example, the processed signal can reach beyond the reference range. In low-power supply circuits this may lead to leakage through forward-biased pn-junctions.

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 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

$$t_{\mathrm{settle}} > \tau_{\mathrm{DAC}} \ln(2^{N+1}). \tag{8.36}$$

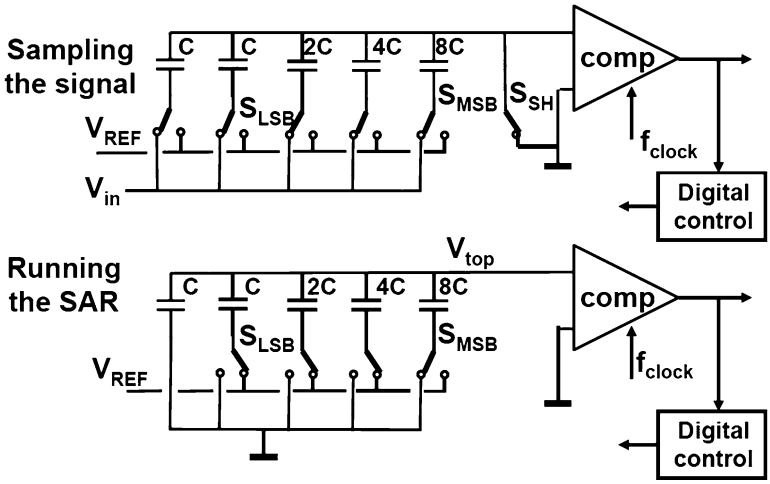


Fig. 8.73 An early implementation of a successive approximation analog-to-digital converter is based on capacitor switching, after [217]

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 microcontroller, 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 [217, 218]. The principle is shown in Fig. 8.73 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

$$V_{\text{top}} = -V_{\text{in}} + \frac{8C}{C+C+2C+4C+8C} V_{\text{ref}}. \tag{8.37}$$

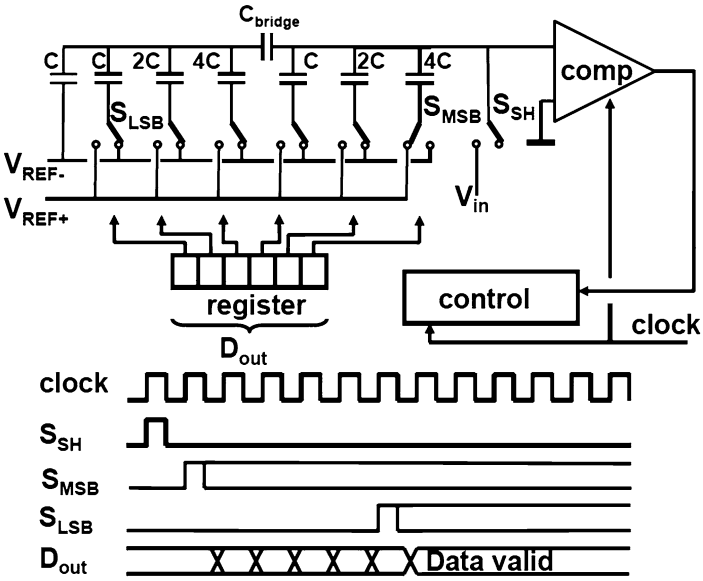


Fig. 8.74 A successive approximation analog-to-digital converter based on capacitor switching [219, 220]

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 8.74 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. 7.25. 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

$$\frac{C_{\text{bridge}} 2^k C}{C_{\text{bridge}} + 2^k C} = C, \tag{8.38}$$

where k is the resolution of the left-hand side array. From this equation the bridge capacitor is found as

$$C_{\text{bridge}} = \frac{2^k}{2^k - 1} C. \tag{8.39}$$

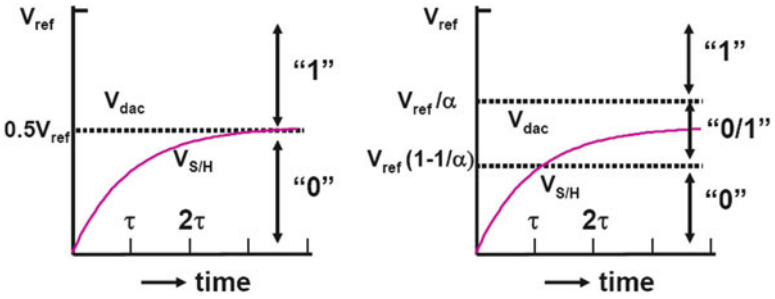


Fig. 8.75 In a base-2 successive approximation the settling of the input signal must reach the final accuracy level before the next step is taken. In non-binary search the base is smaller than 2 and an intermediate region of signal levels exists, where initial decision can be corrected

The bridging capacitor can be rounded to unity if k is large [219], 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 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 [219–221].

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. 8.75. In [222] 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. 8.75) 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). An extension to this nonbinary work is, e.g., found in [223].

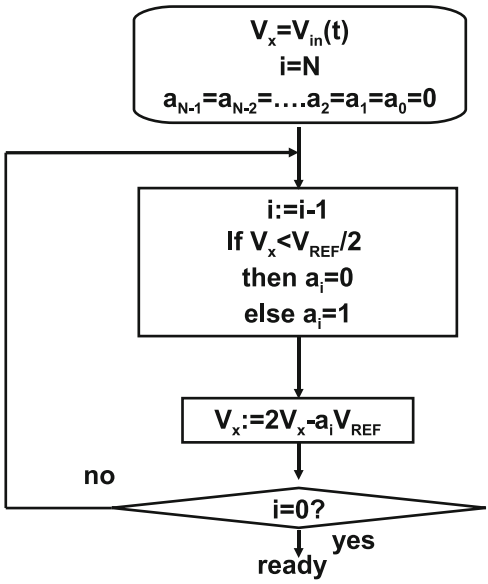


Fig. 8.76 A flow diagram for a cyclic converter

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 [196, 224].

A flow diagram of a basic algorithm is shown in Fig. 8.76. The value V_x is set to the input value and compared to half of the reference. If V_x exceeds half of the 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 elementary successive approximation algorithm. Now errors in the smaller bits count less. Obviously the result of N executions of this flow diagram is

$$V_x = 2^N V_{\text{in}}(t) - V_{\text{ref}}\left(a_{N-1}2^{N-1} + \cdots + a_1 2^1 + a_0 2^0\right). \tag{8.40}$$

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 $(2+\varepsilon)$ the difference between the value at the MSB transition and the (MSB-1LSB) transition (the DNL at that code) equals

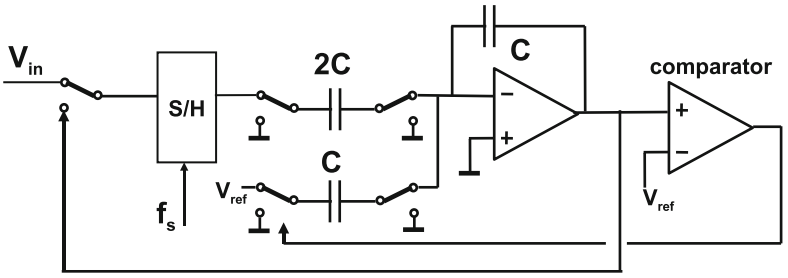


Fig. 8.77 An example of a cyclic analog-to-digital converter, after [199]

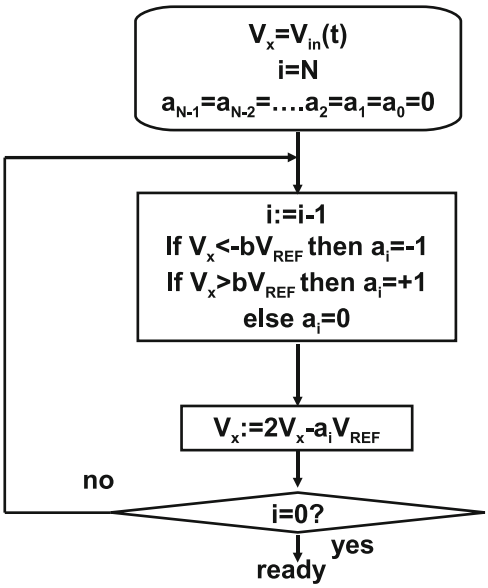
$$D = (2+\varepsilon)^{N-1} - [(2+\varepsilon)^{N-2} + (2+\varepsilon)^{N-3} + \cdots + (2+\varepsilon)^{1} + (2+\varepsilon)^{0}]$$
$$= (2+\varepsilon)^{N-1} - \frac{1-(2+\varepsilon)^{N-1}}{1-(2+\varepsilon)} = \frac{1-\varepsilon(2+\varepsilon)^{N-1}}{1-\varepsilon} \approx 1-\varepsilon 2^{N-1}[\text{ in LSB}]. \tag{8.41}$$

The error in the multiplication factor is itself multiplied by the term 2^{N-1}. In order to keep the DNL sufficiently low, $\varepsilon < 2^{-N}$.

Figure 8.77 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 opamps and comparators 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 Eq. 8.28. The implicit mismatch of the capacitor structure may jeopardize accuracy. However, a careful layout, where the capacitor "2C" is built from two parallel capacitors "C" and properly surrounded by dummy structures, will reduce this error source to the 10–12-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 Eq. 4.5. Differential operation or special switching sequences help to reduce this effect [199].

Several accuracy issues can be removed if some redundancy is built in. The flow diagram in Fig. 8.78 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. 8.79 the algorithm converts a signal $V_{\text{in}} = 0.6V_{\text{ref}}$. The first decision



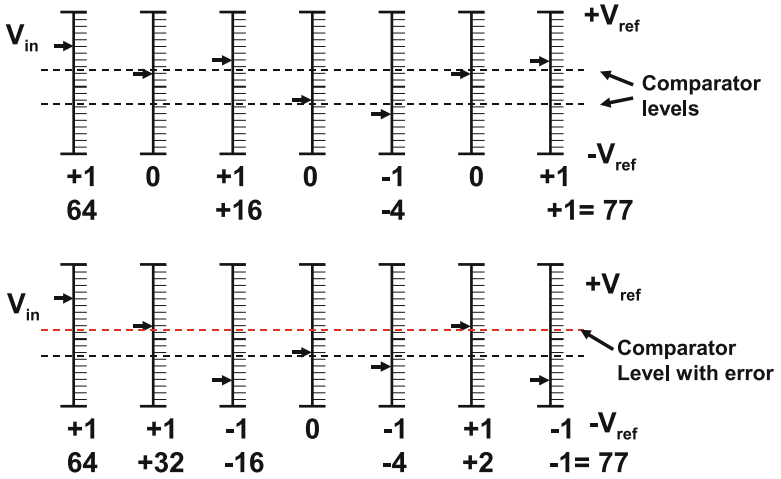


Fig. 8.78 The redundant signed digit algorithm [225]

Fig. 8.79 *Upper sequence*: an input voltage of $0.6V_{\text{ref}}$ is converted in seven cycles with comparator levels at $\pm 0.25V_{\text{ref}}$. *Lower sequence*: the positive comparator level is now at $0.15V_{\text{ref}}$, 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

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 $2 \times 0.2V_{\text{ref}}$. The process repeats until after seven cycles a value of 77 is found: $77/128 = 0.602$.

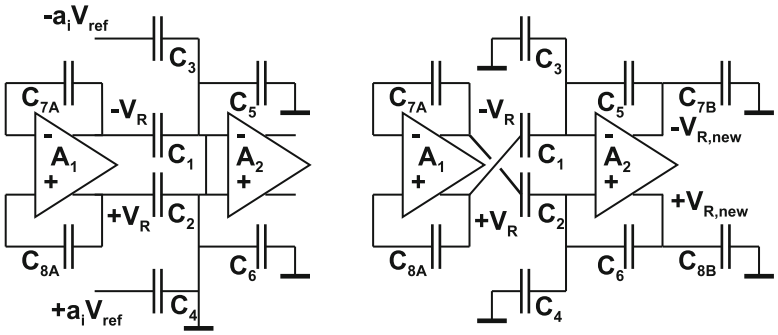


Fig. 8.80 The redundant signed digit (RSD) converter [225]

In the lower part of the figure, the upper comparator level is reduced to $0.15V_{\text{ref}}$, e.g., due to offset. Still the algorithm converges to the same overall result. The RSD algorithm[8] creates in this way robustness for comparator inaccuracies. Figure 8.80 shows an implementation of the RSD principle [225]. 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 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., [226].

Example 8.9. Compare the 1-bit pipeline converter and the successive approximation converter.

Solution. Both converters use a logarithmic approximation of the signal: first the MSB is decided, and based on the outcome a residue is evaluated. Both principles use a track-and-hold circuit to store the signal. Most bandwidth related specifications will therefore be comparable.

Where the standard successive approximation converter is executing the algorithm in time, the pipeline converter uses additional hardware stages. The pipeline converter can reach a high throughput, because the intermediate results are shifted from one hardware section to the next. The successive approximation converter can achieve a similar throughput rate, when N parallel converters are used.

Mismatch affecting comparators is reduced in the pipeline converter by going to reduced base or 1.5-bit schemes with calibration. In a multiplexed successive

[8]The general form of this principle is in [225] identified as the Sweeney-Robertson-Tocher division principle.

approximation converter also a calibration is required, which is mostly done at a architecture level.

A major difference between both converters is that in all forms of pipeline conversion, an amplification stage is used to suppress the errors and noise from the lower bit stages. This feature is missing in successive approximation implementations. The amplification stage is the most power hungry part in a pipeline converter, but it allows to achieve a higher signal-to-noise ratio.

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 multistage 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 8.81 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 [228].

These converters are often applied in image sensors. The pixel array of an image sensor is read out via its columns. A row of column analog-to-digital converters is used to process all pixels on one line. The simple structure of a counting converter fits well to the narrow pitch of less than 10 µm, the relatively slow-speed requirement, but high accuracy of 10–11 bits. In [227] the counting idea is

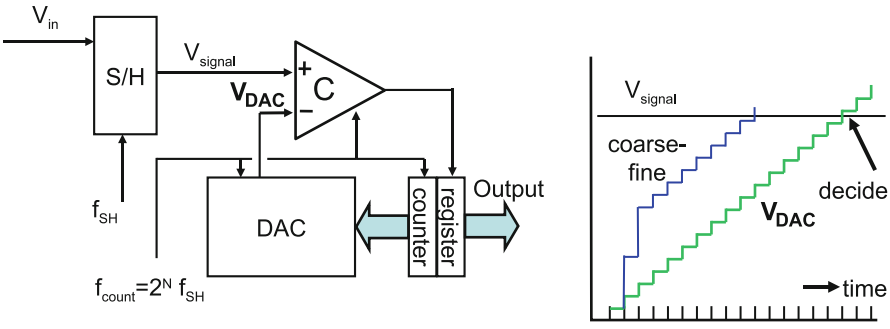


Fig. 8.81 A counting analog-to-digital converter. The *thin line* indicates a coarse step-fine step architecture [227]

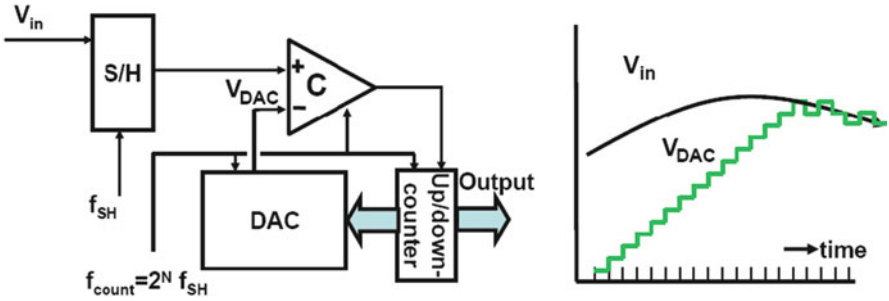


Fig. 8.82 A tracking analog-to-digital converter

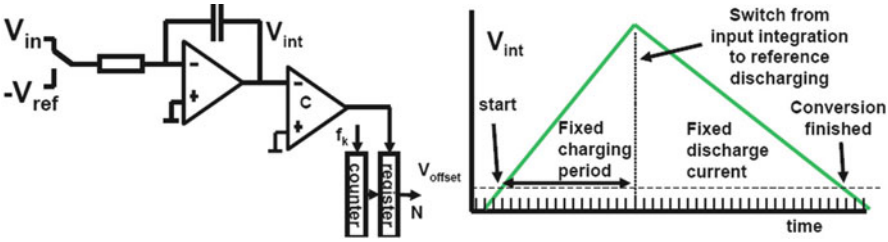


Fig. 8.83 A "dual-slope" analog-to-digital converter

implemented in a phase with increments of $2^n V_{\mathrm{LSB}}$, followed by a phase with V_{LSB} steps, thereby reducing the long conversion time; see the thin line in Fig. 8.81(right). similarities to algorithm converters certainly exist (compare Sect. 8.6.2).

The counting converter can be turned into a tracking analog-to-digital converter [229] by changing the counter in an up-down counter and connecting the comparator output to the up-down switch (Fig. 8.82). The system will operate in a way that the counter and digital-to-analog converter output will follow the input signal. The 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., [230], also show similarity with a tracking analog-to-digital converter.

A "dual-slope" converter is suited for slowly varying input signals. In Fig. 8.83 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:

$$H(\omega) = \frac{\sin(\pi \omega T_{\mathrm{int}})}{\pi \omega T_{\mathrm{int}}}. \tag{8.42}$$

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 (nonlinear capacitor) in Fig. 8.83 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 cannot be tolerated.

Dual-slope converters find their application especially in the harsh industrial environments and multimeters.

Example 8.10. 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.

Solution.

	DNL	INL	Absolute accuracy
Flash	Poor, limited by comparator mismatch	Poor, due to ladder and comparator mismatch	Poor, due to ladder and comparator mismatch
Successive approximation	Good, limited by digital-to-analog converter	Good, limited by digital-to-analog converter	Poor, due to comparator offset
Dualslope	Good, digital-to-analog converter	Good, digital-to-analog converter	Excellent, only drift

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. Demultiplexing 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.

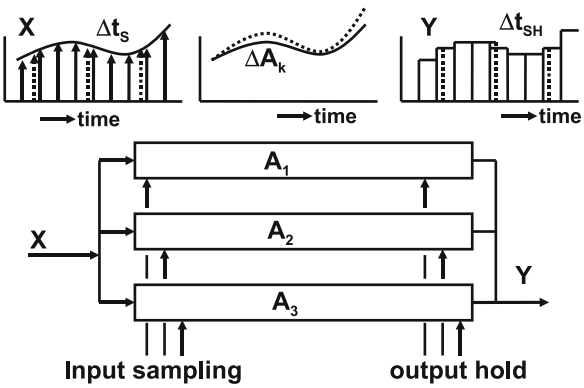


Fig. 8.84 Three different errors in a time-interleaved or multiplexed time-discrete structure

Figure 8.84 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. For example, a one-layer metalization gives asymmetrical crossings. And the speed of signal propagation is roughly half of vacuum, so 1 ps 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 11.7. 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. 8.25. 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.

In Sect. 7.7 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 [231, 232]; see Fig. 8.85 for $N = 2$.

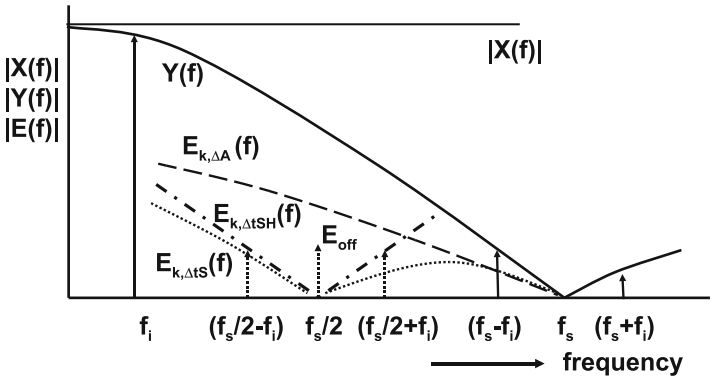


Fig. 8.85 The signal and the trajectory for different error components due to two-times multiplexing

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 $k = 0 \cdots (N-1)$. 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 Eqs. 3.12 and 4.2)

$$Y_k(f) = \sum_{r=-\infty}^{r=\infty} A_k X(j2\pi(f - rf_{sN}))\mathrm{e}^{j2r\pi fkT_s} \frac{\sin(\pi fT_s)}{\pi fT_s} \mathrm{e}^{-j\pi fT_s} \tag{8.43}$$

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 $\sin(\pi fT_s)/\pi fT_s$ 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 nonzero 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

$$Y(f) = \sum_{m=-\infty}^{m=\infty} AX(j2\pi(f - mf_s)) \frac{\sin(\pi f T_s)}{\pi f T_s} \mathrm{e}^{-j\pi f T_s} \tag{8.44}$$

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:

$$E_{\text{off}}(t) = \sum_{r=0}^{r=\infty} V_{\text{off},r} \sin(j2\pi r f_{sN} t). \tag{8.45}$$

- A static time error Δt_s in the input sampling of line k will result in

$$\left| \frac{E_{k,\Delta t}(j2\pi f_r)}{AX(j2\pi f_i)} \right| = \frac{2}{N} \sin(\pi f_i \Delta t_s) \frac{\sin(\pi f_r T_s)}{\pi f_r T_s}. \tag{8.46}$$

- An amplification error ΔA_k in line k leads to

$$\left| \frac{E_{k,\Delta A}(j2\pi f_r)}{AX(j2\pi f_i)} \right| = \frac{1}{N} \frac{\Delta A_k}{A} \frac{\sin(\pi f_r T_s)}{\pi f_r T_s}. \tag{8.47}$$

- A time error Δt_{SH} in the sample-and-hold pulse in line k gives

$$\left| \frac{E_{k,\Delta SH}(j2\pi f_r)}{AX(j2\pi f_i)} \right| = \frac{2}{N} \sin(\pi f_r \Delta t_{SH}) \frac{\sin(\pi f_i T_s)}{\pi f_r T_s}. \tag{8.48}$$

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 8.85 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.

An additional class of problems are frequency dependent deviations. Differences in bandwidths between signal processing sections that form a multiplexed structure, will also give rise to timing-dependent errors. Sections with differing bandwidths, will consequently show different group delays; see Eq. 2.157:

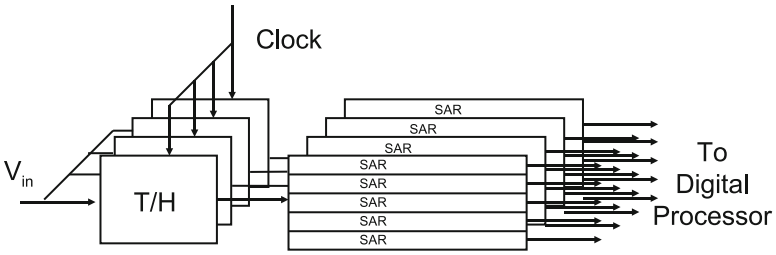


Fig. 8.86 Example of a multiplexed successive approximation analog-to-digital converter, for digitizing a full-spectrum video signal DOCSIS3.0 (BW = 1 GHz, $f_s = 480$ mW, 65-nm CMOS) [233] (courtesy photo: NXP design team)

$$\text{Group delay}(H_{\text{section }i}(\omega)) = \tau_{g,i} = \frac{\tau_i}{1+\omega^2\tau_i^2}, \tag{8.49}$$

where the bandwidth of section i is specified as $1/2\pi\tau_i$. If the input consists of a sine wave $A\sin(\omega t)$, the error signal between two sections i and j can be approximated by an additional term in one of the channels of the form:

$$A\omega(\tau_{g,i} - \tau_{g,j})\cos(\omega t)$$

The resulting error is proportional to the input frequency and bears similarity with the analysis of jitter; see Sect. 3.1.7. At a section bandwidth of approximately two times the sample rate and an input signal frequency of 1/3 of the sample rate, the difference in bandwidths in percent will translate in a similar difference in amplitude error.

Figure 8.86 shows a block diagram of a 64-channel multiplexed 2.6 Gs/s successive approximation converter [233]. The signal is first distributed over four track-and-holds. These circuits each drive the track-and-hold in the four section each containing 16 successive approximation converters. Figure 8.87 gives an impression: the main track-and-hold circuits are placed in the center in order to keep the wires to the four sections as balanced as possible. An overall performance of 48.5 dB in a 65-nm technology is achieved.

Next to interleaving successive approximation converters (see also [211]), also pipeline converters can be multiplexed interleaved and calibrated [234].

8.9 Implementation Examples

Based on a single comparator architecture of Sect. 8.1.8 a full-flash, a successive approximation, and a multistep pipeline converter are described.

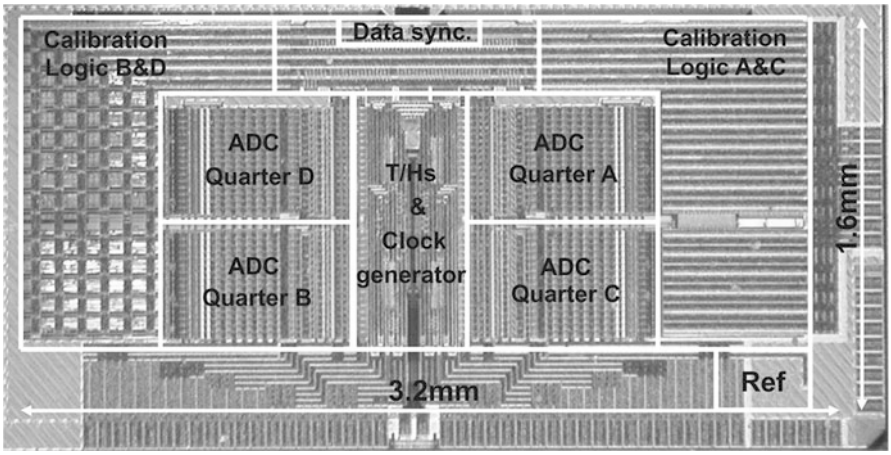


Fig. 8.87 Chip photograph of a multiplexed successive approximation analog-to-digital converter [233]

8.9.1 Full-Flash Analog-to-Digital Converter

The comparator described in the previous section forms the basis for three types of converters: an 8-bit full-flash analog-to-digital converter, a 10-bit successive-approximation analog-to-digital converter, and a 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 eight 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. 7.8.1: 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 require a careful design and layout.

The main characteristics of the analog-to-digital converter are summarized in Table 8.5.

Figure 8.88 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.

Table 8.5 Specifications of the three analog-to-digital converters

A/D converter	Full-flash	Successive app.	Multistep
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 (1 dB)	>70 MHz	20 MHz	70 MHz
Input signal swing	2 V	1.5 V	1.6 V
Ladder resistance	1,200 Ω	4,800 Ω	1,200 Ω
Active area	2.8 mm^2	1.2 mm^2	1.1 mm^2
		0.4 mm^2 (8 bit)	
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

All the converters are based on the comparator shown in Fig. 8.22

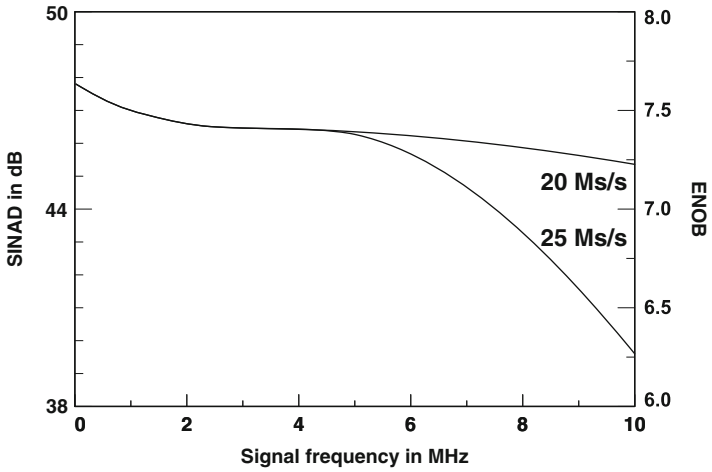


Fig. 8.88 Signal to noise and distortion at 20 and 25 Ms/s as a function of the input frequency

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 [235–237].

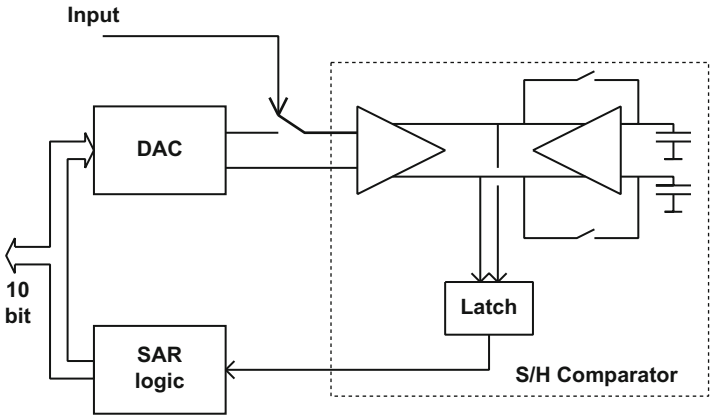


Fig. 8.89 Successive-approximation analog-to-digital converter

8.9.2 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. 8.89). 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 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 microcontroller inputs. Specifications are given in Table 8.5.

8.9.3 Multistep Analog-to-Digital Converter

The multistep analog-to-digital converter[9] (Fig. 8.90) is an 8-bit converter based on a technique involving a combination of successive approximation, flash, and

[9]This section is based on a design by Jeannet v. Rens.

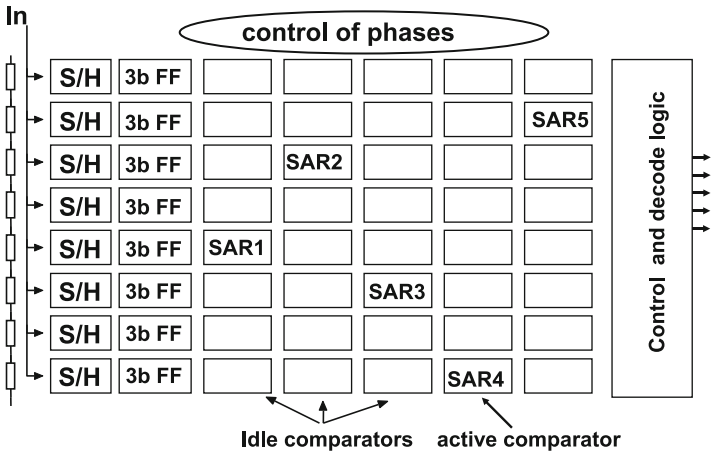


Fig. 8.90 Multi-step analog-to-digital converter (design: J. v. Rens)

time-interleaving. The three 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 seven channels of eight comparators in a flash structure. The channels operate in a time-interleaved manner. The actual conversion takes place in seven clock cycles. First, the input signal is sampled by, and stored in, one channel of eight 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. 8.91.

This converter was produced as a stand-alone product TDA8792 and was the building block for many video integrated circuits.

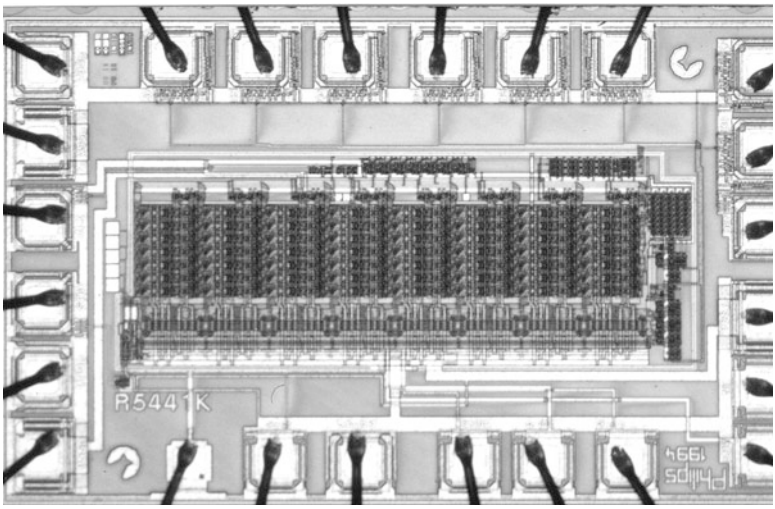


Fig. 8.91 Photograph of multistep analog-to-digital converter TDA8792

8.9.4 Comparison

Table 8.5 shows the main specifications of the three analog-to-digital converters. Basic to all converters is the comparator of Fig. 8.22, 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 have been extensively used in consumer ICs: digital video, picture-in-picture, instrumentation, etc. In 0.5 µm CMOS the 8-bit multistep runs at 50 Ms/s, while the 9-bit version achieves 8.2 ENOB.

Example 8.11. In an IC process input pairs (as used in comparators, gain stages, etc.) suffer from σ_{Vin}= 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 indication of the resolution).

Solution. If a signal range of approximately 1 V is assumed, a 5% error results. The main trade-offs are summarized in (Table 8.6):

Table 8.6 Comparison of converters in case of extreme mismatch

ADC architecture	Remarks	Resolution
Flash converter	2^N parallel comparators, limited by comparator mismatch	N = 4 bit
Pipeline and dual-slope converter	Comparator error will be cancelled, INL,DNL determined by digital-to-analog conversion	$N \approx 10$ bit
Logarithmic approximation, successive approximation	Comparator error results in offset, INL, DNL determined by digital-to-analog conversion	$N \approx 10$ bit

8.10 Other Conversion Proposals

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, which were labeled quantization errors. The sequence of sampling and quantizing can also be reversed. The level-crossing analog-to-digital converter in Fig. 8.92 [238–240] generates a new digital output code at each time moment an amplitude quantization level is passed. In its simplest form, this is a flash converter with unclocked comparators. With infinite time resolution this level-crossing algorithm will lead to a digital representation of the input signal with 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 nT_s, the amplitude error is

$$\Delta A(nT_s) = \frac{\mathrm{d}V_{\mathrm{in}}(t)}{\mathrm{d}t}\Delta T_s. \tag{8.50}$$

Assuming that the signal $V_{\mathrm{in}}(t) = A\sin(\omega t)$ 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

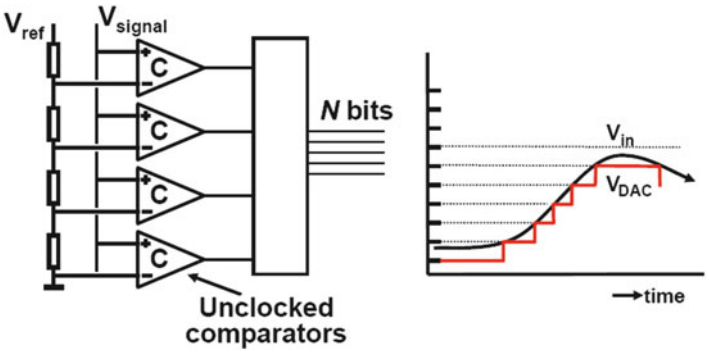


Fig. 8.92 The principle of a level-crossing analog-to-digital converter

$$E((\Delta A)^2) = \left(\frac{\mathrm{d}V_{\mathrm{in}}(t)}{\mathrm{d}t}\right)^2 \times E(\Delta T_s^2) = \frac{A^2\omega^2 T_s^2}{6}. \tag{8.51}$$

The resulting signal-to-noise ratio between signal power and error power is

$$\mathrm{SNR} = 10\log(3) - 20\log(\omega T_s). \tag{8.52}$$

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. A major implementation problem is that the delay between the actual level crossing and the resulting digital transition must be constant, which is not trivial as the delay of latches depends on the overdrive voltage (Sect. 8.1.4).

For low resolutions this principle converges towards asynchronous delta modulation (Sect. 9.8.2 [240]).

8.10.2 Feedback Asynchronous Conversion

Asynchronous converters constantly monitor the signal and create in a feedback path a best-fitting replica. Several low-resolution high-speed examples are reported [241, 242]. An example is shown in Fig. 8.93. The unclocked MSB comparator continuously monitors the signal. The MSB bit is fed into a simple digital-to-analog

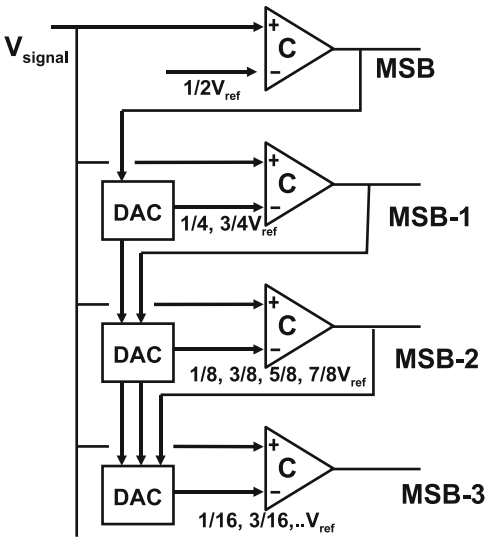


Fig. 8.93 The asynchronous successive approximation analog-to-digital converter [241]

converter that supplies either a $1/4V_{\text{ref}}$ or $3/4V_{\text{ref}}$. 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 Mismatch Dominated Conversion

In this entire chapter mismatch is regarded as a severe problem for performing analog-to-digital conversion. Nevertheless it is certainly possible to use mismatch to the advantage [243]. The comparators in Fig. 8.94 are designed to have a large input-referred mismatch. The group comparators connected to $V_{\text{ref},1}$ will span a voltage range of several tens of mV and follow a Gaussian distribution. In order to increase the range a second group of comparators is connected to $V_{\text{ref},2}$ overlapping the first range. The total network will show a probability density function. A summation network (e.g., a Wallace tree structure) is used to count the number of flipped comparators. This number allows to estimate the input voltage via the probability density function. The Wallace tree is a very power hungry network and the input range is determined by the $A_{V_{\text{T}}}$ coefficient.

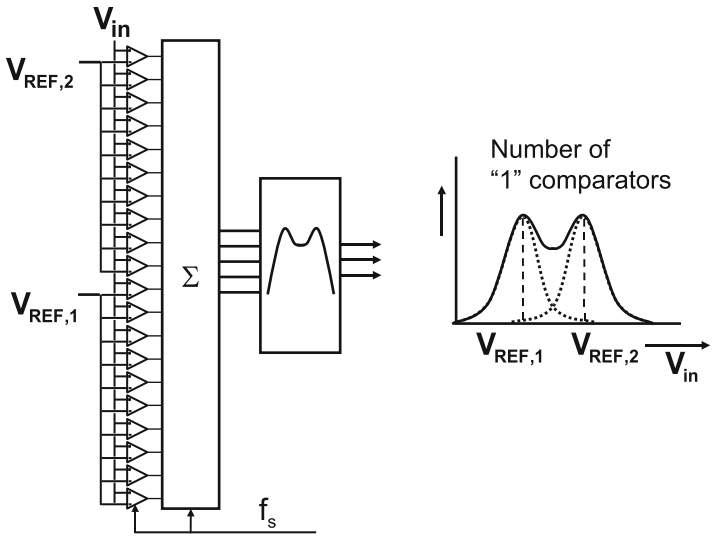


Fig. 8.94 Using mismatch of the comparators to cover an input range

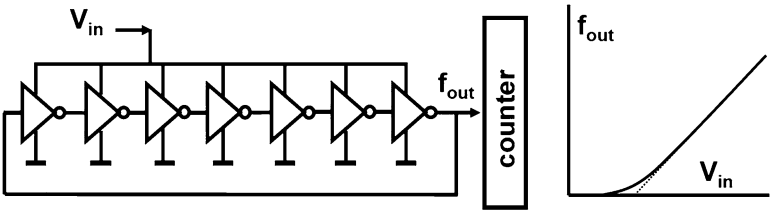


Fig. 8.95 VCO as voltage-to-frequency converter

8.10.4 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. 8.95. 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

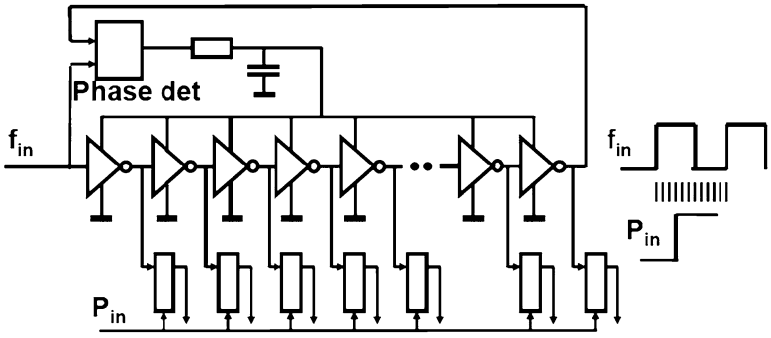


Fig. 8.96 Time-to-digital converter [244]

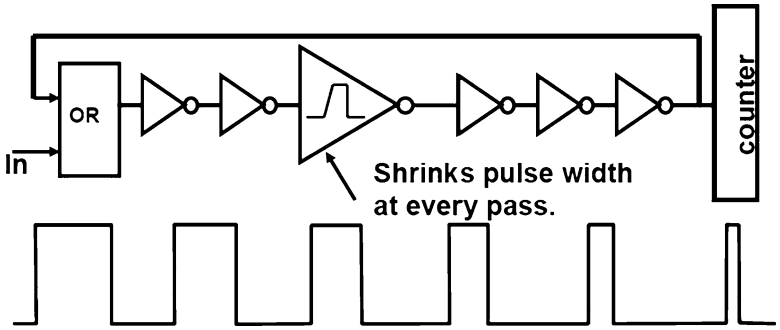


Fig. 8.97 Starving pulse time-to-digital converter [245]

clock is 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. 8.96 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. This method of quantization is often applied in phase-locked loops. The accuracy is limited to the jitter in this system. Power supply variations and substrate interference can also influence the quality. An accuracy in the range of 3–10 ps_{rms} is possible [244].

An elegant implementation is the starving pulse converter [245]. 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. 8.97.

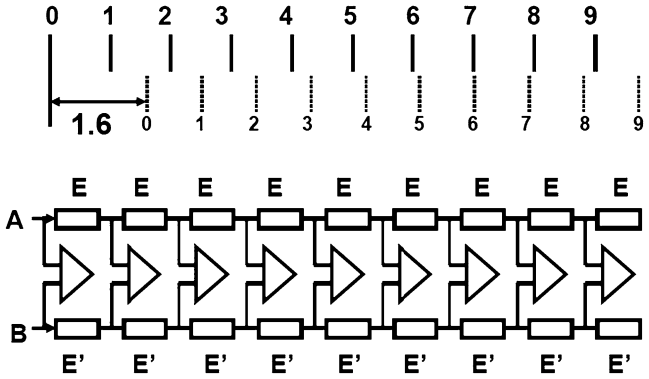


Fig. 8.98 The Vernier or Nonius principle

8.10.5 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 sixteenth- and seventeenth-century Spanish and French mathematicians) is widely used to determine accurately the subdivision in a primary scale. Figure 8.98(upper) shows a primary scale and a secondary scale (dashed). The secondary scale has a subdivision of 0–9 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. 8.98(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 [246]. 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 [247].

8.10.6 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 and signals with a high crest factor, the required resolution varies with the amplitude of the signal; see Sect. 5.3.2. In computing, this sort of problems is addressed by means of floating-point arithmetic. Also in analog-to-digital conversion floating-point

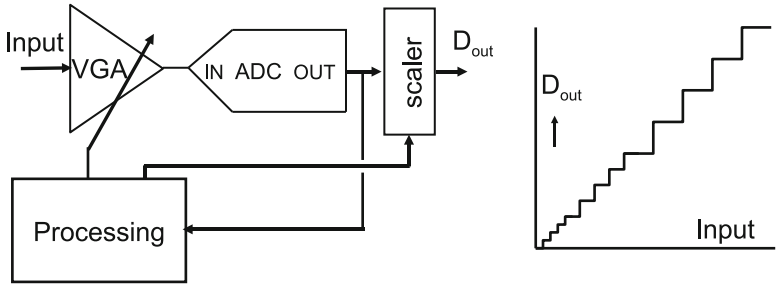


Fig. 8.99 The floating-point principle realized with a variable-gain amplifier (VGA)

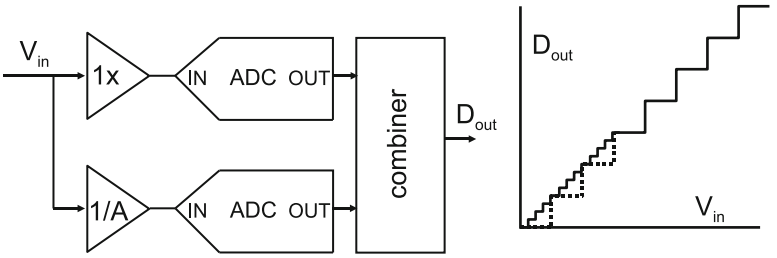


Fig. 8.100 The floating-point principle implemented with two parallel converters

conversion is possible, although most system engineers prefer a fixed-point converter. Figure 8.99 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 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., [248]. The application of this sort of converters is in sensor interfaces.

The floating-point analog-to-digital converter in Fig. 8.100 uses two parallel converters. The upper converter acts on the small signal levels while an attenuated version of input signal is applied to the lower converter. The advantage with respect to the VGA solution is that no time is lost when a range switch must occur. Moreover, the logic in the combiner can easily adjust for undesired gain and offset errors between both signal paths while making use of those signal levels that trigger both converters.

Similar to other multiplexed circuits, the timing of the sample pulses must be accurately controlled as well as the matching of the bandwidths for large and small signals. Matching the $1\times$ buffer with the $1/A$ attenuator is a challenging task.

Exercises

8.1. In a 65 nm process, the input transistors of the comparators are designed with $\mathrm{W/L} = 5/0.1\,\mu\mathrm{m}$. The effective gate-drain capacitance is 0.2 fF/µm. Estimate the maximum ladder disturbance, if a 7-bit full-flash converter with a total ladder impedance of 1 kΩ is designed.

8.2. Deduce from Fig. 8.60 what is the maximum comparator offset that can be tolerated.

8.3. In a 6-bit full-flash 1 Gs/s analog-to-digital converter the wires of the clock lines must be kept as equal as possible. What is the maximum wireline difference that can be tolerated if the converter must operate at Nyquist frequency? Assume a propagation speed of 10^8 m/s.

8.4. Rank the following design parameters for achieving a high BER in order of importance: the gate width of the latch transistors, the length, the current, the gate oxide thickness, and the gate-drain overlap capacitance.

8.5. Repeat the example of Table 8.2 with 0.25 µm technology data and an input range of 2 V. What is surprising?

8.6. Design a decoder for a 5-bit flash converter based on a Wallace tree summation network.

8.7. Show that for a sinusoidal signal the digital output power of an N-bit parallel output port is smaller than for a serialized single pin output.

8.8. At the input of an 8-bit analog-to-digital converter a ramp-type signal rising at 1 LSB per sample period is applied. What is the power consumption at the digital output compared to the maximum power? What happens to the power in case of 1 LSB DNL errors, and what in case of an 1 LSB INL error?

8.9. How can comparator mismatch in a multiplexed successive approximation converter affect the performance?

8.10. A 7-bit flash converter uses 128 resistors of 25 Ω each with a 50 fF parasitic capacitance to each comparator. If the internal signal swing of the comparator is 1 V, calculate the kickback amplitude. Does the kickback vary with input signal level? Where is the worst case level?

8.11. A sine wave must be converted with good absolute accuracy (without DC offsets). No auto-zero or calibration mechanism is available. Give a reason why a flash converter is a better choice than a successive approximation converter.

8.12. In Fig. 8.6 exchange the connections of the $\overline{clock}$ signal and the signals coming from the pre-latch. Does this solve the hysteresis problem and what is the cost?

8.13. Draw a transfer curve of a 2-bit coarse and 4-bit fine analog-to-digital converter. Add the transfer curve in case one of the coarse comparators has an offset of 1% of the input range. Do the same if this offset applies to a fine comparator.

8.14. A 1.6 Gs/s analog-to-digital converter is build by multiplexing 64 successive approximation converters. If offsets are normally distributed calculate σ_{comp} for a 50 dB performance of 95% of the dies. The bandwidth of the T&H circuits is 1.5 GHz. How much variation is allowed on this bandwidth?

8.15. A 1 LSB difference produces 10 mV voltage swing on the nodes of a latch in a comparator with parasitic capacitances of 100 fF. The maximum current in each branch of the latch is 100 µA. The current factor for a square transistor is 200 $\mu A/V^2$. Calculate the BER for a maximum decision time of 1 ns.

8.16. A comparator generates 50 fC of kickback charge. The nodes of a ladder have a parasitic capacitance of 150 fF each. If the ladder can consume 1 mA from a 1 V reference source and the converter must run at 2 Gs/s, what is the maximum resolution?

Chapter 9
Sigma–Delta Modulation

Abstract The sigma–delta modulator is a specific algorithm to convert low-bandwidth signals into a digital data stream. In the first section the effects of oversampling are described both for the analog-to-digital converter as well as for the analog-to-digital converter.

The next step is the noise-shape topology. The feedback mechanism that is applied to the quantization errors allows to free a fraction of the bandwidth of this error energy. The basic mechanisms in the noise shaper are applied to show the effects of increasing the oversample rate and increasing the order of the feedback filter.

A similar approach is adopted to deal with the sigma–delta topology. The time-discrete implementation allows to design first- and second-order modulators. The main extension towards higher orders is found in the cascaded sigma–delta modulator.

The time-continuous sigma–delta modulators use simple transconductance-capacitor filters and their properties differ in a number of aspects from the time-discrete variants. Some implementation aspect are given. The discussion on the advantages of multi-bit conversion is summarized.

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 [249], Cutler [250], and Widrow [251] 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 [252] proposed to place the 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.

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_9,

© Springer Science+Business Media, LLC 2013

9.1 Oversampling

9.1.1 Oversampling in Analog-to-Digital Conversion

In the process of sampling and quantization, an error signal is generated that is fundamentally a distortion component. This quantization error and the multitude of distortion components folded back by various multiples of the sampling frequency are 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 (one is not an integer multiple of the other). This quantization energy has a fixed amount of power $Q^2 = V_{\mathrm{LSB}}^2/12$ for a given resolution N. It is modeled as a noise component that is evenly distributed 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 $f_{s,\mathrm{ny}}$ 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 factor of four in power density and the effective resolution (ratio between signal and unwanted components) increases by 1 bit. The ratio

$$\mathrm{OSR} = \frac{f_s}{f_{s,\mathrm{ny}}} = \frac{f_s}{2f_b} \tag{9.1}$$

is called the oversampling ratio (OSR). The bandwidth BW of the signal is for ease of calculation assumed to range from 0 to f_b. The same result is obtained for bands at other frequency locations.

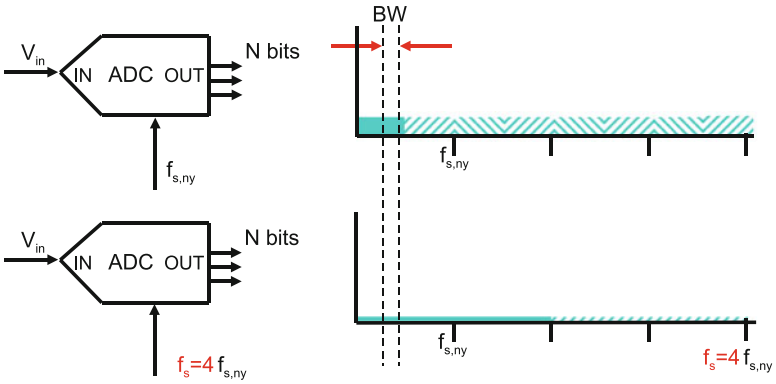


Fig. 9.1 Oversampling in analog-to-digital conversion. In the derivations and figure the lower boundary of the bandwidth is set to 0 Hz

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 $f_s/2$:

$$N_Q(f) = \frac{Q^2}{f_s/2} = \frac{V_{\mathrm{LSB}}^2/12}{f_s/2} = \frac{V_{\mathrm{LSB}}^2}{6f_s}. \tag{9.2}$$

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 noise power

$$Q_b^2 = \frac{V_{\mathrm{LSB}}^2 f_b}{6f_s} = \frac{V_{\mathrm{LSB}}^2}{12}\frac{1}{\mathrm{OSR}}. \tag{9.3}$$

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

$$\Delta\mathrm{SNR} = {}^{10}\log\left(\frac{f_s}{f_{s,\mathrm{ny}}}\right) = {}^{10}\log(\mathrm{OSR}) \tag{9.4}$$

$$\Delta\mathrm{ENOB} = \frac{1}{2}{}^{2}\log\left(\frac{f_s}{f_{s,\mathrm{ny}}}\right) = \frac{1}{2}{}^{2}\log(\mathrm{OSR}). \tag{9.5}$$

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. 3.9). 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.

9.1.2 Oversampling in Digital-to-Analog Conversion

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 setup has a relatively poor

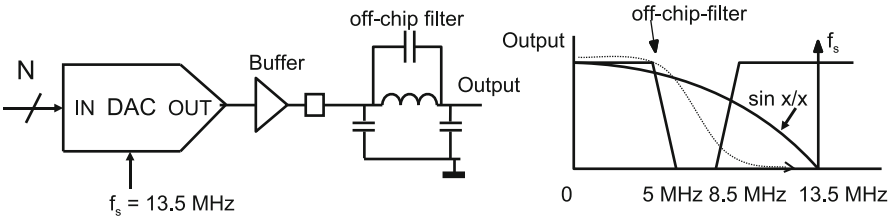


Fig. 9.2 A digital-to-analog converter running at Nyquist rate with the output filtered by an external filter

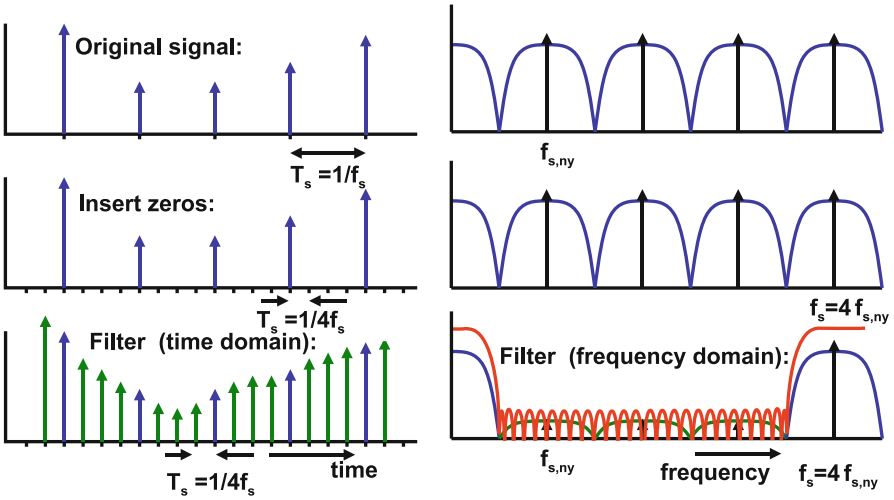


Fig. 9.3 Four-fold digital oversampling in time and frequency domain

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 a sigma–delta output signal 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 noncritical post-filtering (first order) is sufficient to restore the analog signal.

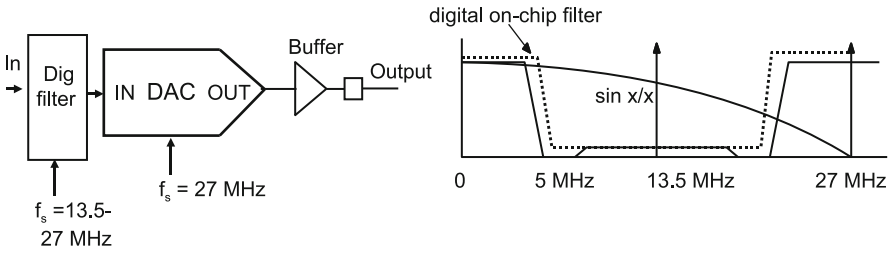


Fig. 9.4 A digital-to-analog converter driven by digital pre-filter that performs a two-time oversampling

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 and power for dig. filter
$\sin(x)/x$ loss of 4 dB	$\sin(x)/x$ loss = 0.5 dB
Standard sample rate	2× sample rate needed

The inherent $\sin(x)/x$ is reduced by an order of magnitude, so no compensation 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.

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 $J(z)$ 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. Therefore there is no feedback stability problem. The inverted quantization error is delayed and added to the signal. Low-frequency components in the error signal are effectively suppressed. At high frequencies the filter delay will cause a phase shift leading to an enhanced error energy level.

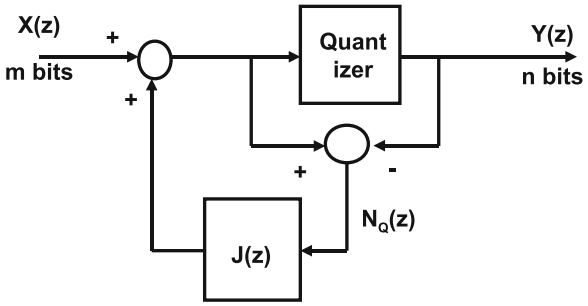


Fig. 9.5 Signal-to-noise improvement with noise shaping

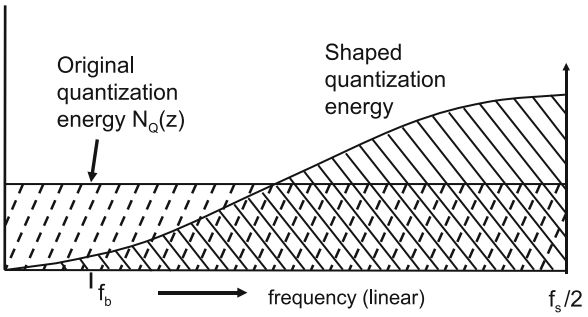


Fig. 9.6 The noise shaper pushes the quantization energy to higher frequencies

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 $z \leftrightarrow \mathrm{e}^{j\omega T_s}$:

$$Y(\omega) = X(\omega) + (1 - \mathrm{e}^{-j\omega T_s})N_Q(\omega). \tag{9.8}$$

In order to determine the noise transfer to the output (NTF), the absolute value of the function in brackets is evaluated:

$$\begin{aligned} |\mathrm{NTF}(\omega)|^2 &= \left|\frac{Y(\omega)}{N_Q(\omega)}\right|^2 = |1 - \mathrm{e}^{-j\omega T_s}|^2 \\ &= 2 - 2\cos(\omega T_s). \end{aligned} \tag{9.9}$$

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

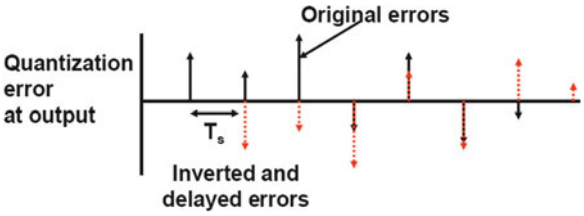


Fig. 9.7 The output of the noise shaper contains both the original error sequence and its inverted and delayed version. Both are not visible at the output as in this plot, because they are part of the quantized signal. This plot helps in realizing that the DC content of the error combination is low

function $J(z) = z^{-1}$. Integration of the noise power density function over the band from 0 to $f_s/2$ gives

$$\frac{V_{\mathrm{LSB}}^2}{6f_s} \int_{f=0}^{f=f_s/2} 2 - 2\cos(\omega T_s)\mathrm{d}f = \frac{V_{\mathrm{LSB}}^2}{6} \tag{9.10}$$

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

$$\frac{V_{\mathrm{LSB}}^2}{6f_s}(2 - 2\cos(\omega T_s)) \approx \frac{V_{\mathrm{LSB}}^2}{6f_s}\left(\frac{2\pi f}{f_s}\right)^2. \tag{9.11}$$

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:

$$\begin{aligned}\int_{f=0}^{f_b} \frac{V_{\mathrm{LSB}}^2}{6f_s}|1 - \mathrm{e}^{-2j\pi f T_s}|^2\mathrm{d}f &= \frac{V_{\mathrm{LSB}}^2}{3}\left[\frac{f_b}{f_s} - \frac{\sin(2\pi f_b/f_s)}{2\pi}\right] \\ &\approx \frac{V_{\mathrm{LSB}}^2}{12}\frac{\pi^2}{3}\left(\frac{2f_b}{f_s}\right)^3 = \frac{V_{\mathrm{LSB}}^2}{12}\frac{\pi^2}{3\mathrm{OSR}^3}.\end{aligned} \tag{9.12}$$

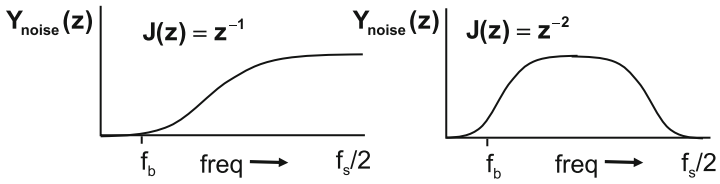


Fig. 9.8 *Left*: The noise spectrum in the output with a single sample pulse delay. *Right*: the spectrum with two sample pulses delay

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 (Eq. 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 $\sin(x) = x - x^3/6$

$$\begin{aligned}\text{Noise power reduction} &= \frac{4}{3}(\pi f_b/f_s)^2\\ \text{Noise amplitude reduction} &= \frac{2}{\sqrt{3}}(\pi f_b/f_s).\end{aligned} \tag{9.14}$$

The total in-band noise power is related to the OSR by a cube power in Eq. 9.12. One OSR term comes from the oversampling mechanism; the remaining OSR^2 term comes from the noise shaping loop. Doubling the OSR 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.

Example 9.1. In a noise shaper the feedback path function equals $J(z) = z^{-2}$. Sketch the noise behavior at the output of this noise shaper.

Solution. Figure 9.8 shows the noise spectrum in the output. In case $J(z) = z^{-2}$, the shape of the noise in the output is $1 - J(z) = 1 - z^{-2}$, which can be transformed to the frequency domain via $z \leftrightarrow \mathrm{e}^{j2\pi f}$.

9.2.1 Higher Order Noise Shaping

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 nth-order polynomial,

$$1 - J(z) = (1 - z^{-1})^n. \tag{9.15}$$

The frequency transfer function becomes

$$Y(\omega) = X(\omega) + \left(1 - \mathrm{e}^{-j\omega T_s}\right)^n N_Q(\omega). \tag{9.16}$$

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 nth-order noise shaper is found in a similar way as the first order [8, formula 300]:

$$\begin{aligned}
&\int_{f=0}^{f_b} \frac{V_{\mathrm{LSB}}^2}{6f_s} |1 - \mathrm{e}^{-2j\pi f T_s}|^{2n} \mathrm{d}f \\
&\quad = \int_{f=0}^{f_b} \frac{V_{\mathrm{LSB}}^2}{6f_s} (2\sin(\pi f/f_s))^{2n} \mathrm{d}f \\
&\quad = \frac{V_{\mathrm{LSB}}^2}{6\pi} \frac{(2n)!}{(n!)^2} \left[\pi \frac{f_b}{f_s} - \cos\left(\pi \frac{f_b}{f_s}\right) \sum_{k=0}^{k=n-1} \frac{2^{2k}(k!)^2 (2n)! \sin^{2k+1}(\pi f_b/f_s)}{(2k+1)!} \right].
\end{aligned} \tag{9.17}$$

The signal-to-noise ratio gain is found by dividing this result by the total noise in a Nyquist band $V_{\mathrm{LSB}}^2/12$:

$$\text{SNR gain} = \frac{2}{\pi} \frac{(2n)!}{(n!)^2} \left[\pi \frac{f_b}{f_s} - \cos\left(\pi \frac{f_b}{f_s}\right) \sum_{k=0}^{k=n-1} \frac{2^{2k}(k!)^2 (2n)! \sin^{2k+1}(\pi f_b/f_s)}{(2k+1)!} \right]. \tag{9.18}$$

Figure 9.9 shows the signal-to-noise gain for first till sixth order noise shaping.

Assuming that f_b/f_s is sufficiently small, a simpler expression can be obtained for the noise power:

$$\begin{aligned}
\int_{f=0}^{f_b} \frac{V_{\mathrm{LSB}}^2}{6f_s} |1 - \mathrm{e}^{-2j\pi f T_s}|^{2n} \mathrm{d}f &\approx \int_{f=0}^{f_b} \frac{V_{\mathrm{LSB}}^2}{6f_s} (2\pi f/f_s)^{2n} \mathrm{d}f \\
&= \frac{V_{\mathrm{LSB}}^2}{12\pi} \frac{(2\pi f/f_s)^{2n+1}}{2n+1} \\
&= \left(\frac{V_{\mathrm{LSB}}^2}{12}\right) \times \left(\frac{1}{\mathrm{OSR}}\right) \times \left(\frac{\pi^{2n}}{(2n+1)\mathrm{OSR}^{2n}}\right).
\end{aligned} \tag{9.19}$$

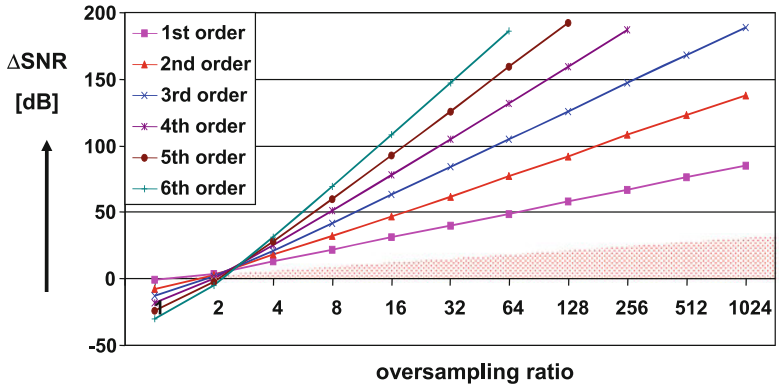


Fig. 9.9 The gain in signal-to-noise ratio for an nth-order noise shaper as calculated in Eq. 9.18. The *shaded triangle* is the portion that originates from oversampling alone. This plot is equally valid for sigma–delta conversion. The same curves apply to nth-order sigma–delta conversion

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 $(2n+1) \times 6\,\mathrm{dB}$ for a doubling of the oversampling rate. The SNR gain of nth-order noise shaping over Nyquist conversion is given as

$$\Delta\mathrm{SNR} = 10^{10}\log\left((2n+1)\frac{\mathrm{OSR}^{2n+1}}{\pi^{2n}}\right) = 20^{10}\log\left(\frac{\mathrm{OSR}^{n+0.5}\sqrt{2n+1}}{\pi^{n}}\right). \tag{9.20}$$

This expression has an error of less than 1 dB with respect to Eq. 9.18 and Fig. 9.9 for OSRs larger than 4.

This expression is based on the assumption in Eq. 9.15 that $1 - J(z)$ 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

$$Y(z) = \frac{X(z)}{1-cz^{-1}} + \frac{1-J(z)}{1-cz^{-1}}Q(z) \tag{9.21}$$

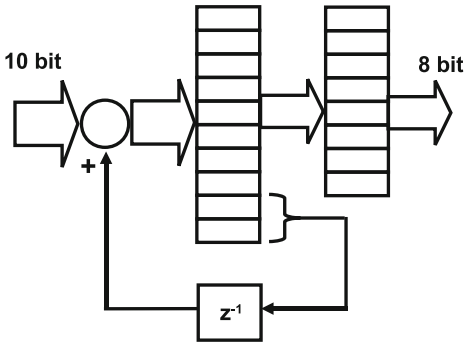


Fig. 9.10 Noise shaping in digital rounding

which implies a linear signal distortion for the input signal and a 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.

Example 9.2. A digital signal processor delivers a 10-bit result. This result must be processed to fit optimally to an 8-bit bus. Give a solution using noise shaping.

Solution. The problem to truncate an 10-bit digital data word to 8 bits can be solved by means of a simple form of a first-order noise shaper as shown in Fig. 9.10. The truncation is a form of quantization. The two residue bits form the quantization error. These 2 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.

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] as shown in Fig. 9.11 [252–256]. Sigma–delta conversion as a form of signal quantization 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 [252] 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.

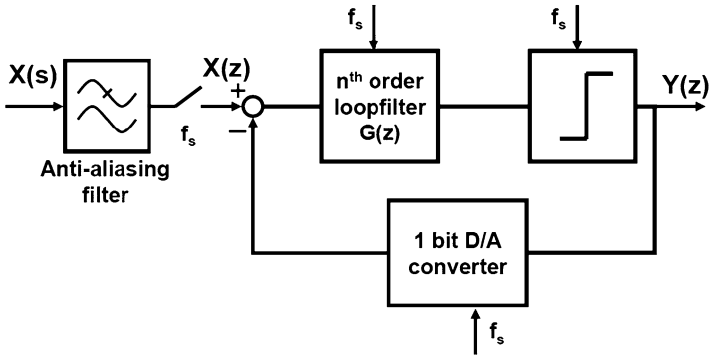


Fig. 9.11 Basic scheme of a sigma–delta modulator

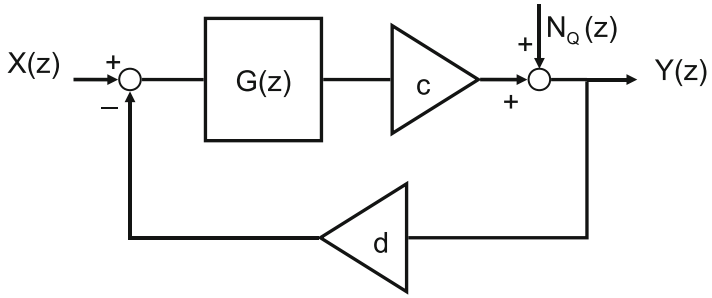


Fig. 9.12 The non-linear quantizer has been replaced by a gain factor and a noise injection point

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 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 behavior of modulators.

As the quantizer is a nonlinear element, it cannot be fully modeled in a linear system. In every model description the quantizer has to be approximated by linear components. Figure 9.12 shows that the quantizer is replaced by a gain factor c and the quantization errors N_Q have been made explicit by adding them with a summation node. When the quantizer's effective gain is modeled as c and the digital-to-analog converter gain as d, the input-output relation becomes

$$Y(z) = \frac{cG(z)}{1+cdG(z)}X(z) + \frac{1}{1+cdG(z)}N_Q(z) \tag{9.22}$$

$G(z)$ is designed to have a high gain within the desired bandwidth of $X(z)$. For those frequencies the equation reduces to

$$Y(z) = X(z) + \frac{1}{cdG(z)} N_Q(z). \tag{9.23}$$

If this equation is compared to the derivation for the noise shaper, Eq. 9.6, then a mathematical equivalence will occur if

$$cdG(z) \leftrightarrow \frac{1}{1 - J(z)} = \left(\frac{1}{1 - z^{-1}}\right)^n.$$

The last term can be implemented as a time-discrete integrator. The transfer function of a sigma–delta modulator with a loop filter dominated by an nth-order filter is mathematically comparable to an nth-order noise shaper. The same signal-to-noise improvement as in Fig. 9.9 is obtained.

The gain factor c of 1-bit quantizers is created by the comparator. The output of the comparator is given $(-1, +1)$; however, the input signal can be very small. Consequently a rather large gain can be expected. In a design the gain is mostly established from simulating the input power versus the output power (see also Sect. 9.5). For multi-bit quantizers the ratio between the quantizer LSB size V_{LSB} and the corresponding digital-to-analog converters output step is a good measure for c.

The feedback loop shapes the noise power from a flat spectrum into a shape which is determined by the loop filter $G(z)$. If the amplification of the digital-to-analog conversion is set to $d = 1$, the substitution of $z \leftrightarrow \mathrm{e}^{j\omega T_s}$ results in a frequency domain representation:

$$Y(\omega) = \frac{cG(\omega)}{1 + cG(\omega)} X(\omega) + \frac{1}{1 + cG(\omega)} N_Q(\omega). \tag{9.24}$$

At high values of $cG(\omega)$ the quantization errors are suppressed by this amount, and the output signal equals the input signal. The slope of the noise spectrum is the inverse of the slope of the filter. In the frequency band where the filter does not produce much gain, the quantization power in the output signal increases (see Fig. 9.13). An increase of noise above the original noise spectral density of the quantization is visible close to the half of the sample rate. Here also patterns are observed that resemble frequency modulation products (a sigma–delta modulator behaves here as a VCO).

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 (NTF):

$$\begin{aligned} Y(\omega) &= \mathrm{STF}(\omega) X(\omega) + \mathrm{NTF}(\omega) N_Q(\omega) \\ \mathrm{STF}(\omega) &= \frac{cG(\omega)}{1 + cG(\omega)} \\ \mathrm{NTF}(\omega) &= \frac{1}{1 + cG(\omega)}. \end{aligned} \tag{9.25}$$

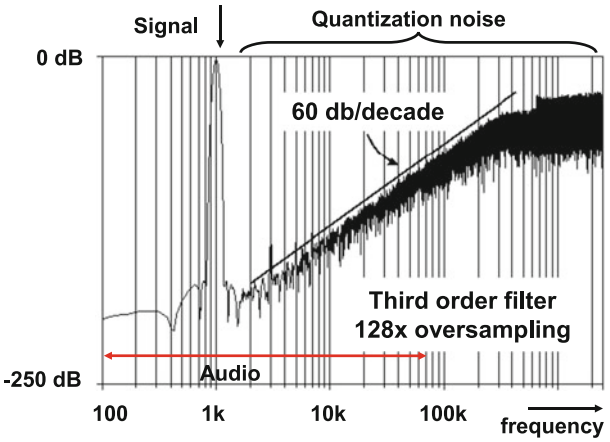


Fig. 9.13 Example of a frequency spectrum at the output of an audio sigma–delta modulator. The inverse third-order filter characteristic can be observed in the noise pattern

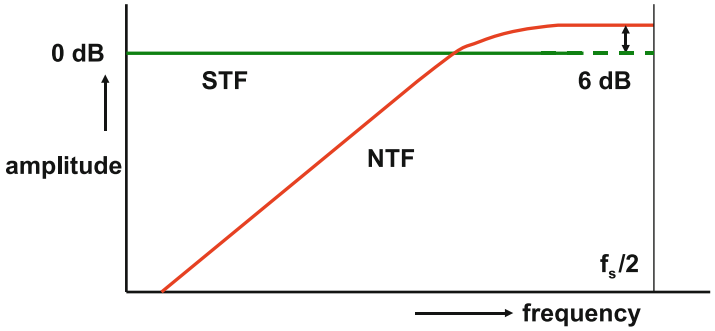


Fig. 9.14 The ideal signal transfer and a first-order noise transfer function

Figure 9.14 shows both transfer functions. Often system designers like to see a flat STF: the signal passes through the sigma–delta modulator without experiencing filtering. However, at high frequencies where the active elements lose their gain, bumps and other irregularities may appear. The NTF is designed to suppress the noise at the band of interest. In that band it mimics the inverse filter function.

9.3.1 Overload

In the early realizations of sigma–delta conversion the quantizer is a comparator. The output signal consists of a 1-bit signal with two levels: either $0, 1$ or more symmetrical $-1, +1$. In Fig. 9.15 a sine wave is shown with such a 1-bit representation. The sine wave is here at 85% of the range of the feedback digital-to-analog

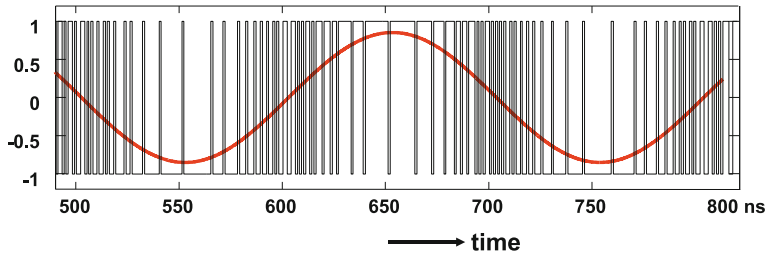


Fig. 9.15 A sine wave with an amplitude of 85% of the full scale is represented by a pulse density signal of the digital-to-analog converter

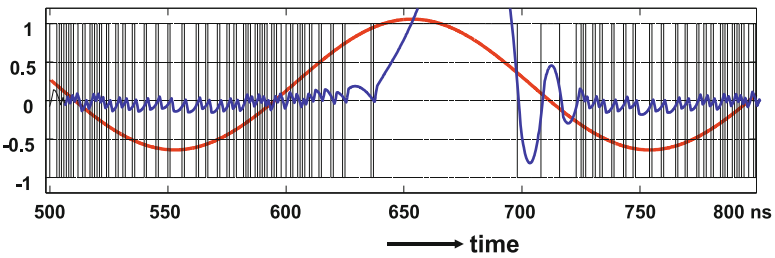


Fig. 9.16 A sine wave with an amplitude of 85% of the full scale and 20% bias creates an overload situation with instability. Shown are the input signal, the signal after the loop filter and the digital code

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 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 period of instability (see Fig. 9.16).

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

$$V_i(nT_s) \text{ is bounded by } \pm V_{\text{DAC,max}}, \forall n, \forall i. \tag{9.26}$$

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. 7.5 the 1-bit representation of a sine wave generated by sigma–delta modulation shows a lot of quantization or 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:

$$\text{Power of 1-bit digital} \leftrightarrow \text{Power of maximum sine wave}$$

$$P = \frac{V^2}{R} = \frac{(\pm V_a)^2}{R} = \frac{V_a^2}{R} \leftrightarrow P = \int \frac{(\alpha_{OL} V_a \sin(\omega t))^2}{R} \mathrm{d}t = \frac{\alpha_{OL}^2}{2} \frac{V_a^2}{R} \tag{9.27}$$

shows that more than half of the power in a 1-bit signal consists of unwanted distortion. Next to this unwanted power, the overload factor causes another -2 to -4 dB loss in the practical implementation.

The 1-bit representation of a sigma–delta output signal is therefore largely composed of high-order distortion products. Any nonlinearity 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 nonlinear operations such as rounding.

Example 9.3. What indicates "instability" in a sigma–delta converter? What measures can be taken to recover from instability?

Solution. Instability for a sigma–delta converter is a non-oscillating output or an oscillating output whose time average does not follow the input signal. A sigma–delta converter is stable if it oscillates in a manner that the time average corresponds to the input signal.

For a (controlled) oscillation, a loop gain of 1 and a phase delay of 180° is needed. In case of instability, adjusting the state of the integrators is needed to return to a controlled oscillation. These adjustments can be implemented by limiting the integration or by a reset switch.

9.3.2 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

[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.

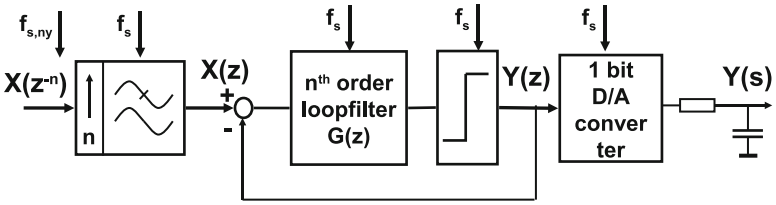


Fig. 9.17 A digital-to-analog sigma–delta modulator

width to a smaller word width (see Fig. 9.17). 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 intact. The resulting bit stream, e.g., a 1-bit stream, contains the required signal. In some applications with a high OSR, a passive filter is sufficient to obtain an analog signal of high quality. As mentioned before, the 1-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 [257, 258] or even 19 bit [259] with various optimization tricks. A full 20-bit signal in a 20 kHz bandwidth corresponds to a jitter of 8 $\mathrm{ps_{rms}}$, 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. 7.3.5.

9.4 Time-Discrete Sigma–Delta Modulation

9.4.1 First-Order Modulator

Comparison of the transfer function of the noise shaper and the sigma–delta modulator suggests that the filter function $G(z)$ can be obtained by rewriting $J(z)$. This is not the optimum strategy. A noise shaper is in fact not a feedback loop allowing more freedom to design $J(z)$. The filter $G(z)$ in a sigma–delta modulator is an essential part for stable conversion.

Figure 9.18 shows a first-order time-discrete sigma–delta modulator. The loop filter is implemented as a first-order integrator. A small DC input signal is applied in the simulation shown in Fig. 9.19. The dominant 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 provide sufficient conditions to implement a stable oscillation (Eq. 2.142). The DC input signal is superimposed on this oscillation.

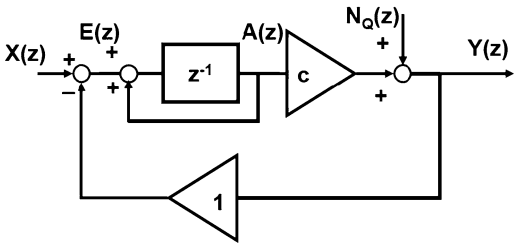


Fig. 9.18 A first-order sigma–delta modulator

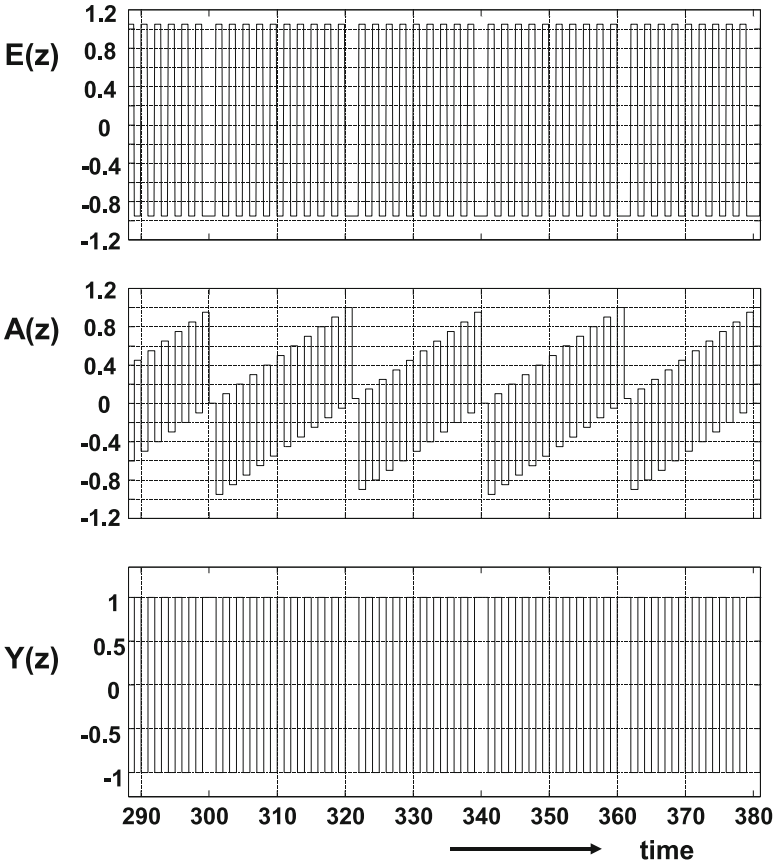


Fig. 9.19 The signals in a first-order time-discrete sigma–delta modulator with a small DC input signal. *Top*: the signal after the summation node. *Middle*: the integrator output. *Bottom*: the comparator output

The input causes the integrator to slightly increase its output value. After repeated passes of the oscillation signal, it reaches a level where the returning block wave is not sufficient to reach the threshold of the comparator. The oscillation wave skips one transition. In the long-term output pattern these skipped transitions result after averaging the representation of the DC input signal. The input signal modulates the oscillation.

The above transfer function for the first-order sigma–delta configuration is split in the STF and the NTF. With $c \gg 1$ and $d = 1$ and an ideal integrator filter $G(z) = z^{-1}/(1 - z^{-1})$ these functions reduce to

$$\begin{aligned} \mathrm{STF}(z) &= \frac{Y(z)}{X(z)} = \frac{cG(z)}{1 + cdG(z)} \approx z^{-1} \\ \mathrm{NTF}(z) &= \frac{Y(z)}{N_Q(z)} = \frac{1}{1 + cdG(z)} \approx 1 - z^{-1}. \end{aligned} \tag{9.28}$$

The signal passes unaltered through the modulator, delayed by one sample pulse. The noise is filtered, as can be seen in the frequency domain

$$|\mathrm{NTF}(\omega)|^2 = (2\sin(\omega T_s/2))^2 = 2 - 2\cos(\omega T_s/2). \tag{9.29}$$

At half of the sample rate $\omega T_s/2 \approx \pi/2$, the NTF equals 2 (see Fig. 9.14).

The total noise power in a band from 0 to f_b is found in the same way as for the noise shaper:

$$\int_{f=0}^{f_b} \frac{V_{\mathrm{LSB}}^2}{6f_s} |1 - \mathrm{e}^{-2j\pi f T_s}|^2 \mathrm{d}f \approx \frac{V_{\mathrm{LSB}}^2}{12} \frac{\pi^2}{3} \left(\frac{2f_b}{f_s}\right)^3 = \frac{V_{\mathrm{LSB}}^2}{12} \frac{\pi^2}{3\mathrm{OSR}^3}. \tag{9.30}$$

The NTF is in this example a first-order high-pass filter: suppression of noise close to DC while the noise is two-times amplified close to half of the sampling rate.

Figure 9.19 reveals another phenomena in sigma–delta modulation. The lower trace shows an output pattern with a dominant $f_s/2$ component. The regular interruption of this pattern caused by the DC offset creates low-frequency signal components. These frequency components are called "idle tones." At low-signal levels these tones can pop up in the desired frequency band [260]. Idle tones are a major drawback of first-order converters. These correlated components, see Fig. 9.20, can present serious problems in applications.[4] With higher-order modulators or the addition of helper signals (dither) these tones are reduced to acceptable levels.

9.4.2 *Cascade of Integrators in FeedBack or Feed-Forward*

A first-order sigma–delta modulator oscillates stably as long as the signal is sufficiently within the range of the feedback digital-to-analog converter. The addition of

[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.

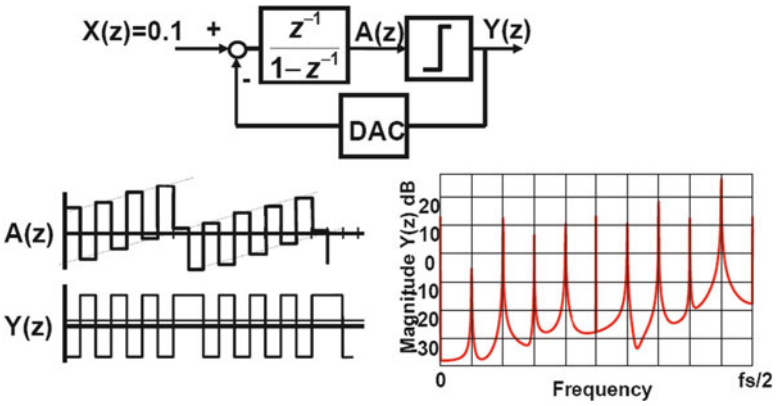


Fig. 9.20 A small DC offset in a first-order sigma–delta modulator creates a pattern with frequency components in the signal band. These idle tones occur in low-order sigma–delta modulators

a second integrator stage creates a loop where the negative feedback (180°) plus the contribution of two integrator stages $2 \times 90°$) 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 only conditionally stable. Proper conversion depends on input signal level, feedback, and delay. Filter topologies must combine first-order and second-order signals to reduce the overall phase shift below 360° and guarantee proper operation.

In Sect. 3.2 two types of time-discrete filters were introduced: finite impulse response filters are based on the addition of delayed copies of the signal and infinite impulse response filter uses feedback in addition. In a similar way the second-order filters inside a sigma–delta modulator can be classified: cascade of integrators in feed forward (CIFF) and cascade of integrators in feedback (CIFB). Figure 9.21 shows second-order filter constructed as CIFF and CIFB.

In the CIFF structure the input signal and feedback signal are subtracted in the first integrator. This stage must be carefully designed as the sharp transitions of the 1-bit digital-to-analog converter easily can drive the opamp into distortion. If the input stage provided sufficient amplification, the following stages process a filtered error signal. Their noise and nonlinearity are less relevant due to the high amplification of the first stage. This allows a low-bandwidth low-power design for these filter sections. The topology of a CIFF filter will show a zero term in the STF transfer. This term can lead to some peaking behavior in the STF transfer leading to reduced signal ranges.

The filter function is shaped by weighted addition of the outputs of the successive integrators. The 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 results in the feed-forward topology and is discussed in Sect. 9.8.3.

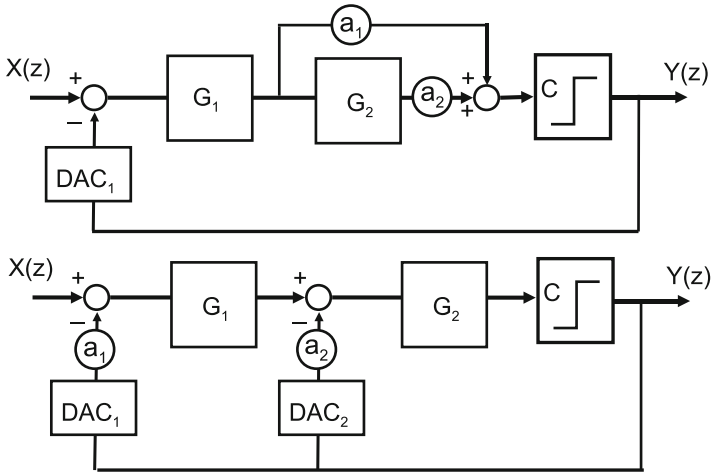


Fig. 9.21 Two methods to extend a first-order sigma–delta modulator: the cascade of integrators feed-forward and feedback topologies

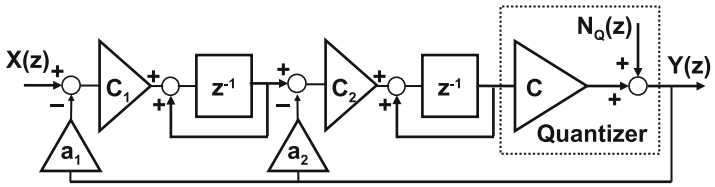


Fig. 9.22 A second-order sigma–delta modulator with two feed-back paths. The weighting coefficients are in the feedback path

The same filter function can also be realized by a feedback topology (CIFB). Now two digital-to-analog converters are needed, one for every pole in the filter. The subtraction scheme allows more freedom to choose various signal levels. The feedback factor is not exactly unity, which affects the STF. Some peaking in the STF can occur.

In the CIFB topology each feedback path requires to design for maximum performance. Moreover the subtraction of the feedback signal is not completely concentrated in the first stage, that therefore has to amplify the input signal at maximum linearity. Scaling of the signal is needed to optimally utilize the available signal range.

9.4.3 Second-Order Modulator

Figure 9.22 shows the feedback extension for a time-discrete second-order topology. The feedback path coefficients are a_1 and a_2. If a_2 is too small, the first-order

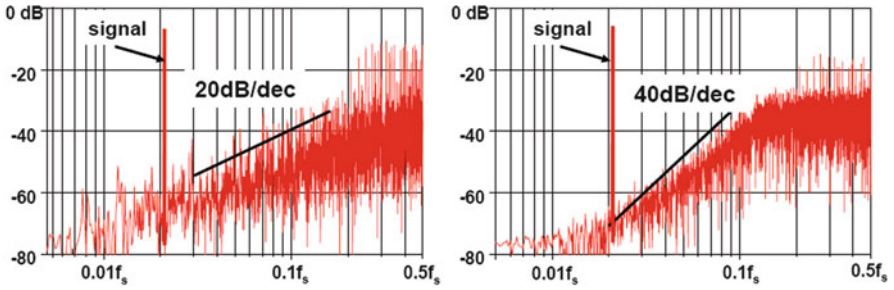


Fig. 9.23 The spectra of a first- and second-order sigma–delta modulator

content is insufficient to keep the modulator free of instability. If a_2 is too large, the modulator will behave as a first-order system. The coefficients c_1 and c_2 are chosen to scale the signal after the summation point back into the range. With $c_1 = c_2 = 1$ the transfer is

$$Y(z) = \frac{X(z)z^{-2} + N_Q(z)(1-z^{-1})^2}{a_1 + a_2 z^{-1}(1-z^{-2}) + (1-z^{-1})^2}.$$

Popular values are $a_1 = 1$ and $a_2 = 2$ resulting in a unity denominator.

For low frequencies the gain of the first integrator clearly surpasses the ratio a_2/a_1. At higher frequencies the first-order path starts to contribute. As the comparator also provides gain, the unity gain frequency of the filter will be chosen considerably lower than $f_s/2$.

The NTF in Fig. 9.22 is of the form

$$\mathrm{NTF}(z) = \frac{Y(z)}{N_Q(z)} \propto (1-z^{-1})^2$$

resulting in a behavior equivalent to a second-order noise shaper, with its noise power approximated by

$$\int_{f=0}^{f_b} \frac{V_{\mathrm{LSB}}^2}{6f_s} |1 - \mathrm{e}^{-2j\pi f T_s}|^4 \mathrm{d}f \approx \frac{V_{\mathrm{LSB}}^2}{12} \frac{\pi^2}{5} \left(\frac{2f_b}{f_s}\right)^5 \tag{9.31}$$

which is a similar description as was found for a second-order noise shaper (see Fig. 9.9).

The output spectrum second-order sigma–delta modulator is compared to a first-order modulator in Fig. 9.23. 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/decade slope and much less idle tones. The dual integration of the signal and the two feedback paths from the quantizer create a much more complex pattern, which appears to contain less correlated

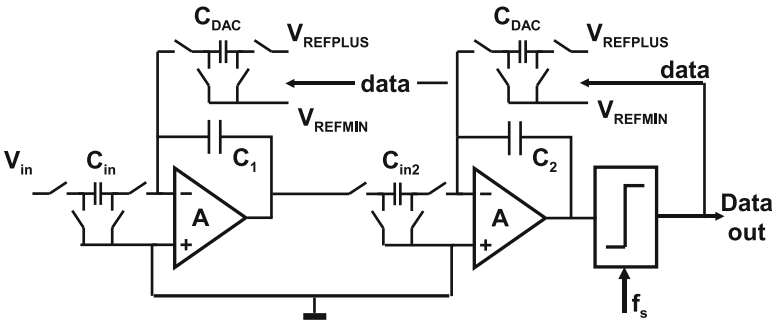


Fig. 9.24 A switched-capacitor realization of a second-order sigma–delta modulator

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. 9.13). 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 9.24 shows the realization of a second-order switched-capacitor sigma–delta modulator. The input network samples the input voltage and transfers this 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.

9.4.4 Circuit Design Considerations

The analog summation point in the sigma–delta analog-to-digital converter must process high-frequency signal 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 delayed 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 is acceptable for reaching the performance.

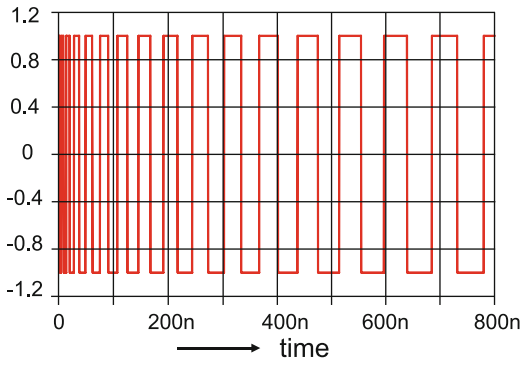


Fig. 9.25 The output signal of a second-order sigma–delta modulator with a single feedback via two ideal integrators and zero volt input signal

Designing the input with a sufficiently low distortion requires to spend a lot of power.

Often some form of scaling is needed to reduce the signals to levels that can be handled by the integrators [261]. Another approach is to add gain before the quantizer and to use a scaled digital-to-analog conversion output signal. The scaling operation sacrifices useful dynamic range. The thermal noise contribution of the analog circuits becomes relatively more important with the risk of losing signal-to-noise ratio. Without scaling, some topologies (prominently CIFB structures) are limited by nonlinearities and overload conditions. Architectures, such as the feed-forward topology in Sect. 9.8.3, limit inherently the internal voltages and are less prone to nonlinearity.

In sigma–delta converters with very high sample rate, the comparator will be playing a more important role. A comparator, just as in Nyquist rate conversion, should generate an instantaneous decision. However, a delay period can arise or even a nondecision, as described in Sect. 8.1.4. The delay will cause phase-margin impairment in the loop and is in high sample-rate designs compensated by an excess-delay path (see Sect. 9.5.3).

Metastability in the comparator generates a noise-like component and can either be taken into the design as a parameter [262] or must be reduced along the same line as in Sect. 8.1.4.

Example 9.4. In a second-order time-discrete sigma–delta modulator the first-order feedback path is removed $a_2 = 0$. Draw the output signal.

Solution. Without a first-order path the feedback signal will be delayed by two times 90°. Any additional delay will cause the feedback to be too late. So an edge generated by the comparator will be processed one clock period later. For zero-input signal a $f_s/2$ waveform is expected, but now successive edges will be processed with increasingly larger intermediate delays (see Fig. 9.25 for this typical form of sigma–delta instability).

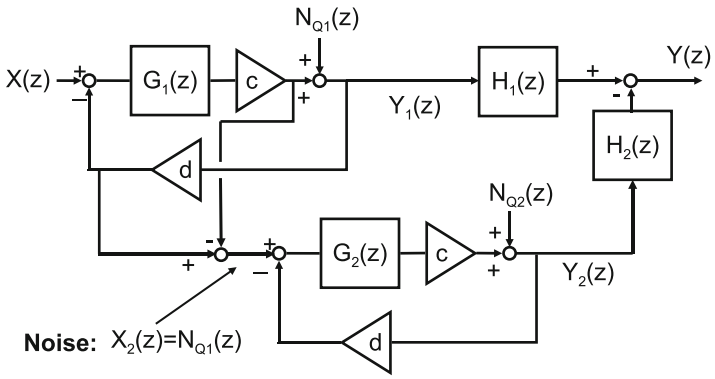


Fig. 9.26 The cascaded sigma–delta modulator or multi-stage noise shaper (MASH)

9.4.5 Cascaded Sigma–Delta Modulator

Extending the time-discrete sigma–delta converter to a third- or higher-order is possible [263]; however, the loop filter has to be designed in such a way that a stable oscillation is guaranteed [264, 265]. This can be done by inserting zeros in the transfer function. Yet, stability problems may occur. The gain of the quantizer in this analysis is assumed to be linear. In reality, the quantizer gain is a nonlinear function and varies with the input amplitude. An indication (but not more than that) for stability of 1-bit sigma delta converters is based on simulations and called Lee's rule: $|\mathrm{NTF}(f)| < 1.5 - 2$ for all frequencies.

More noise power can be suppressed in a fixed bandwidth either by increasing the oversampling rate or by designing a higher-order loop filter without stability problems. The cascaded sigma–delta approach [266] solves this issue in a third way: multistage noise shaping (MASH) (see Fig. 9.26). Cascaded sigma–delta converters allow higher-order noise shaping without higher-order loop filters. 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 Eq. 9.22. Although this sigma–delta loop suppresses the quantization errors $N_{Q1}(z)$, this noise can be further suppressed by considering that this quantization error is a deterministic distortion signal that can be measured and subtracted from the output (see Fig. 9.26). The noise of the first sigma–delta converter is isolated by creating an analog copy of the output signal (via the digital-to-analog converter "d"). This signal is then subtracted from 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 the same way as $N_{Q1}(z)$ is shaped in by the loop of the first sigma–delta modulator. $Y_2(z)$ has to be divided by the loop 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 ($d = 1$)

$$\begin{aligned} Y_1(z) &= \frac{G_1(z)}{1+G_1(z)}X(z) + \frac{1}{1+G_1(z)}N_{Q1}(z) \\ Y_2(z) &= \frac{G_2(z)}{1+G_2(z)}N_{Q1}(z) + \frac{1}{1+G_2(z)}N_{Q2}(z) \\ Y(z) &= \frac{H_1(z)G_1(z)}{1+G_1(z)}X(z) + \left(\frac{H_1(z)}{1+G_1(z)} - \frac{H_2(z)G_2(z)}{1+G_2(z)}\right)N_{Q1}(z) \\ &\quad - \frac{H_2(z)}{1+G_2(z)}N_{Q2}(z). \end{aligned} \tag{9.32}$$

With G_1, G_2 sufficiently large and $H_2(z)G_1(z) = H_1(z) = z^{-k}$:

$$Y(z) = X(z) + \frac{1}{G_1(z)G_2(z)}N_{Q2}(z). \tag{9.33}$$

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" or "2–1 MASH." 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 and high-quality results can be reached [267].

A time-continuous cascaded sigma–delta converter needs a calibration mechanism to trim the digital filter to the analog filter [268] or careful design [269]. Also in sigma–delta digital-to-analog conversion noise-subtraction techniques are used. An interesting example is found in [147].

Example 9.5. Determine a suitable order of the analog filter and OSR to obtain 100 dB SNR gain in the desired bandwidth.

Solution. With the help of Fig. 9.9 various combinations are found: 16 times oversampling requires a fifth or sixth order, while 256 times oversampling needs only a second-order filter.

9.5 Time-Continuous Sigma–Delta Modulation

9.5.1 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 a loop filter based on time-continuous components. The sampling takes place after the filter (see Fig. 9.27). 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 9.28 shows the waveforms for a 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

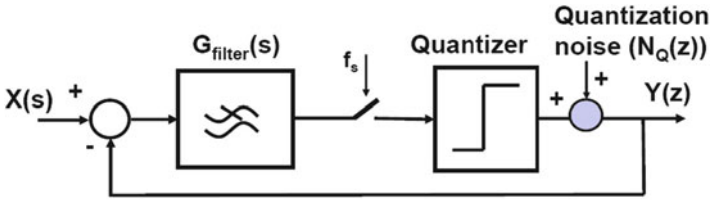


Fig. 9.27 A time-continuous sigma–delta modulator samples after the filter

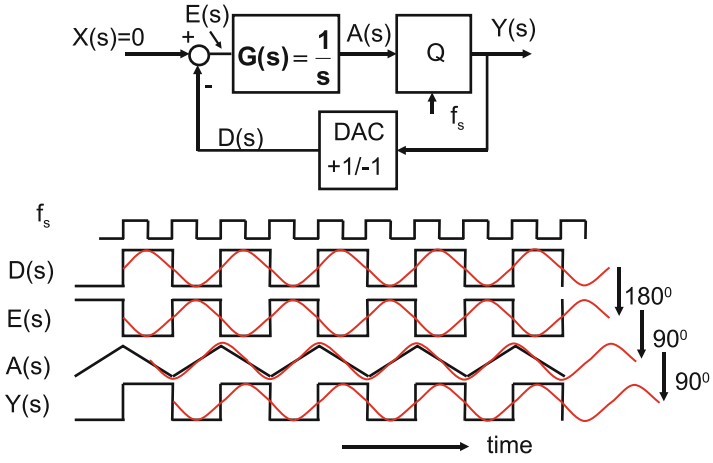


Fig. 9.28 The signal flow in a first order time-continuous sigma–delta modulator. The quantizer switches on the rising edge of f_s. The fundamental frequency for all block and *triangle-shaped* signals is indicated to see the phase relations

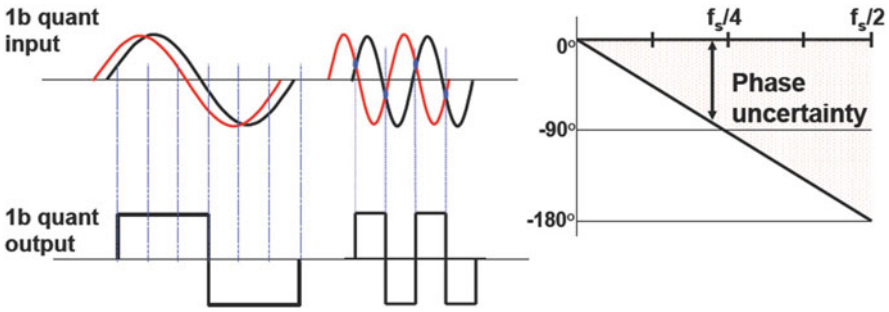


Fig. 9.29 Two examples of comparator input signals that result in the same comparator decisions. The waiting period before sampling translates in a phase range for each signal frequency

with minimum 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 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 an oscillating loop 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 will be early (see Fig. 9.29). 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.

In the presence of some DC offset at the input, Fig. 9.30, the waveforms in the time-continuous sigma–delta converter resemble the waveforms in the time-discrete case in Fig. 9.19. The integration of the high-frequency component and the input signal is now well visible.

The STF and a NTF are formulated in the s or Laplace domain:

$$\begin{aligned} Y(s) &= \mathrm{STF}(s)X(s) + \mathrm{NTF}(s)N_Q(s) \\ \mathrm{STF}(s) &= \frac{cG(s)}{1+cG(s)} \\ \mathrm{NTF}(s) &= \frac{1}{1+cG(s)}, \end{aligned} \tag{9.34}$$

where c is the gain of the comparator (see Fig. 9.12). The complication of this description is the accuracy of the modeling of the sampling and quantization after the time-continuous filter.

A second-order time-continuous sigma–delta modulator is created by adding an integration stage. The second-order filter will suppress the quantization errors better.

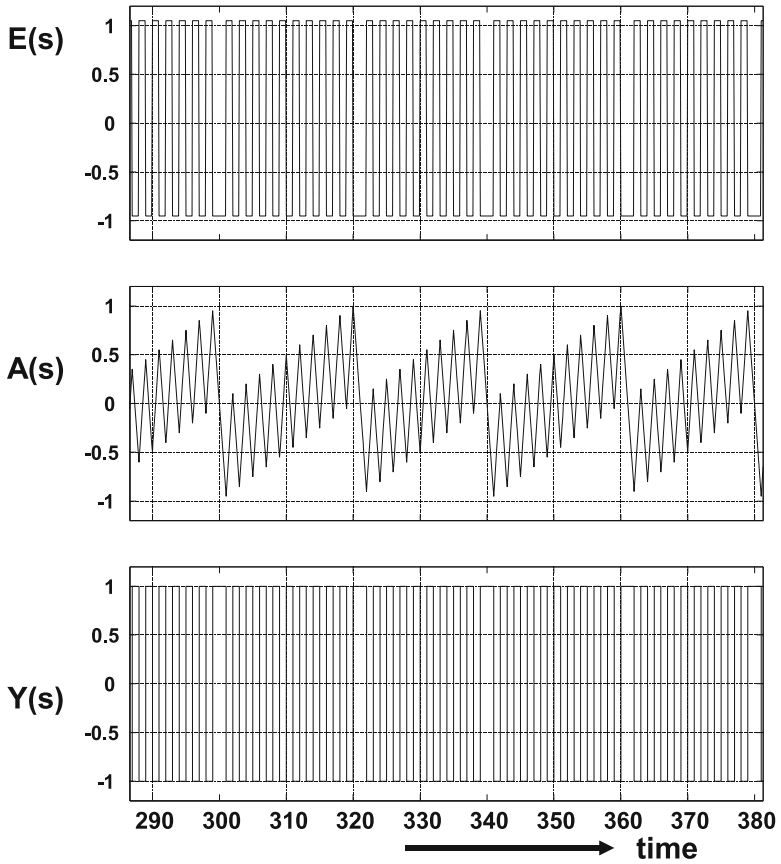


Fig. 9.30 The signals in a first-order time-continuous sigma–delta modulator with a small DC input signal

This becomes visible in less correlated error signals as can be seen from Fig. 9.31. This configuration in a time-continuous sigma–delta modulator has an advantage that (a part of) the alias filtering is performed by the loop filter.

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 zeros 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.

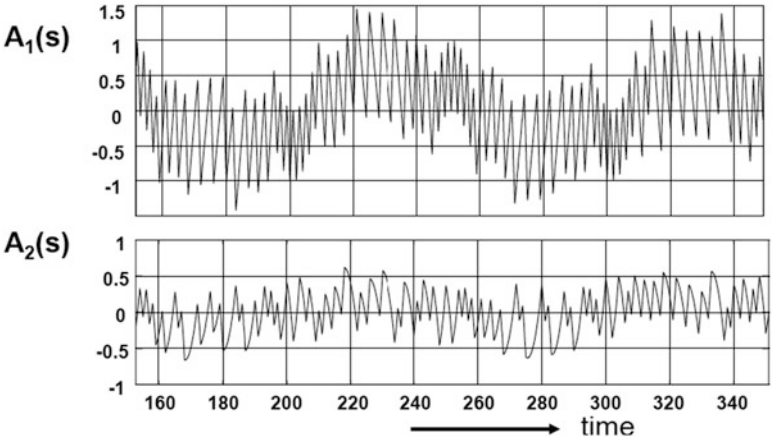


Fig. 9.31 A comparison of the signal after the filter in a first-order and a second-order sigma–delta modulator

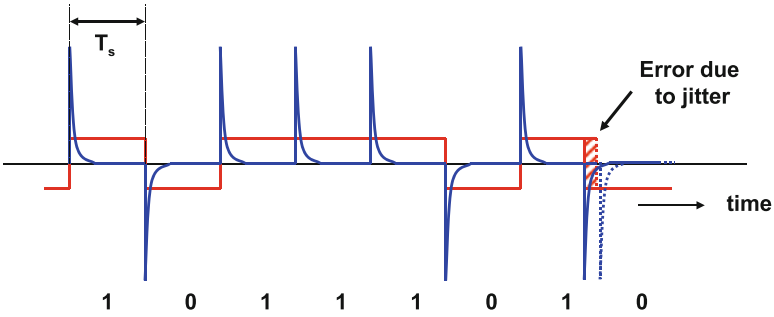


Fig. 9.32 A time-continuous DAC feedback signal is a continuous pulse. A switched-capacitor DAC generates charge spikes. Jitter just shifts the position of a switched-capacitor pulse, while it modulates the volume of a time-continuous pulse

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 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 sometimes 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. 9.32 and Sect. 7.5). Asymmetry in the feedback path, see Fig. 9.33, 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

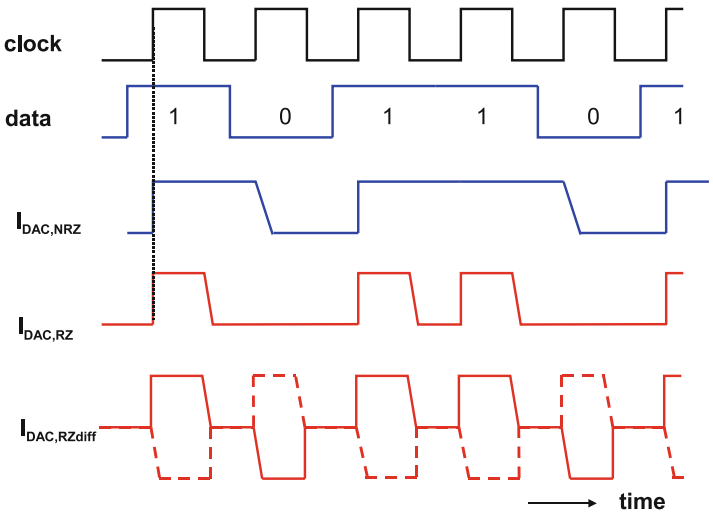


Fig. 9.33 Asymmetry in the digital-to-analog conversion pulses $I_{\mathrm{DAC,NRZ}}$ can be mitigated with a return-to-zero technique $I_{\mathrm{DAC,RZ}}$ or a differential return-to-zero technique $I_{\mathrm{DAC,RZdiff}}$

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 [270]. Also a charge-based digital-to-analog converter can be used.

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. 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. For achieving low-noise levels in wider bandwidths without increasing the sample rate into the GHz range, also in time-continuous sigma–delta converters the filter order must be increased.

Figure 9.34 shows a sigma–delta converter with a fourth-order feed-forward filter. 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.

The design of the filter function has to take into account that the requirement for stable oscillation in the loop requires a feedback path with a total phase shift before the quantizer of less than 180° at the oscillation frequency (close to $f_s/2$). Therefore, the design of higher-order filters is possible if the filter is designed in such a way that it shows a high-order behavior for low frequencies, while turning

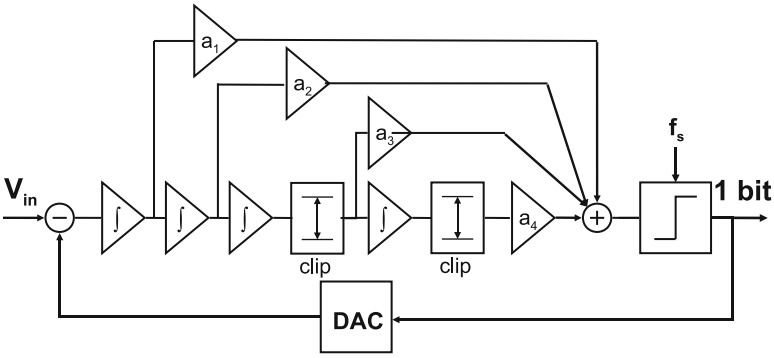


Fig. 9.34 A fourth-order feed-forward sigma–delta modulator [271]

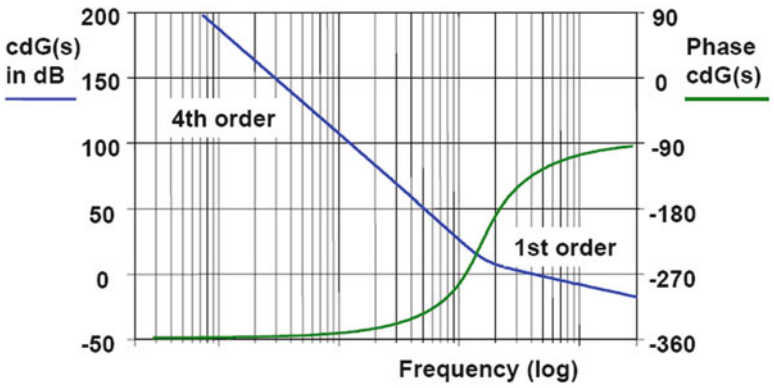


Fig. 9.35 The fourth-order filter transfer function turns back to a first-order transfer function close to the frequency region where the modulator acts as an oscillator

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 360° overall phase shift is reached: 180° from the inversion at the summation node and >90° from the filter and the rest is consumed in the delay before the sampling moment, the phase uncertainty.

In case of instability due to an excess voltage at the input, the outputs of the last two integrators may saturate. The clippers in Fig. 9.35 effectively switch off the contribution of these integrators to the overall filter function. The modulator will behave as a third- or second-order sigma–delta converter. Although the performance is temporarily degraded, still some signal conversion is performed. This behavior is called "graceful degradation." After removal of the excess signal the integrator will recover and help again to shape the filter function.

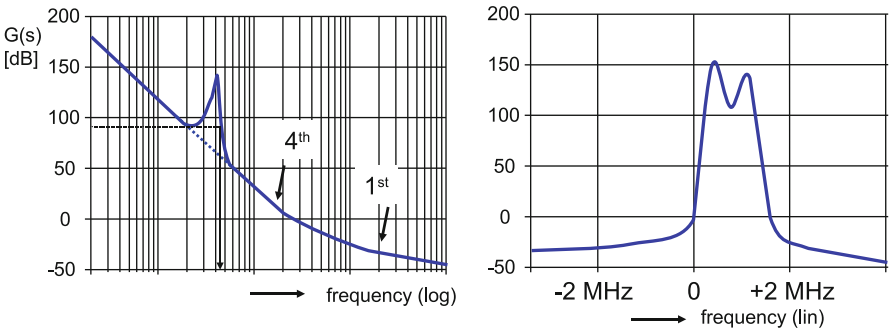


Fig. 9.36 Resonators and other filter elements are used to modify the loop filter curve. On the *left* side a pole in the filter increases the gain in a limited frequency range, thereby stretching the noise suppression. *Right*: a complex filter allows to create suppression in the positive frequency range [268]

The shape of the filters in a time-continuous sigma–delta converter depends on the ratios of the frequency determining elements. The absolute 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.

Many techniques known from analog filter design have been applied to modify the noise transfer characteristics. Inserting resonators, see Fig. 2.65, allows to stretch the noise suppression frequency range in Fig. 9.36(left). More poles in the filter on different frequencies can be applied to shape the noise transfer. Poles in the loop filter turn into zeros in the NTF.

Quadrature signals are very common in advanced communication. Complex sigma–delta modulators, Sect. 9.8.1, use two parallel modulators driven by in-phase and quadrature signals. The filters of these two modulators are coupled to perform complex filtering. Figure 9.36(right) shows a filter characteristic that suppresses only for positive frequency ranges [268].

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 nonlinear, the description of the gain is not a simple mathematical expression. 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 $T \gg T_s$:

$$(cd)^2 = \frac{\int v^2_{\text{quantizer,out}} \mathrm{d}t}{\int v^2_{\text{quantizer,in}} \mathrm{d}t} = \frac{(\pm 1)^2}{(1/T) \int\limits_T v^2_{\text{quantizer,in}} \mathrm{d}t}. \tag{9.35}$$

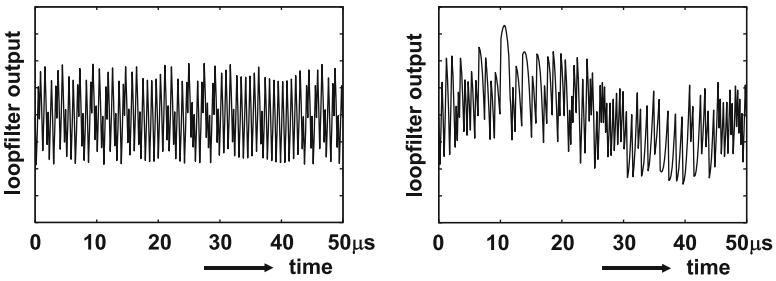


Fig. 9.37 The simulated input for the quantizer, *left* with 0.1% signal input to the modulator, *right*: with maximum input signal, from [271]

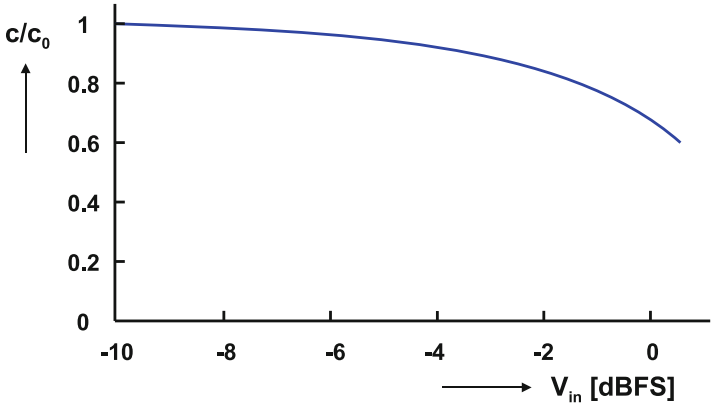


Fig. 9.38 The gain as function of the input signal to the modulator, from [271]

It can be necessary to do a numerical comparison of the input power with the rather simple output power of the comparator. In a simulation example the gain does not vary much over the input amplitude range (see Figs. 9.37 and 9.38). 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.

Figure 9.39 shows the resulting design of a fourth-order loop filter, including some guard-banding for the gain inaccuracy. Also the phase uncertainty region is indicated. This region indicates where the phase must be in order to be in-time for sampling. The oscillation will chose a frequency 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 nonlinear circuit to verify proper functionality for all input levels.

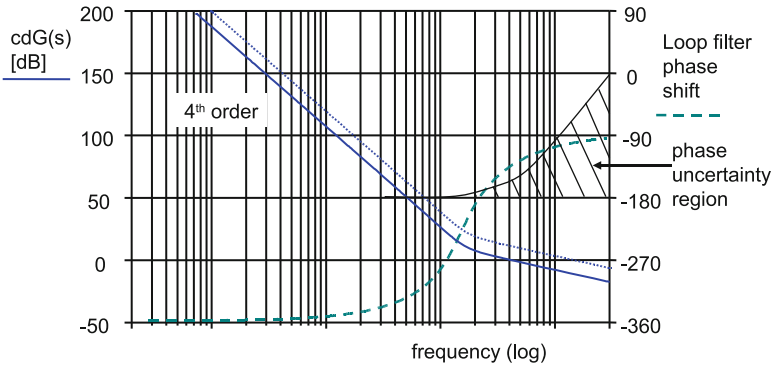


Fig. 9.39 The fourth-order filter transfer function with gain tolerance and the phase plot including phase uncertainty

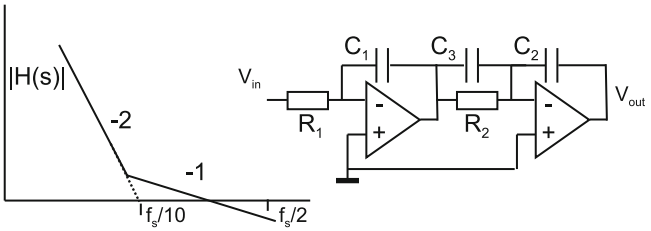


Fig. 9.40 Filter and transfer function for a second-order time-continuous sigma–delta converter

Example 9.6. 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 digital-to-analog converter can be assumed ideal. The gain uncertainty and the sampling uncertainty of the comparator need to be taken into account. Setup a second-order filter using capacitors, resistors, and ideal opamps that gives a maximum SNR ratio in this band.

Solution. The oversample ratio is 2.5 MHz/20 kHz = 125. With an ideal second-order 1-bit sigma delta modulator this would give 6 dB from the 1-bit quantizer and 92 dB shaping improvement (see Fig. 9.9). A second-order integrator can be designed as a cascade of two opamp-*R*-*C* sections (Fig. 9.40). However, the loop would be instable. Therefore an additional capacitor C_3 is added. So a component choice $R_1 = R_2 = 100\,\text{k}\Omega$ and $C_1 = C_2 = 3\,\text{pF}$ gives a unity gain for the two integrators of 500 kHz. $C_3 = 4\,\text{pF}$ will create a zero a factor 3–4 before the unity-gain frequency. At $f_s/2$ the overall filter gain will be around $-20\,\text{dB}$. The gain of the comparator has to compensate this attenuation. The filter will amplify 55 dB at 20 kHz, which is a loss of about 30 dB compared to the ideal situation. Figure 9.41 shows a simulation with this filter in a sigma–delta loop.

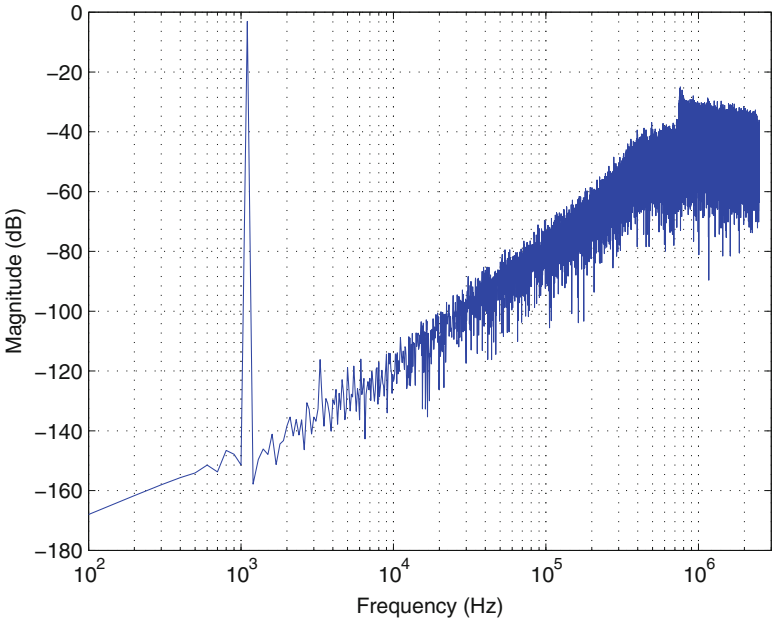


Fig. 9.41 Conversion of a 1.1 kHz input signal (courtesy: Hugo Westerveld UT)

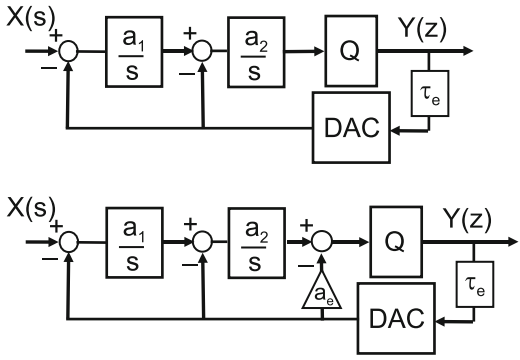


Fig. 9.42 Delay in the quantizer and the digital-to-analog converter causes excess loop delay. An additional feed-back path compensates for this delay

9.5.3 Excess Loop Delay in Time-Continuous Sigma–Delta Conversion

A major problem in time-continuous sigma–delta conversion is the delay from the sample pulse to the actual feedback signal. This excess loop delay is caused by comparator delays and the digital-to-analog converter. In Fig. 9.42(top) an ideal second-order sigma–delta converter with loop delay is shown. Assuming that the

comparator does not influence the path, the ideal transfer function is described with the coefficients $\mathbf{a}_1$ and $\mathbf{a}_2$:

$$\frac{Y(s)}{X(s)} = \frac{\mathbf{a}_1\mathbf{a}_2}{s^2 + \mathbf{a}_2 s + \mathbf{a}_1\mathbf{a}_2}$$

The excess delay is modeled by the function $f_{\tau e}(s) = \mathrm{e}^{-s\tau_e}$. This function is approximated with the help of a Taylor series Table 2.8: $\mathrm{e}^{-s\tau_e} \approx 1 - s\tau_e + s^2\tau_e^2/2$. Now the transfer function is

$$\frac{Y(s)}{X(s)} = \frac{a_1 a_2}{s^2 + a_2 s f_{\tau e}(s) + a_1 a_2 f_{\tau e}(s)} \approx \frac{a_1 a_2}{s^2(1 - a_2\tau_e + a_1 a_2 \tau_e^2/2) + s(a_2 - a_1 a_2 \tau_e) + a_1 a_2}.$$

The terms for s^3, s^4 are neglected as their coefficients are small. If the coefficient $a_2 = \pi f_s$ (the unity gain frequency of the second integrator is close to half of the sample rate) and the excess delay is 1/6 of a sample period, then $(1 - a_2\tau_e) \approx 0.5$. It is obvious that the transfer function is not accurately implemented.

Figure 9.42(bottom) shows the compensation scheme: an additional path weighted with a_e is added to the transfer

$$\frac{Y(s)}{X(s)} \approx \frac{a_1 a_2}{s^2(1 - a_2\tau_e + a_1 a_2 \tau_e^2/2 + a_e) + s(a_2 - a_1 a_2 \tau_e) + a_1 a_2}.$$

This transfer function approximates the ideal second-order function if all terms match the terms of the ideal transfer. Equations $1 - a_2\tau_e + a_1 a_2 \tau_e^2/2 + a_e = 1$, $(a_2 - a_1 a_2 \tau_e) = \mathbf{a}_2$, and $a_1 a_2 = \mathbf{a}_1\mathbf{a}_2$ yield

$$\begin{aligned} a_1 &= \frac{\mathbf{a}_1}{1 + \mathbf{a}_1\tau_e} \\ a_2 &= \mathbf{a}_2(1 + \mathbf{a}_1\tau_e) \\ a_e &= \mathbf{a}_2\tau_e(1 + \mathbf{a}_1\tau_e/2). \end{aligned}$$

For very fast time-continuous sigma–delta modulation this compensation is absolutely necessary; more analyses are found in [272].

9.5.4 Latency

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.

9.6 Time-Discrete and Time-Continuous Sigma–Delta Conversion

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. Figure 9.43 summarizes a number of architectural choices and consequences for the designer of sigma–delta converters. 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 (see Table 9.2).

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 zeros 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 et al. [268] shows 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.

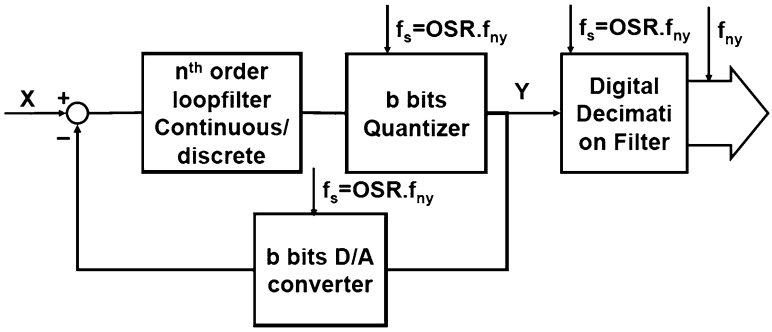


Fig. 9.43 Several design choices play a role during the implementation of a sigma–delta converter: the oversampling factor OSR, the levels in the quantizer, the type and order of filter

Table 9.2 Comparison of time-discrete and time-continuous sigma–delta modulation

	Time discrete	Time continuous
Filter arrangement	Often CIFB to stabilize the second order	Often CIFF to reduce demands on second- and higher-integrator stages
Relative positions of poles and zeros in filter	Defined by capacitor ratios	Matching of capacitor ratios and ratios between resistive elements
Absolute value of filter parameters	Capacitor ratios and sample rate	Filter curve modulated by absolute spread of parameters
Relative and absolute accuracy of poles and zeros	$<0.1\%$, 0.1%	0.5%, 15%
Single loop	Practical implementations are second order	Up to fifth order
Higher-order noise shaping	Cascade of lower order sigma–delta modulators	Increase filter order
Linearity	Summation nodes and integrators	Mainly in first integrator
Anti-alias filter	Extra time-continuous filter required	Partly or completely by loop filter
Out-of-band signals	Can be modulated back	Suppressed by loop filter
Opamp constraint	Runs at sample rate	Determined by filter
Power	In fast opamps	In linearity and noise of first integrator

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. 3.9). 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 [273, 274]. 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 OSR 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 counterpart 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 set 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 7.30 shows the impact of jitter on pulse-width modulated signals. This impact of jitter in a time-continuous converter is its 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 [268, 275] or apply a FIR digital-to-analog converter to limit jitter influence [276]. Figure 9.44 compares the time-continuous and switched-capacitor sigma–delta converters that have been published on the ISSCC. Switched-capacitor sigma–delta modulators are more present in the high-resolution region, while time-continuous converters push forward towards higher bandwidths. In Fig. 12.10 the comparison is extended with other converter architectures. It is obvious that in the last years time-continuous sigma–delta conversion enters the domain of pipeline conversion [277, 278].

Example 9.7. Discuss the various noise sources in a sigma–delta converter.

Solution. For many applications quantization distortion/noise can be strongly reduced by a proper choice of oversample ratio and filter order.

Jitter and thermal noise are the main contributors. Jitter is especially dominant in fast 1-bit feed-back loops. It must be minimized by proper design of the timing path and a high-quality sample-rate source. In many architectures thermal noise is

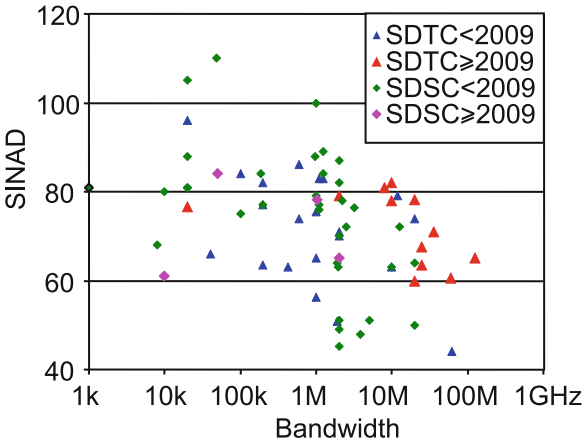


Fig. 9.44 Bandwidth and SINAD comparison for time-continuous and switched-capacitor sigma–delta converters as published on the International Solid-State Circuits conference (from: B. Murmann, "ADC Performance Survey 1997–2012," [Online]. Available: http://www.stanford.edu/~murmann/adcsurvey.html)

most critical at the first summation point. Linearity demands do not allow excessive gains, and the dynamic range has to be sufficient to accommodate the sum of input and feedback signals.

Very low bandwidth converters may suffer from a significant $1/f$ contribution. Chopping the bias paths will largely remove this term.

Finally some contributions from metastability in the comparator and residual components from digital-to-analog converters with data-weighted averaging can contribute to the total noise budget.

The initial design of a typical SNDR $= 80\,\mathrm{dB}$ sigma–delta converter running at several tenths of MHz bandwidth will allocate the noise contributions roughly in three parts: 30% for jitter and the same for thermal noise and the rest for the other components.

9.7 Multi-bit Sigma–Delta Conversion

In sigma–delta conversion the signal-to-noise improvement by oversampling and subsequent noise shaping is huge. The signal-to-noise gain obtained by a multi-bit quantizer has from that perspective 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. This last argument enabled designers to improve harmonic distortion performance beyond the levels set by multi-bit quantization with their inherent nonlinearities. Yet multi-bit quantizers with output waveforms as in Fig. 9.45 enjoy an increasing popularity [268, 279]. 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:

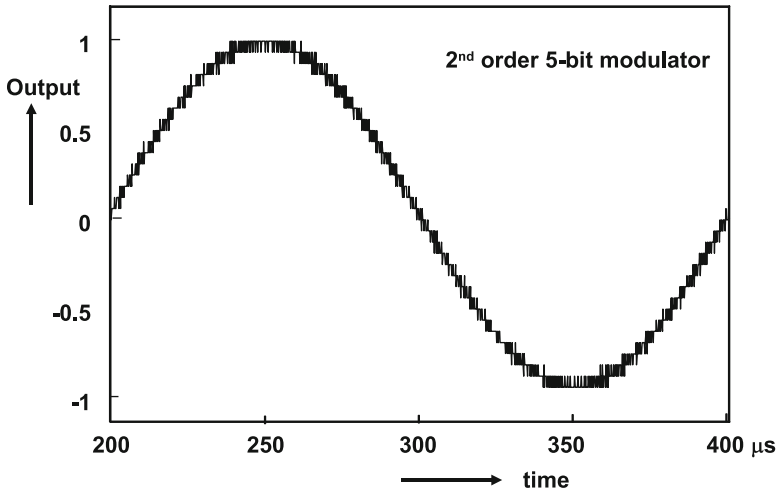


Fig. 9.45 The reconstructed output wave form of a 5-bit quantizer in a sigma–delta modulator

Table 9.3 Simulated noise level reduction for multi-level quantization

Resolution	ΔSNR (dB)
1→2	7.0
2→3	6.2
3→4	6.1
4→5, etc.	6.0

- A single-bit quantizer generates a lot of noise in the loop. This noise has to be treated in a linear way because nonlinearities will cause mixing products in the desired band. The lower noise level of a multilevel 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 9.3 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 multilevel digital-to-analog converter in the feedback loop.
- The swing of the residue signal in the first integrator of a multilevel quantizer is reduced, so a larger integrator gain is possible reducing distortion.
- A multilevel quantizer requires more comparators, more digital hardware, and a more extensive digital-to-analog converter, which translates in more power.

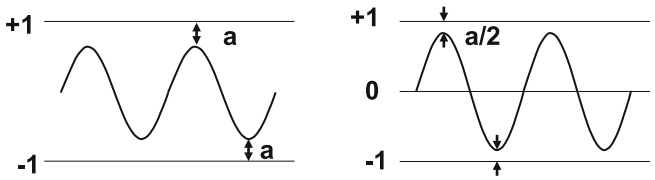


Fig. 9.46 Signals close to overload in a 2-level and 3-level quantizer

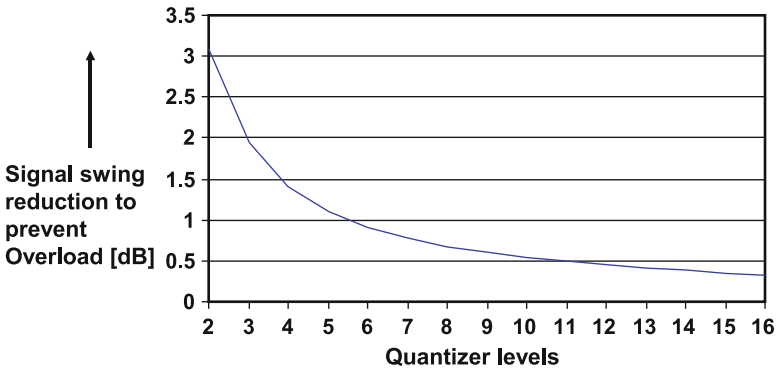


Fig. 9.47 The required signal reduction to prevent overload reduces with increasing number of quantization levels

- The gain factor in a single-bit quantizer has to be determined by comparing the input and output power. In a multilevel 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 multilevel quantizer.
- The overload level of a multilevel modulator improves [280] and thereby increases the signal-to-noise ratio. In Fig. 9.46 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. van Veldhoven et al. [280] assume that in a multilevel 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 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 9.47 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 2-bit quantizer and a 2-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 multilevel 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 [153], see Sect. 7.7.3, 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 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.

A provocative (yet instructive) case against 1-bit sigma–delta is made in [281].

9.8 Various Forms of Sigma–Delta Modulation

9.8.1 Complex Sigma–Delta Modulation

In digital communication systems signals are modulated with an in-phase carrier and its 90° shifted quadrature component. These I/Q signals are converted to the digital domain with a complex sigma–delta converter (Fig. 9.48). 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 NTF. These properties are beneficial for designing an optimum conversion for various communication systems.

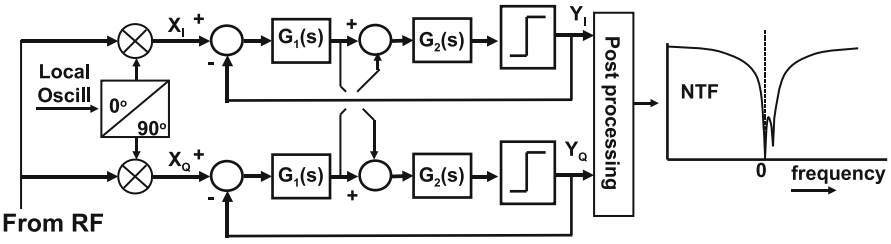


Fig. 9.48 The complex sigma–delta converter consists of two coupled sigma–delta modulators, e.g., [268, 282]

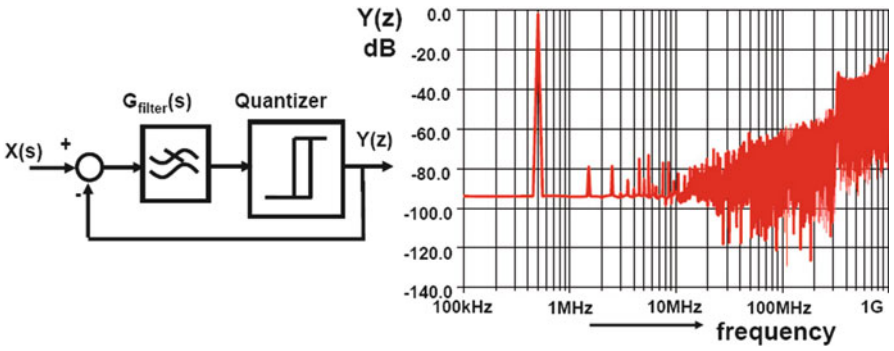


Fig. 9.49 The asynchronous sigma–delta converter and output spectrum

9.8.2 Asynchronous Sigma–Delta Modulation

The asynchronous sigma–delta converter [283] in Fig. 9.49 uses the same topology as the time-continuous converter with the exception of the quantizer. The clocked quantizer is replaced by an non-clocked nonlinear element, such as a continuous comparator, or a comparator with some hysteresis. As there is initial no quantization involved, there is no quantization energy created and neither to be shaped. The loop operates as a kind of multivibrator on a self-oscillation frequency. This 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. 9.49(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).

9.8.3 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

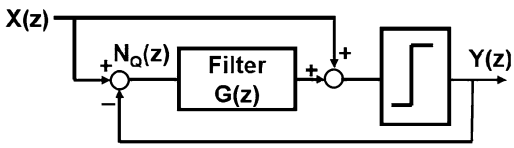


Fig. 9.50 A feed-forward sigma–delta modulator

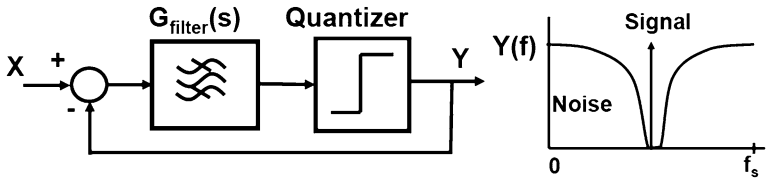


Fig. 9.51 A basic set-up for a band-pass sigma–delta modulator

input feed-forward architecture as shown in Fig. 9.50. The input signal is fed into the chain just before the quantizer. This leads to a transfer function ($c = d = 1$):

$$Y(z) = X(z) + \frac{N_Q(z)}{G(z)+1}. \tag{9.37}$$

So STF = 1. 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 [284] with BW = 25 kHz, SNDR = 95 dB, and $P = 0.87$ mW.

9.8.4 Band-Pass Sigma–Delta Converter

The dominant direction in the implementation of RF communications systems is the shift of analog external components into the converter. As a result designers are looking for methods to integrate more components of the RF-to-digital interface at the intermediate frequency (IF) level of a superheterodyne receiver architecture. The antenna 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. 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 directly convert the narrow-band signal at the IF frequency by means of band-pass sigma delta conversion (Fig. 9.51). Instead of a low-pass filter a band-pass filter will clear the IF frequency band from quantization noise. Engelen et al. [285] show a sixth-order

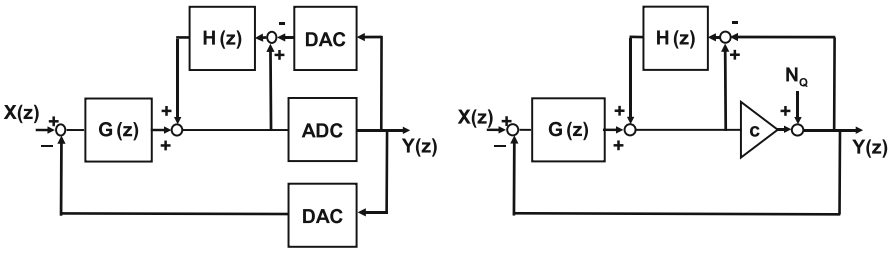


Fig. 9.52 A sigma–delta converter with embedded noise shaper. *Left* is the design and *right* the equivalent scheme [255]

sigma–delta analog-to-digital converter with an FM-radio IF of 10.7 MHz, BW = 200 kHz, ENOB = 10.2, and $P = 60$ mW, and Barkin et al. [147] implement a digital-to-analog converter at one quarter of the sample rate. A television band-pass converter is shown in [286] and [287] uses a sixth-order subsampling bandpass converter for the 2.4 GHz RF band. Harrison et al. [288] implement the band pass with an LC tank, while [289] uses resonator-type filters and compares recent designs. On system level there is ambiguity in this concept. The construction of the band-pass filter is tedious, and running this filter at high frequencies costs power and area. In the RF 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 [270, 282, 290–292] provide an alternative. van der Zwan et al. [290] show a sigma–delta analog-to-digital converter with an IF of 10.7 MHz, BW = 200 kHz, and carrier noise = 79 dB with a power of 11 mW.

9.8.5 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 OSR to allow the transfer to 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. 9.52(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

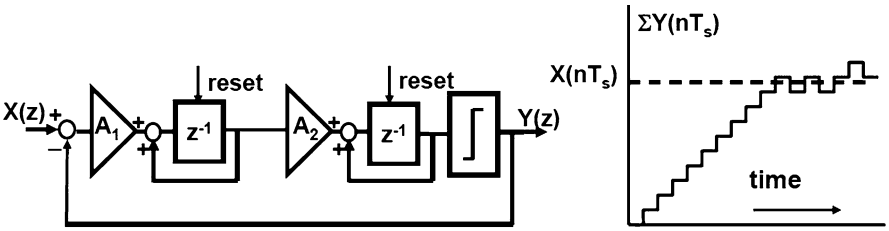


Fig. 9.53 A second-order incremental sigma–delta modulator

noise shaper serves as a modified quantizer in the overall sigma–delta loop. With the help of the equivalent scheme in Fig. 9.52(right) the transfer is calculated as

$$\begin{aligned} Y(z) &= \frac{cG}{(1+cG)+H(c-1)}X(z)+\frac{1-H}{(1+cG)+H(c-1)}N(z) \\ &\approx X(z)+\frac{1-H}{G}N(z). \end{aligned} \tag{9.38}$$

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 setup allows to gain some additional noise suppression without the disadvantages of the other solutions.

9.8.6 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 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 [293] (see Fig. 9.53). 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 $x(i) = x_{in}\ i = 0 \cdots n$ 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:

$$\begin{aligned} &\sum_{i=0}^{n-1}(x_{in}-y(i)) \\ &\sum_{m=0}^{n-1}\sum_{i=0}^{m-1}(x_{in}-y(i)), \end{aligned} \tag{9.39}$$

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

$$\frac{1}{n}\sum_{i=0}^{n-1} y(i) \approx x_{\text{in}}. \tag{9.40}$$

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 is 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

$$\frac{2}{n(n-1)}\sum_{m=0}^{n-1}\sum_{i=0}^{m-1} y(i) \approx x_{\text{in}}. \tag{9.41}$$

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 $n = 300\cdots600$ a 16-bit resolution is possible.

Exercises

9.1. Compare the heater-thermostat-living room system to a sigma–delta modulator.

9.2. A 6-bit analog-to-digital converter has to convert a signal bandwidth from DC to 50 kHz, but the converter can run up to 200 Ms/s. How much resolution can be obtained? A dither source is available. Can it be used to improve the accuracy?

9.3. A DC voltage is applied to a first-order 1-bit sigma–delta modulator running at 1 MHz. What output frequencies are generated for voltage values of $0\%, 10\%, \ldots, 100\%$ of the range. What changes if the quantizer is replaced by a 3-bit analog-to-digital converter?

9.4. A first-order sigma–delta converter is extended with an integration stage in the feedback path. Draw the STF and the NTF. Is this now a second-order sigma–delta modulator?

9.5. An oversample ratio of 64 is available to reach 110 dB SNR gain. How many first-order cascade stages are needed to achieve this result?

9.6. The digital-to-analog converter in the feedback path shows a 0.1% distortion component. Sketch or simulate the effect.

9.7. The second integrator of a second-order signal-delta modulator shows a 0.1% distortion component. Sketch or simulate the effect.

9.8. A time-continuous sigma–delta modulator runs at a sample rate of 100 Ms/s. The filter has single poles at 30 and 60 kHz and a zero at 10 MHz. Calculate the NTF and the STF.

9.9. Give an expression for the NTF and STF in a time-discrete sigma–delta modulator running at 100 Ms/s and with poles at 30 and 60 kHz.

9.10. A 1–1 cascaded sigma–delta modulator converts a bandwidth of 10 kHz with an oversampling factor of 512. There is a 2% mismatch between the gain of the analog filter in the first loop and the digital inverse filter. What will happen?

9.11. A first-order sigma–delta modulator runs at a sampling rate of 10 Ms/s. The feedback signal switches between 0 and 1 V. What is the output frequency if the input equals 0.5, 0.6, and 0.9 V?

9.12. A first-order sigma–delta modulator is part of a 1–1 cascade sigma–delta modulator. Its time-discrete filter is described by $z^{-1}/(1-z^{-1})$. Give a time-discrete description for the digital compensation filter.

9.13. Is it possible to multiplex two identical first-order sigma–delta modulators that run on a sample-rate clock and a half-period delayed sample rate.

9.14. A third-order time-continuous sigma–delta converter is used to convert a base band of 2 MHz with an oversample ratio of 64. At what level and frequency will a 255 MHz interferer signal of equal strength as the input signal show at the output? What happens if this converter is changed into a time-discrete implementation?

9.15. What is the maximum SNR that can be obtained in a 1 MHz bandwidth that is 128× oversampled if the feedback DAC clock jitters 2 $\mathrm{ps_{rms}}$?

9.16. A first-order noise shaper is cascaded with a second first-order noise shaper. Repeat the derivation in Eq. 9.32 for this situation and define the digital filter.

9.17. A third-order time-continuous 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 need to be taken into account. Set up a third-order filter using capacitors, resistors, and ideal opamps that gives a maximum SNR ratio in this band.

Chapter 10
Characterization and Specification

Abstract The specification of the converter is a dominant mechanism to align the wishes of the user to the possibilities of the designer. Directly coupled to the specification is the measurement technique that serves to establish a numerical value for a theoretical concept. This chapter discusses the fundamentals of the characterization and measurement techniques, such as histogram testing.

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 [107, 294, 295]. 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), and input impedance. The correct application of a converter depends crucially on a good characterization, test setup, and parameter extraction.

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_10,

© Springer Science+Business Media, LLC 2013

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

10.1 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. 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.
- 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. 10.2 a technique is shown where the surface-mounted device is pushed onto the connection electrodes of the printed-circuit card. This setup 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.

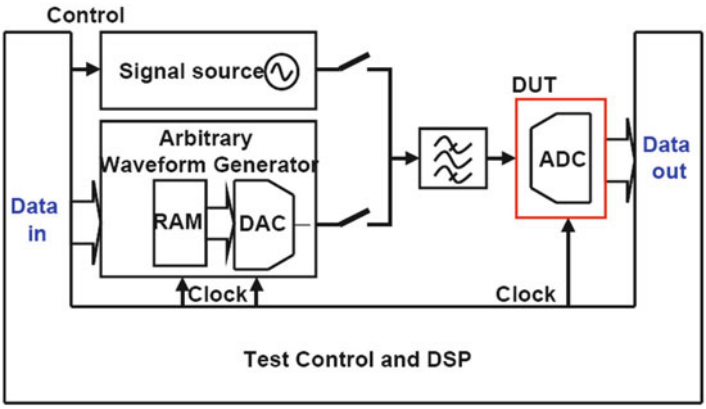


Fig. 10.1 Test set-up for digital-to-analog conversion

- 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 setup for characterizing an analog-digital converter uses a computer to control the setup and to analyze the measurement results; see also the relevant IEEE standardization documents [106–108]. 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, setup 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 DUT. 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. 10.1. This setup avoids that remaining distortion components, noise of various origins as well as cross talk of the generators internal processing, disturb the measurement. 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. 10.2. 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

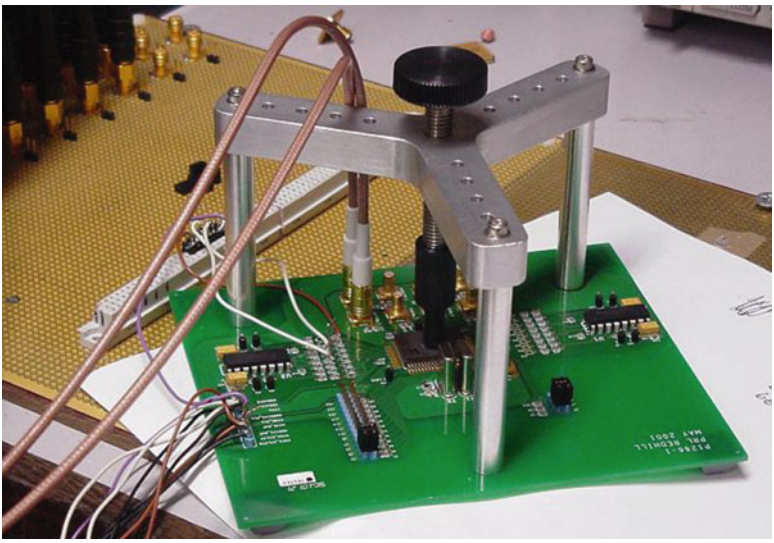


Fig. 10.2 Test board for measuring an analog-to-digital converter (courtesy: R. v. Veldhoven)

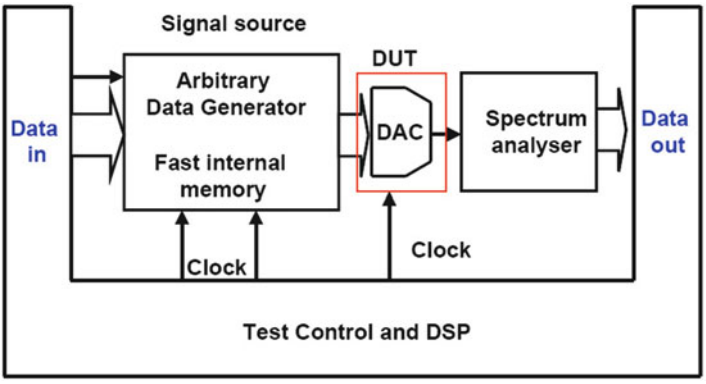


Fig. 10.3 Measurement setup for a digital-to-analog converter

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 holds similarly for digital-to-analog converters. Figure 10.3 shows a potential setup 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

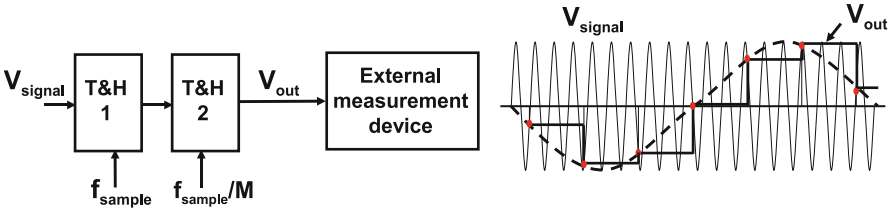


Fig. 10.4 The first track-and-hold circuit is tested at high input and high sample rate. The second device runs a sample rate that is an integer factor slower. The resulting output signal contains signals that correspond to the first track-and-hold output signal and all harmonics

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.

Example 10.1. How can a high-speed track-and-hold circuit be tested without having to accurately measure high-frequency output signals?

Solution. Subsampling can be used to test a sampling device such as a track-and-hold circuit. Two devices are cascaded as in Fig. 10.4. The first device is operated at a high sample and signal rate. The second track-and-hold circuit samples the output of the first device at an integer fraction of the sample rate. The original output signal of the first track-and-hold circuit and its harmonics are subsampled to a low output frequency that can be easily measured. The "subsample" method requires quite some skills to interpret correctly the resulting frequency components.

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

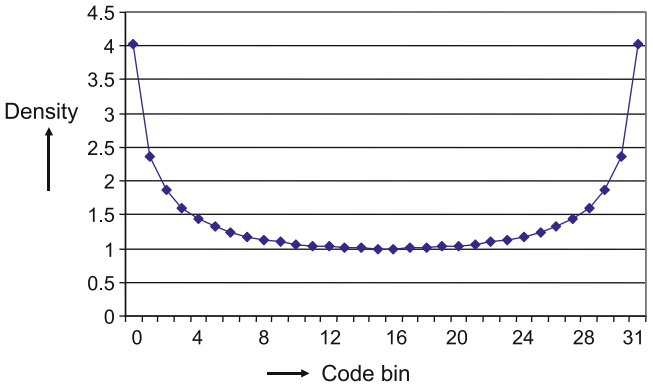


Fig. 10.5 The ideal distribution of hits when a full-scale sine wave is applied to a 5-bit analog-to-digital converter

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

$$y(t) = 0.5 - 0.5\cos(2\pi t/T_{\mathrm{sw}}). \tag{10.1}$$

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 $(T_{\mathrm{sw}}/2)$ corresponds to the total amount of samples. The ratio $\Delta t_y/(T_{\mathrm{sw}}/2)$ is now the fraction of hits that end up in bin Δy.

$$\begin{aligned} t &= \frac{1}{\omega_{\mathrm{sw}}}\arccos(1-2y) \\ \frac{\mathrm{d}t}{\mathrm{d}y} &= \frac{1}{\omega_{\mathrm{sw}}\sqrt{y-y^2}} \\ \frac{\Delta t_y}{T_{\mathrm{sw}}/2} &= \frac{2}{T_{\mathrm{sw}}}\frac{\mathrm{d}t}{\mathrm{d}y}\Delta y = \frac{\Delta y}{\pi\sqrt{y-y^2}} \end{aligned} \tag{10.2}$$

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.

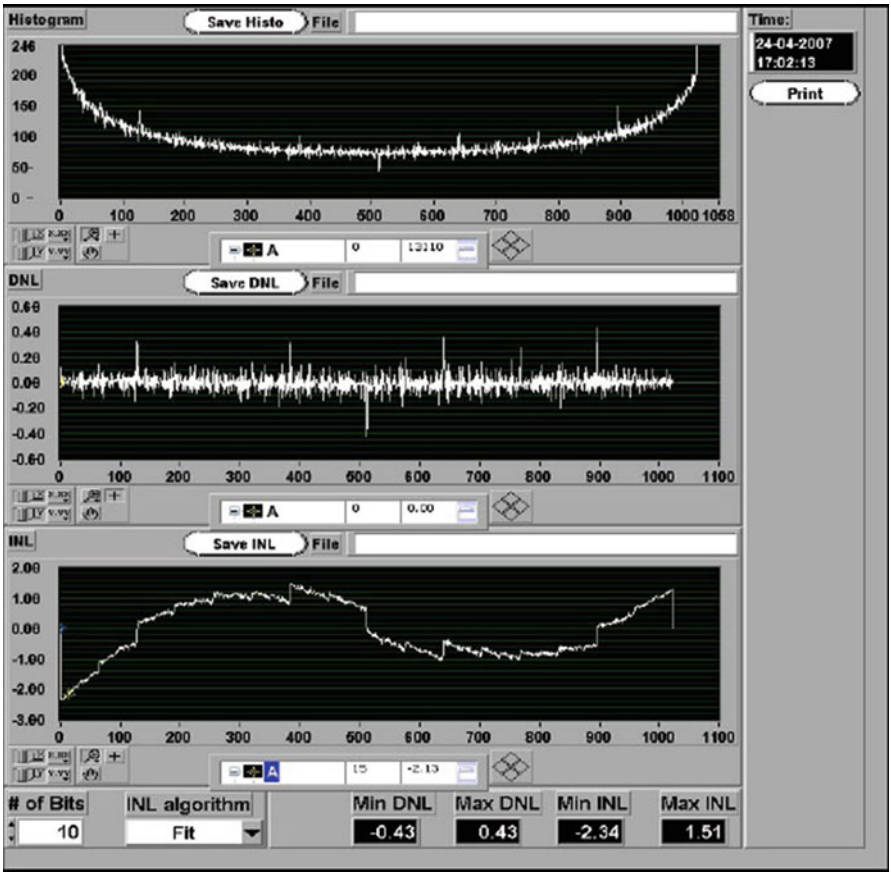


Fig. 10.6 Output of an automated test set-up. *Top*: histogram output, *middle*: DNL, *bottom*: INL

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 sawtooth test 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, routines exist that will allow also amplitudes that extend over the input range. A reconstruction of the input is also possible; see Fig. 10.7.

Example 10.2. How many samples must be acquired to specify the accuracy of the INL with 0.1 LSB?

Solution. In the case of an ideal sawtooth the number of samples in a bin with a size of 1 LSB is determined by the slope of the sawtooth. If the sawtooth rises from

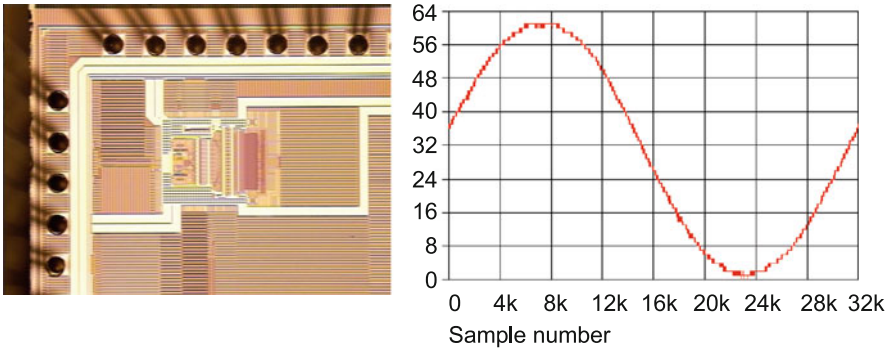


Fig. 10.7 Reconstructed wave form of a 311 MHz signal sampled at 1.44 Gs/s [184]

minimum reference to maximum reference in N_{ST} sample periods, then the average number of hits per bin will be $N_{ST}/2^N$. In order to obtain an accuracy of 0.1 LSB, N_{ST} must exceed 10×2^N. The above calculation for the histogram method 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

$$\frac{\text{hits in bin at } y = 0.5}{\text{total samples}} = \frac{\Delta t_{y=0.5}}{T_{\text{sw}}/2} = \frac{\Delta y}{\pi\sqrt{y - y^2}} = \frac{\Delta y}{\pi/2} = \frac{1}{\pi 2^N/2}. \tag{10.3}$$

With a sine wave the number of samples is $\pi/2$ times larger than when applying a sawtooth signal. To obtain an accuracy of 0.1 LSB in INL and DNL a minimum of $10 \times \pi 2^N/2$ samples are required.

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 frequency 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; see Fig. 10.9. 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. 10.10.

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. 10.11.

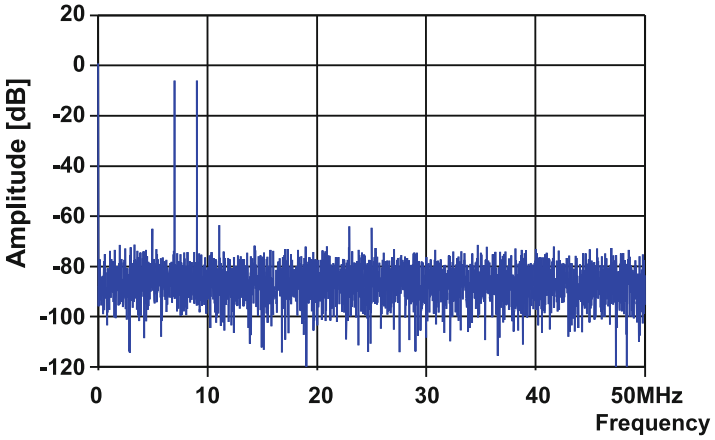


Fig. 10.8 Dynamic measurement of an analog-to-digital converter on intermodulation at f_s = 100 MHs/s

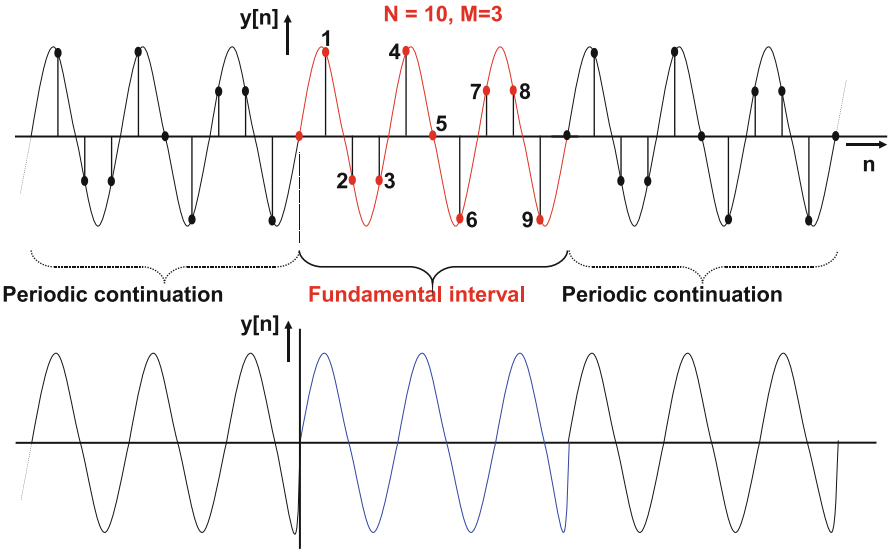


Fig. 10.9 The measured period of a signal is expanded on both sides to enable a fourier transformation. If the signal and the sample frequency do not fit to the window (*below*) frequency leakage will occur

The basic requirement for a good test that avoids both problems is called the "coherent testing" condition:

$$T_{\text{meas}} = \frac{M_s}{f_s} = \frac{M_{\text{signal}}}{f_{\text{signal}}}, \tag{10.4}$$

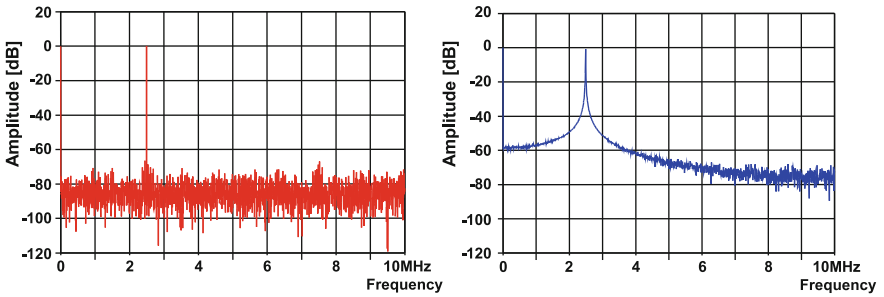


Fig. 10.10 Frequency leakage because of one missing sample: *left*: 4,000 samples, *right plot*: 3,999 samples

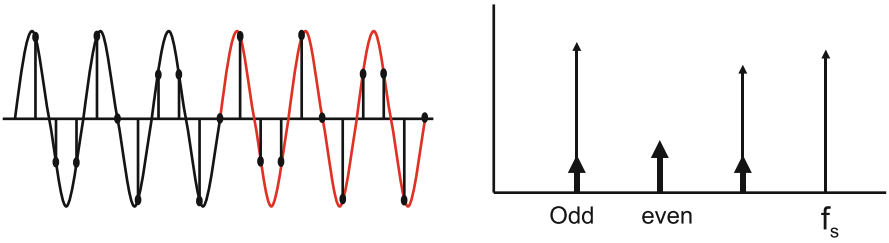


Fig. 10.11 If the sample rate is an integer multiple of the sample rate, a part of the measured samples are simple duplicates of the earlier sequence. In the frequency domain this may lead to the masking of harmonics behind other harmonics or behind the fundamental frequency

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 a sufficiently large M_{signal}, M_s, and T_{meas}.

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}.

Figure 10.12 compares spectra taken with 200 and 2,000 samples.

Example 10.3. Explain the noise floor in Fig. 10.12.

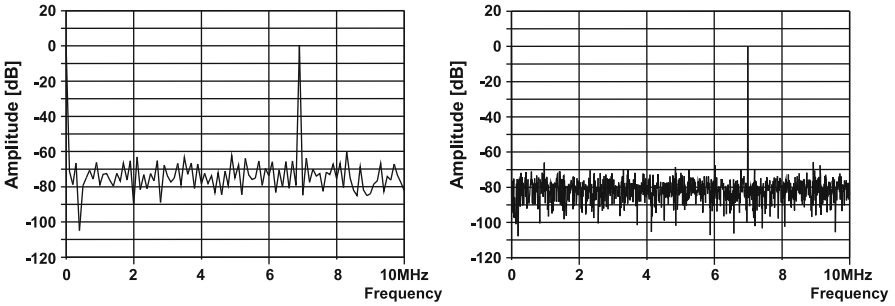


Fig. 10.12 Increasing the measurement period and the number of samples by a factor of 10, reduces the bin size with that factor and lowers the noise floor by 10 dB

Solution. The 8-bit analog-to-digital converter has a theoretical maximum signal-to-noise ratio of $1.76 + 8 \times 6\,\text{dB} = 49.8\,\text{dB}$. A measurement and Fourier transform with 200 samples will result in 101 bins. The quantization energy in Fig. 10.12 is distributed over these 101 bins, so the "noise floor" in the spectrum is expected at a $100\times$ lower energy level: at $49.8 + {}^{10}\log(100)\,\text{dB} \approx 70\,\text{dB}$ 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.

Example 10.4. An ADC is tested during 1 ms at a sampling speed of 20 Ms/s; the performance at 3 MHz signal frequency is required. Calculate an appropriate set of test conditions.

Solution. With $f_s = 20$ Ms/s and $T_{\text{meas}} = 1$ ms, a total of $N = 20{,}000$ samples is generated. For a 10-bit ADC this would allow an accuracy of approximately 0.1 LSB. A 3 MHz input sine wave would show 3,000 periods, and no coherent conditions can be observed. Changing the input frequency to 2.999 MHz will do. At a test period of 1 ms, the spectral resolution (frequency bin) is $1/T_{\text{meas}} = 1$ kHz.

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 km 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 create 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 built-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.

Example 10.5. 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 seen in the FFT plot. (f) What can be the technical disadvantage of a long test time?

Solution. With 25 ns clock period a ramp signal will take $1024 \times 25\,\text{ns} = 25.6\,\mu\text{s}$. In case a sinusoidal signal is used 40 µs is needed. Probably the DSP processing will limit this test. All codes need to be accessed to be sure there are no missing codes. A 1 ms FFT period results in a 1 kHz FFT bin size. For a 40 Ms/s sample rate and a 19.999 MHz input signal, $N = 40{,}000$ and $M = 19.999$ which numbers are mutually prime. The number of bins in the FFT is $40,000/2 + 1 = 20{,}001$ bins of 1 kHz. The 10-bit converter should have an ideal noise level at $10 \times 6 + 1.76 = 62$ dB. If this noise level is spread over 20,000 bins the level will drop another $10^{10}\log(20,000) = 43$ dB. The total quantization noise level can reach -105 dB, so most likely some thermal noise source will dominate. A long test period allows to perform a detailed FFT that may reveal more details of the analog-to-digital converter's performance.

Exercises

10.1. Compare the advantages and disadvantages of testing the performance of an analog-to-digital converter by connecting a high-performance digital-to-analog converter to the output or by analyzing the digital output in a signal processor.

10.2. Propose a sine-based equivalent test method for a digital-to-analog converter.

10.3. The histogram measurement method uses a sine wave. Set up a test scheme along the same lines using a uniform distributed random signal.

10.4. Can a sigma–delta modulator be tested with a histogram method?

10.5. An 8-bit 20 Ms/s analog-to-digital converter is tested during 0.1 ms with a half-scale sine wave. The result is processed via an FFT. Make a drawing of the expected FFT result.

10.6. A 6-bit ADC needs to be tested dynamically at 4 Gs/s in a DSP-based environment. Determine the minimum test time needed to have accessed all codes. There is 40 µs test time available for the FFT. Determine the input frequency at the Nyquist edge for a good test. What is the number of bins in the FFT? Determine the approximate noise level of the FFT.

10.7. An 8-bit 600 Ms/s analog-to-digital converter is used in a communication system where a spurious-free dynamic range of 80 dB in 2 MHz bandwidth is required. What test is required.

10.8. An analog-to-digital converter is part of a system on silicon. The sample clock is generated on chip and cannot be accessed separately. Define a method to quantify the jitter of the clock.

Chapter 11
Technology

Abstract Technology is the foundation on which every circuit is build. Over the last four decades technology developments were dominated by the rapid evolution of CMOS technology. This chapter examines a few aspects that are relevant to the analog-to-digital designer.

A major worry in design is the variation of parameter values. With the improved control over processing, the batch-to-batch variation is largely under control. However, now a new class of phenomena has appeared: statistical variations. In conventional ICs, analog circuits with a differential operation (e.g., analog-to-digital converters) were already affected by this random parameter spread. The remaining variation between otherwise identical components is generally described by "mismatch" parameters. Next to these random phenomena also systematic errors called "offsets" play an increasingly important role. Understanding and mitigating these effects require statistical means and models.

Some sections describe the modeling of systematic and random effects that originate from physical, electrical, thermal, and lithography effects in devices causing intra-die variations.

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, and packaging and substrate interference.

11.1 Technology Road Map

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,

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_11,

© Springer Science+Business Media, LLC 2013

and tolerances. Many of these specifications in CMOS are derived from the ITRS road map [20]. 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 (polysilicon) 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 map. The logic gate density is measured on a minimum height two-input NAND. Such a low-height pitch would in practical designs lead to a massive wiring congestion. Normally cell heights between 9 and 14 times the minimum design rule 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.

Table 11.1 Some characteristic data from the ITRS road map 2011, supplemented with older data

Year (ITRS analog/RF)	1995	1998	2000	2002	2004	2007	2010	2013	2015	2018
ITRS node (nm)	350	250	180	130	90	65	45	32	21	15
V_{DD} (V)	3.3	2.5	1.8	1.2	1.2	1.1	1.1	1	0.8	0.73
Min gate length(nm)	350	250	100	75	53	37	28	22	17	13
d_{ox} (nm)	7.0	5.7	3.6	3.2	2.2	1.9	1.4	1.3		
g_m/g_{ds} ($5\times L_{min}$)	200	150	120	47	32	30	30	30	55–73	58–65
$1/f$ noise ($\mu V^2\mu m^2$/Hz)	500	300	180	180	160	100	80	70	38–58	42–46
A_{V_T} matching (mVμm)	7	6	6	6	6	5	5	5	1.6–2	0.6
Peak f_T (GHz)	20	40	70	100	170	240	320	400	510	750
Nfmin (dB)					0.33	0.25	0.2	<0.2		
DRAM size	64M	128M	256M	512M	1G	4G	8G	8G	8G	16G
Tr density logic (tr/cm^2)				39M	77M	154M	309M	617M	2,500M	5,000M
Logic density (gates/mm^2)	20k	40k	100k	200k	400k	800k	1.6M	3M	6M	12M
SRAM density (bits/mm^2)				350k	700k	1.4M	3M	6M	20M	48M
Digital clock speed (GHz)				1.68	3.99	6.74	11.5	19		

The range in the last two columns corresponds to three transistor options: traditional bulk, fully depleted SOI, and metal gate

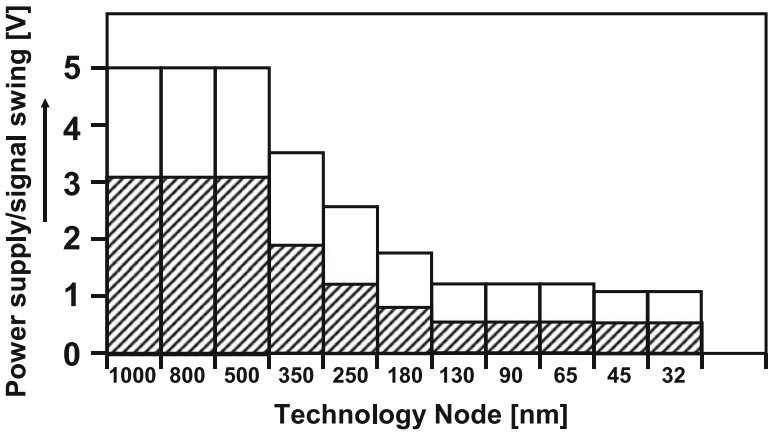


Fig. 11.1 The power supply evolution as a function of process generation. The signal swing (*shaded*) is derived by using the minimum power supply minus a threshold and a bias voltage

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 [296]

$$\text{Power} = N_i f (C_{\text{load}} + C_{\text{gate}}) V_{\text{DD}}^2 + P_{sc} + P_{\text{leak}}. \tag{11.1}$$

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 (fCV^2) is dominant, while in standby mode the leakage current is the main power surge. Leakage originates mostly from the subthreshold 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 cutoff frequency f_T for the same current. Figure 11.2 shows the cutoff frequency as a function

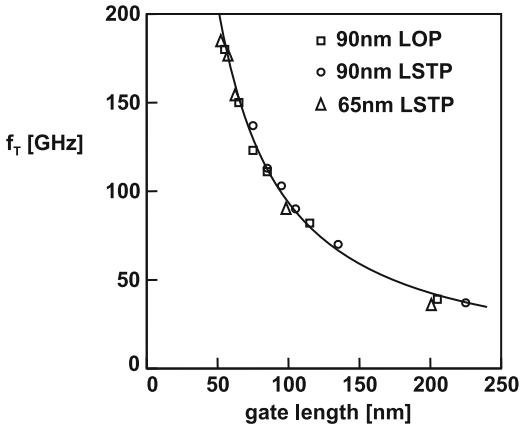


Fig. 11.2 Evolution of the cutoff frequency f_T for minimum gate length NMOS transistor, from [297]. LOP means low-operating power, and LSTP means low standby power

of technology node. The cutoff frequency is the ratio of transconductance and load capacitance ($f_T = g_m/2\pi C$). These cutoff frequencies are determined by RF designers at a high gate voltage of $V_G \approx V_{\mathrm{DD}}$. 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.

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 [298]:

- No extra masks, no process adaptations
  The extent of analog characterization and monitoring of specific process parameters. Examples are MOST noise, matching, temperature, and voltage dependencies of passive and active components. Proper characterization (of, e.g., matching) reduces margins and allows to design for minimum power consumption. The reduction in process tolerances also falls in this category; see Table 11.2.

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 (°C)	25	85	120	140
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×

- No extra masks, minor process adaptations
  The definition of MOST thresholds is crucial for low-voltage operation. Analog designers want high-threshold transistors for input stages and low-threshold transistors for switches. Structures that are normally considered parasitics in the baseline process are useful in an analog design. Examples are vertical PNP transistors, large MOSTs as capacitors, and resistors and multilayer 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 the following: multiple thresholds in order to optimize for analog switching, lightly doped drain (LDD) suppression on the source side of a MOS transistor, definition of resistors by removing the silicide layer, and a special dielectric 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 are second gate oxide for high analog voltages and second polysilicon layer. A double polysilicon 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

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 nonvolatile 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 is required.

In the scope of this chapter only phenomena that can be considered static within the observation period are discussed.

Table 11.3 Time dependencies of some variability effects

Static	Seconds	Microseconds	Picoseconds
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		
Differential design, calibration	Compensation circuitry	Wire width, chopping, common-centroid placement	Decoupling, shielding, clock cleaning

The lower row lists the methodologies the designer has at his disposal to mitigate these 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 www.springerlink.com and was co-authored with Hans Tuinhout and Maarten Vertregt.

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

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 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 multimillion 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 into two classes: deterministic and stochastic and 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 (Table 11.5).

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

Table 11.5 Deterministic and statistical effects

Deterministic	Pseudo-statistical	Statistical
Process	Substrate noise	Line-edge roughness
Supply voltage	Dyn IR drop, jitter	Noise $(1/f, kT)$
Temperature	Temp. gradient	Granularity
Offsets, IR drop	Cross talk	Dopant fluctuation
WPE, stress, STI	Substrate noise	Mobility fluctuation
Proximity effects	CMP density	Interface states
Soft breakdown	Drift	Work-function fluctuation

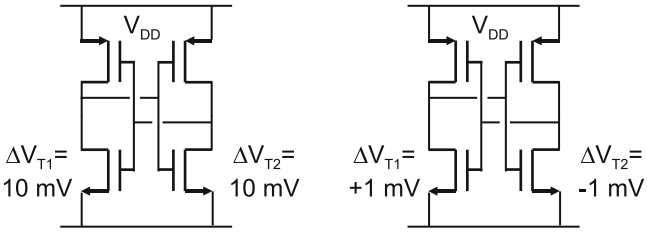


Fig. 11.3 Two latches in a meta-stabler condition experience global and local offsets

source changes the behavior of every individual device with respect to the average, the effect will be described with stochastic means; see Table 11.2.

Example 11.1. The latch in Fig. 11.3(left) is in a metastable position. Due to a global change in back-bias voltage the threshold voltage of the NMOS transistor shifts by 10 mV. Another metastable latch (right) experiences cross talk from neighboring wiring equivalent with a positive, respectively, negative shift on both nodes. What will happen with these two latches?

Solution. Global excitations and variations are canceled by differential structures such as the differential latch topology of Fig. 11.3. The global change of threshold voltage will not impact the latch and the metastable condition will continue.

However, a local variation will like asymmetrical cross talk and will affect also differential circuits. Now the latch will amplify the cross-talk signal and flip into a 1–0 state.

11.3 Deterministic Offsets

Deterministic or systematic offsets between pairs of resistors, CMOS transistors, or capacitors are due to electrical biasing differences, mechanical stress, and lithographic and technological effects in the fabrication process [299].

It may seem trivial, but good matching requires in the first place that matched structures are built from the same material and are of equal size. This implies that

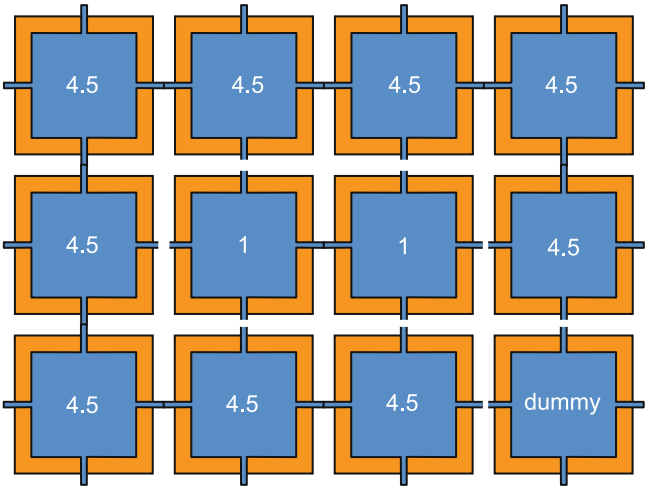


Fig. 11.4 Example of a 1:4.5 ratio, laid out with 11 units. For symmetry reasons one dummy capacitor is left in the lower right corner

a 4.5:1 ratio is constructed as a 9:2 ratio, with 11 identical elements, arranged in a common centroid; see Figs. 11.21 and 11.4. The required 4.5:1 ratio holds for every aspect of the combination: area, perimeter, coupling, etc. Any variation will scale with the same ratio. This technique is applied in capacitive filters or binary-weighted capacitor banks, digital-to-analog conversion elements, etc. Also larger ratios or less trivial integer ratio's like 22/7 for π are preferred for matching over simply multiplying the width by 3.14. In case the required ratio cannot be approximated sufficiently within the boundary conditions, an optimum ratio is chosen and one cell is increased or reduced in size to obtain the required ratio.

A pitfall can occur when an existing layout is scaled and the resulting dimensions rounded off to fit the new layout grid. The rounding operation can result in unexpected size deviations in originally perfectly matched devices.

11.3.1 Offset Caused by Electrical Differences

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. On a building block level DC offset can be introduced by phenomena as self-mixing; see Sect. 3.1.3. Deterministic effects can indirectly be caused by electrical currents or voltages, e.g., heat gradients due to power dissipation and temporary charging effects due to phenomena as negative bias temperature instability [300].

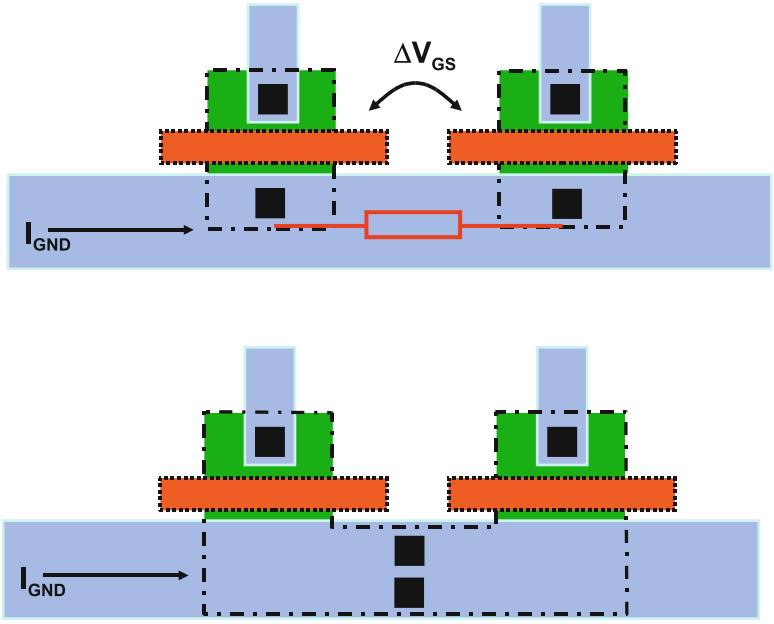


Fig. 11.5 A small resistance in the source connection wire will result in offset in the gate voltage. A common connection point is better from an offset point of view

Figure 11.5 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 starlike connection is well suited to avoid this problem. A rigorous inspection of the layout in which the main current paths have been identified is always needed when offset problems are suspected.

Example 11.2. The NMOS transistors in Fig. 11.5 have dimensions of 20/0.2 in a 0.18 µm process and are biased with 0.2 V drive voltage. The sources are connected with a 0.2 µm metal wire of 0.1 Ω. How much offset will occur?

Solution. With a current factor of $\beta_\square = 300\,\mu A/V^2$ (see Table 2.20) each transistor will conduct a current of 0.6 mA. Between the sources there is some 25 µm distance, resulting in a resistance of 12.5 Ω. So the voltage offset is 7.5 mV translating into 18 µA of current difference.

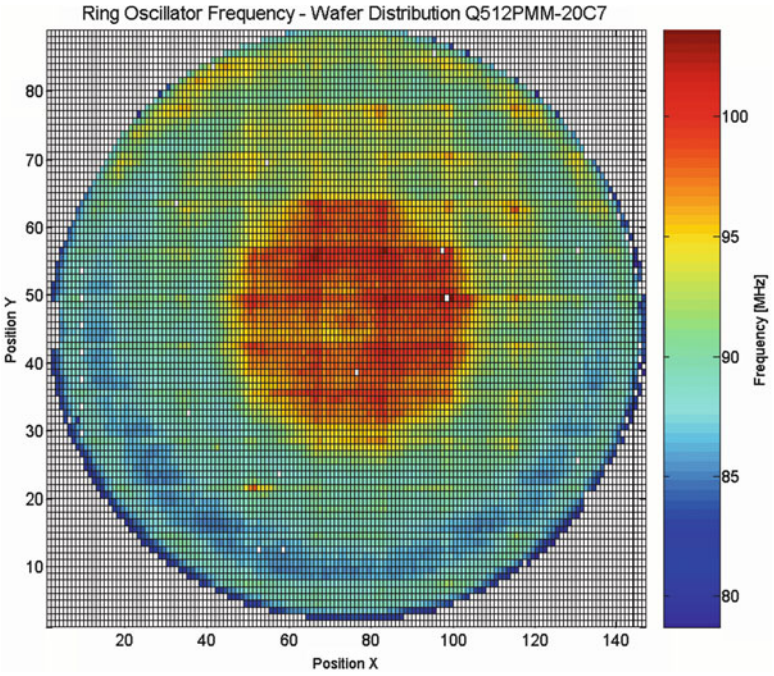


Fig. 11.6 A wafer from a 90-nm CMOS production lot on which a free-running oscillator frequency is measured. The frequency deviation (80–100 MHz) is largely an indication for the variation in gate length due to mask-plate dimensional errors, lithography, and the processing steps that determine the electrical gate length (courtesy: B. Ljevar, NXP)

11.3.2 Offset Caused by Lithography

During the lithography process the structures drawn in the layout are transferred to a mask and in the resist. Accuracies on masks in advanced processes range from 1 to 5 nm on silicon.[2] 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 11.6 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 nonuniformity 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 seven devices.

[2]Price is the determining factor.

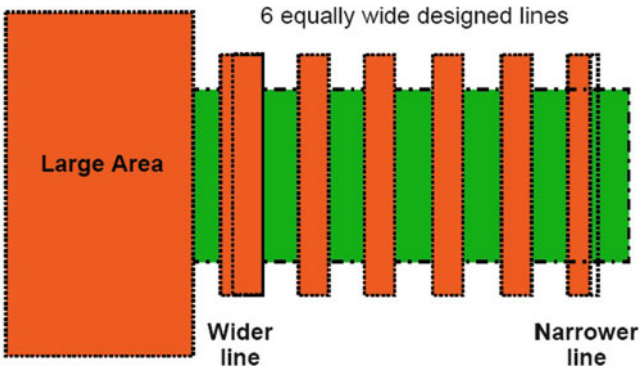


Fig. 11.7 The proximity effect: lines with large neighboring structures grow in size, while lines next to open space shrink

This $\approx$400 mm^2 reticle is repeated to pattern the entire wafer. Random effects are averaged over the nearly 100 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.

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 layout phase by hand.

11.3.3 Proximity Effects

Figure 11.7 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 layout 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

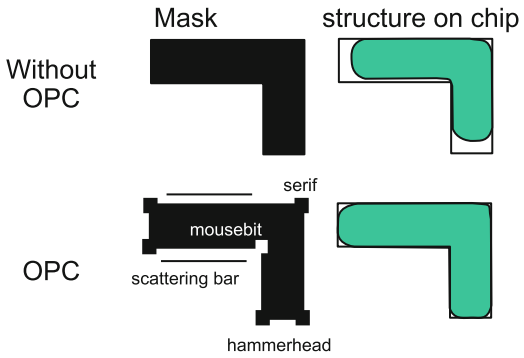


Fig. 11.8 Standard lithography results in rounded structures that are too short. With the help of optical proximity corrections in the mask the net result on a wafer comes much closer to the desired dimensions

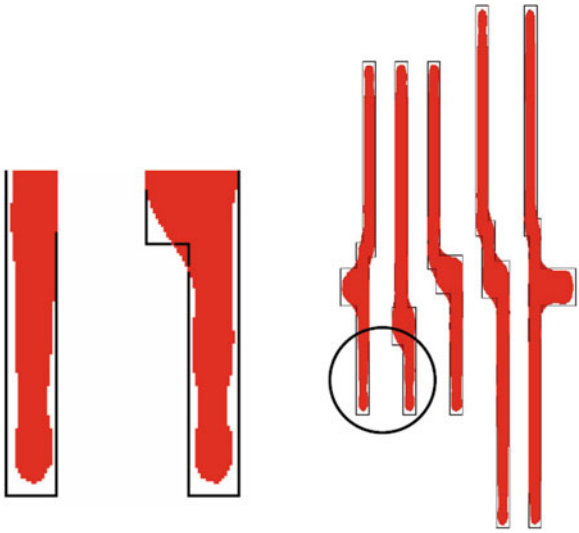


Fig. 11.9 The tips of the structures are enhanced by the OPC tool in this lithography simulation

width patterns pre-distortion is applied to the mask in the form of optical proximity correction (OPC) (Fig. 11.8). An example of dimensional deformations of an advanced lithographic tool is visible in the lithography simulation of Fig. 11.9.

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.

11.3.4 Implantation-Related Effects

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.

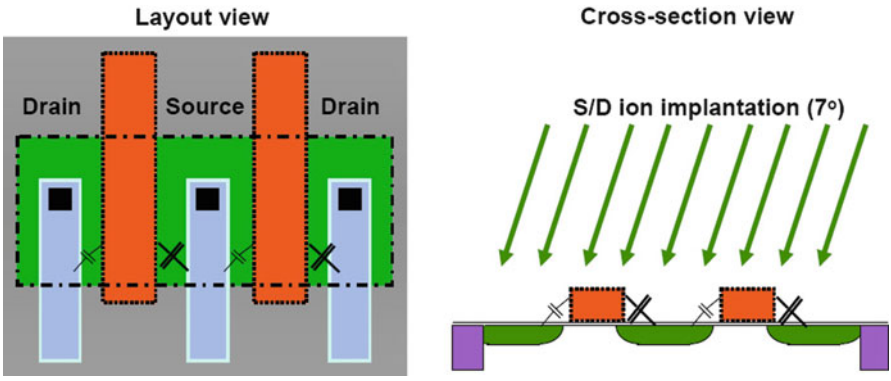


Fig. 11.10 The source and drain diffusions are implanted under an angle. This causes asymmetry for drain and sources

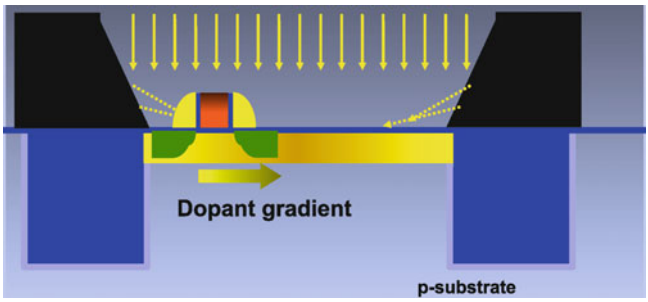


Fig. 11.11 The well-proximity effect occurs during the well implant. The drawn transistor is fabricated afterwards, but resides in a well with a horizontal doping gradient [301, 302]

This effect is called "channeling." During the ion-implantation steps in 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. 11.10. 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 antiparallel. In integrated circuit manufacturing there are more processing steps that can cause similar asymmetries.

The well-proximity effect in Fig. 11.11 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 photoresist boundary and cause a horizontal gradient in the well implantation dose. Variations ranging from 1 µm [301] to 2 µm [302] have been reported. It is advisable to use an overlap well beyond the minimum layout design rule for implantation masks, where possible.

Many effects can occur in a process that affect the reproducibility of transistors and other components. Figure 11.12 shows three transistors in an *n*-well. A larger

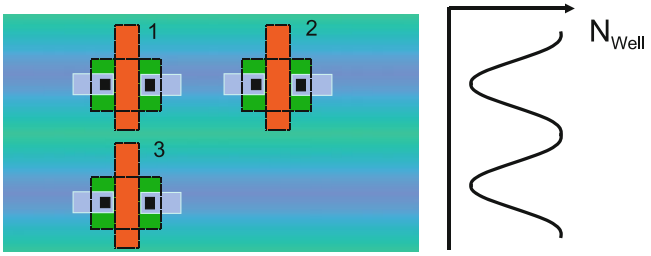


Fig. 11.12 A large area is implanted by scanning the implanter beam over the wafer area. A sharply focussed beam will cause stripes and a doping profile in a direction perpendicular to the scanning direction. Differences in threshold voltages will occur

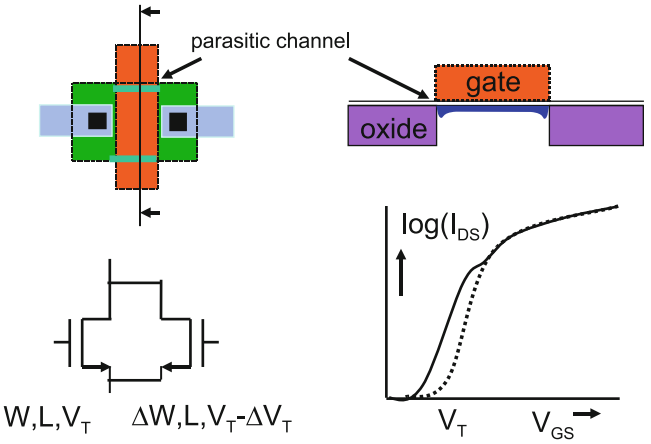


Fig. 11.13 The effect of parasitic channels along the field oxide

structure like a well is implanted by scanning the implantation beam stripe by stripe over the area. A sharply focussed beam will create a distinguishable stripping pattern with a doping gradient running perpendicular to the scanning direction. As a consequence transistors 1 and 2 will match, but transistor 3 will deviate.

A lot of attention is normally focussed on the gate to source and drain perimeter. The boundary of the gate to the surrounding isolation material, such as field oxide or shallow-trench layers, is equally important. In Fig. 11.13 the doping profile next to the isolating oxide is such that a parasitic channel can occur. This can happen when a well doping and the substrate doping locally compensate each other. The effect is that the desired transistor is flanked by two narrow width but low-threshold voltage transistors. Especially in low-current regime and in weak inversion the parallel transistors will generate a considerable current, often indicated by a bump in the subthreshold $I_{DS} - V_{GS}$ characteristics. Slight differences in biasing between transistors of a matched pair lead to considerable current deviations.

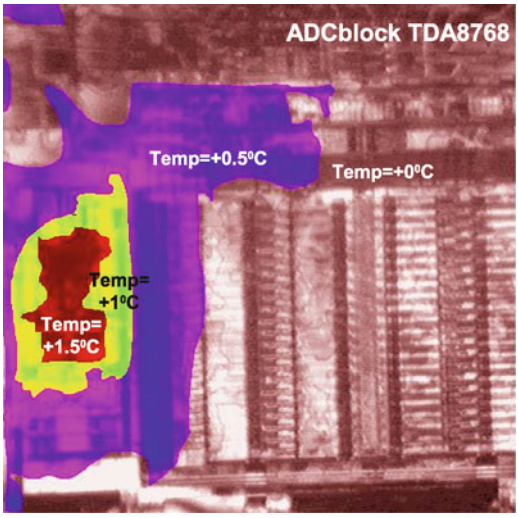


Fig. 11.14 Temperature difference can be visualized by means of a liquid crystal technique. Here the track-and-hold circuit generates the heat, which spreads out via the wiring into the ladders of this analog-to-digital converter

11.3.5 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.[3] 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. 11.14. In a system-on-chip the different blocks show a great variety of power dissipations. On-chip memories show 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 W 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 mV 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.

[3] An estimate for a technological gradient can be approximated by dividing the difference between the maximum and minimum of the parameter by half of the wafer diameter.

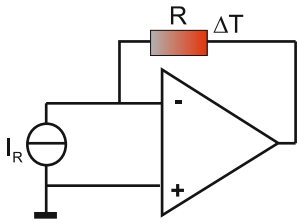


Fig. 11.15 A high-quality sinusoidal current is converted into a distorted output voltage because the resistor will heat up

Example 11.3. A current source $I_R(t)$ delivers a perfect sinusoidal current. This current is converted into a voltage by means of a high-gain high-bandwidth opamp and an on-chip resistor (Fig. 11.15). The resistor has a negligible voltage coefficient and a temperature coefficient κ. After processing, a severe distortion appears at the output for low-frequency signals. What is a potential reason for this distortion?

Solution. During a sine wave the current through the resistor varies and consequently the momentary heating. As the thermal relaxation of IC components is in the range of microseconds, the temperature of the resistor will follow the magnitude of the dissipated power. At any moment the temperature rise will be proportional to the dissipated power and the thermal conductivity: $T - 25\,^\circ\mathrm{C} = I^2{}_R(t)R\kappa$. So the effective resistance is $R(T) = R(T = 25\,^\circ\mathrm{C})(1 + TC \times (T - 25\,^\circ\mathrm{C}))$.
So the output voltage now is $V_{\text{out}} = I_R(t)R(25\,^\circ\mathrm{C})(1 + TC\kappa I^2{}_R(t)R(25\,^\circ\mathrm{C}))$. In case of a sine wave with zero bias, a third-order distortion component will appear. The magnitude of the distortion will reduce at higher signal frequencies.

11.3.6 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 layout 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. 11.18. Some major causes for stress are:

- 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.

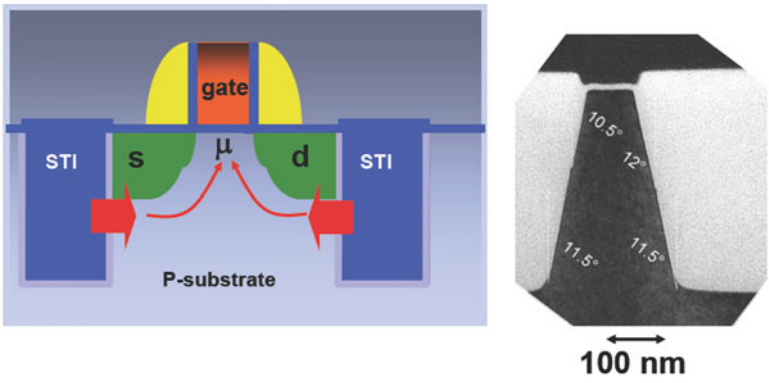


Fig. 11.16 Stress in integrated circuits is caused by thermal expansion of various materials. The blocks of shallow trench isolation (STI) create a considerable stress in the active area of the transistor, causing current variations up to 10%. The photo shows a cross section, where the slopes of the STI are indicated (courtesy: C. Detcheverry, NXP)

- 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 airflow.
- In modern isolation techniques a trench is etched in the substrate and silicon dioxide is deposited and planarized: shallow trench isolation (STI) (Fig. 11.16). 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 STI has been subject of various studies [302–305]. At temperatures of over 1,000°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 [303]. The mechanical stress deforms the lattice and causes the mobility in the transistors to vary.

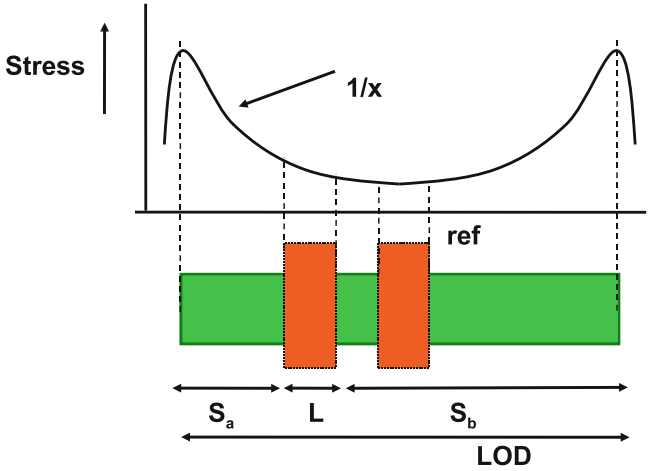


Fig. 11.17 A transistor is placed asymmetrically with respect to the reference device. The edges between active and STI inflict stress in the channel region; however due to the asymmetry the effect of the stress is different for both devices

Figure 11.17 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 [303]:

$$\begin{aligned} S &= \frac{1}{S_a + L/2} + \frac{1}{S_b + L/2} \\ \mu_{\text{eff}} &= \frac{1 + K_{s,\mu} S}{1 + K_{s,\mu} S_{\text{ref}}} \mu_{\text{ref}} \\ V_{\text{T}} &= V_{\text{T,ref}} + K_{s,V_{\text{T}}} (S - S_{\text{ref}}), \end{aligned} \tag{11.2}$$

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 [305] an experiment is reported where the STI-to-active edge of the source and drain is varied. Figure 11.18 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 [302, 303].

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 a digital circuit; however, in precision analog design these effects must be taken into account. The generalized model uses

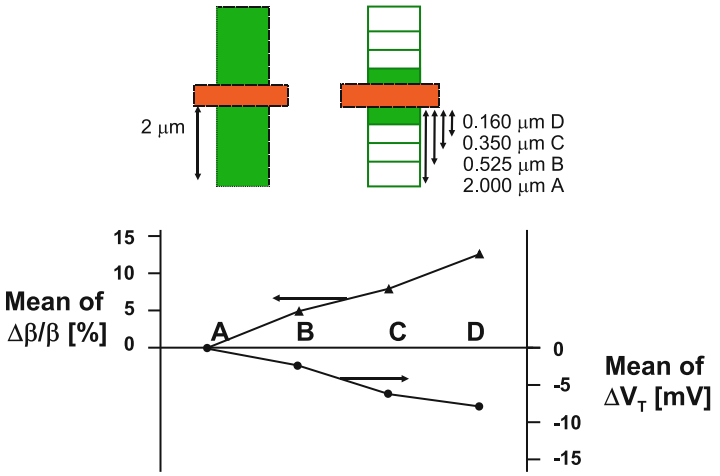


Fig. 11.18 An experiment shows the influence of the STI edge on the drain current and threshold voltage [305]. *Top*: a 65-nm 2/0.5 µm NMOS reference transistor is designed with the STI edge of the source and drain at 2.0 µm. A second device has a similar STI distance (A) or at 0.525 (B), 0.35 (C), and 0.16 µm (D)

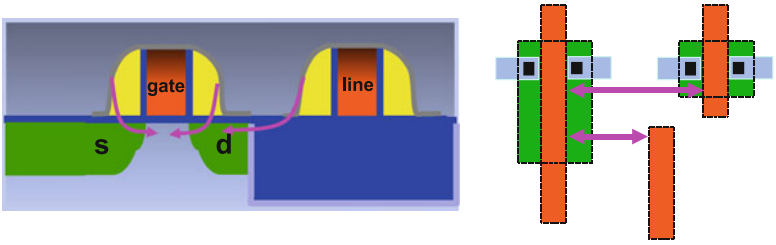


Fig. 11.19 The poly-space effect is caused by the etch-stop layer

$$S = \sum_{i=1}^{n} \frac{\pm W_i}{2W} \frac{1}{S_i + L/2}, \tag{11.3}$$

where the first term relates the width of the edge to the gate width and the ± sign to an STI-active edge or an active-STI edge.

The poly-space effect [306] in Fig. 11.19 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 [307] and to variations in the stress pattern. When a wiring pattern is placed at different spacings on one side of a bipolar pair [308], the resulting current

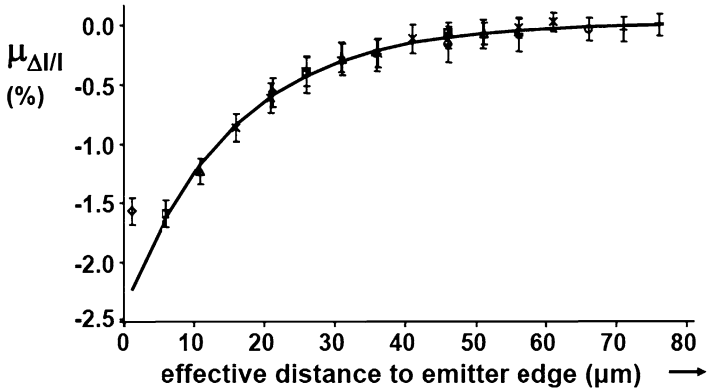


Fig. 11.20 An aluminum wire placed at a certain distance of the emitter causes current deviations that can be measured up to 40 µm [308]

Table 11.6 Potential magnitudes of deterministic variability effects

Effect	Magnitude
Power supply voltage drop	Voltage shifts up to 100 mV in poorly designed power grids
Lithography, etch depletion, etc.	Dimensional errors up to 100 nm, in >0.25 µm CMOS
Proximity effects	Dimension errors up to 20 nm in advanced mask making
Temperature gradient	1°C over 20–50 µm close to strong local heat source. With 2 mV/°C threshold sensitivity
LOD or STI effect [302, 303, 305]	Mobility changes 12% between minimum drain extension and large drain
Well proximity[301, 302]	Threshold shift up to 50 mV if mask edge closer than 2 µm
Metallization[308]	Mobility changes 2% between close and 40 µm far metal track
Metal coverage of gates[307]	Currents may deviate up to 10–20%

variation is in the order of 1%. Even wiring on top level (e.g., tiling patterns) can cause stress [309]. Copper wiring behaves similar to aluminum wiring: deviations in the 1–2% range.

Figure 11.20 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 layout is required for analog circuits that should yield offsets below 1%. Table 11.6 lists a number of effects.

11.3.7 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

Table 11.7 Guide lines for the design of matching components

1	Matching components are of the same material and have the same form, dimensions, and orientation
2	The potentials, temperatures, pressures, and other environmental factors are identical
3	Currents in components run in parallel, not antiparallel 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, and stress. Do not switch off the tiling pattern generation, but define a symmetrically placed tiling pattern

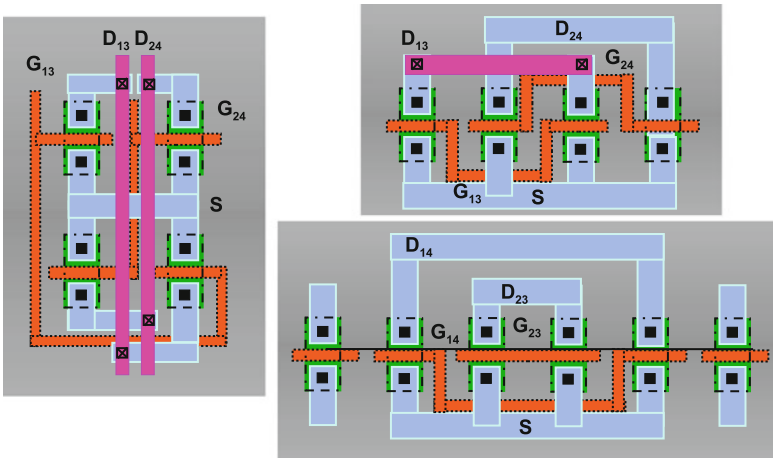


Fig. 11.21 Some common centroid arrangements

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 [38]. Despite the complex nature of some variations, a number of guidelines can be formulated to minimize the effect of these offset causes; see Table 11.7 [310].

Common centroid structures are used to reduce the gradient effects [311]. Applying a common centroid geometry is not trivial; an asymmetry in the wiring scheme can easily cause more problems than are solved; see Fig. 11.21. 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. Large resistor structures are sensitive to substrate coupling.
- Capacitors: The absolute value is usually well defined in a double polysilicon 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.

Example 11.4. In Fig. 11.22 a cross-coupled transistor layout is shown. What is the most important mistake in the layout?

Solution. There are several minor mistakes in this layout: the upper-right transistor is partially covered by metal-2 (purple). This will create stress and potentially annealing problems. The connections of the drains in the middle are far from symmetrical. The current will flow from top side to bottom side and become part of a current loop. However, the most important mistake is the placement of the joint source "S" connection. In the connection from the lower left transistor to the "S" terminal now flow the currents of three transistors, instead of the current of a single transistor. It is better to connect the "S" terminal in the middle of the left side metal-1 connection.

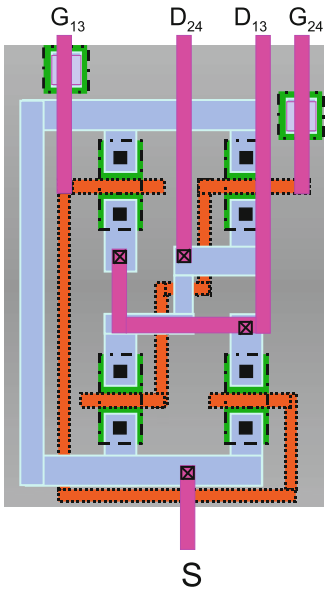


Fig. 11.22 A common centroid arrangement

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 are 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. Various authors have investigated random effects in specific structures: capacitors [24–26, 312, 313], resistors [314, 315], MOS transistors [316–319], and bipolar devices [320].

In a general parameter fluctuation model [31], 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

$$\begin{aligned}\Delta P(x_{12},y_{12}) &= P(x_1,y_1) - P(x_2,y_2)\\ &= \frac{1}{\text{area}}\left[\int\int_{\text{area}(x1,y1)} P(x',y')\mathrm{d}x'\mathrm{d}y' - \int\int_{\text{area}(x2,y2)} P(x',y')\mathrm{d}x'\mathrm{d}y'\right].\end{aligned} \tag{11.4}$$

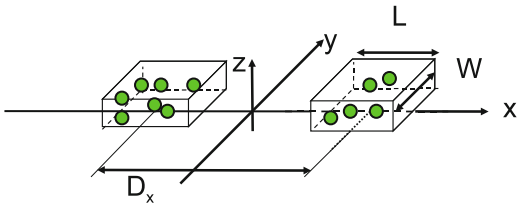


Fig. 11.23 Definition of the area function $h(x,y)$

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 multiplication and allows separating the geometry-dependent part from the mismatch source:

$$\Delta\mathcal{P}(\omega_x,\omega_y) = \mathcal{G}(\omega_x,\omega_y)\mathcal{P}(\omega_x,\omega_y). \tag{11.5}$$

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 dependence function $\mathcal{G}(\omega_x,\omega_y)$. These two components are analyzed separately.

The geometry function as shown in Fig. 11.23 for a pair of rectangular devices with area WL is defined as

$$h(x,y) = \begin{cases} \dfrac{1}{WL}, & (D_x/2 - L/2) < x < (D_x/2 + L/2), -W/2 < y < W/2 \\ \dfrac{-1}{WL}, & (-D_x/2 - L/2) < x < (-D_x/2 + L/2), -W/2 < y < W/2 \\ 0, & \text{elsewhere.} \end{cases} \tag{11.6}$$

For convenience it has been assumed that both areas are at a distance D_x along the x-axis. Some mathematical manipulation results in a geometry function for the difference in **paired**-transistors parameters (Fig. 11.23):

$$\mathcal{G}(\omega_x,\omega_y) = \frac{\sin(\omega_x L/2)}{\omega_x L/2}\frac{\sin(\omega_y W/2)}{\omega_y W/2}[2\sin(\omega_x D_x/2)]. \tag{11.7}$$

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. 11.21(left) has a geometry function where the last term in brackets in Eq. 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 [31]:

$$\sigma^2_{\Delta P} = \frac{A^2_P}{WL} \tag{11.8}$$

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 [321].

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. 11.6. 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 [31]:

$$\sigma_{\Delta P}^2 = \frac{A_P^2}{WL} + S_P^2 D_x^2 \tag{11.10}$$

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.[4]

For a group of four cross-coupled transistors is found in a similar way:

$$\sigma_{\Delta P}^2 \approx \frac{A_P^2}{2WL} + S_P^2 D_x^2 \frac{D_y^2}{\text{wafer diameter}^2}. \tag{11.11}$$

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

$$V_\mathrm{T} - V_{FB} = \frac{Q_B}{C_\mathrm{ox}} = \frac{qN_x z_d}{C_\mathrm{ox}} = \frac{\sqrt{2q\varepsilon N_x \phi_b}}{C_\mathrm{ox}}, \tag{11.12}$$

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. 11.24) is defined by its width W, length L, and a depletion region depth $z_d = \sqrt{2\varepsilon\phi_b/qN_x}$, then the volume of the

[4]The importance of this model extension is more in signaling a potential problem in the process than in accurately modeling a phenomenon.

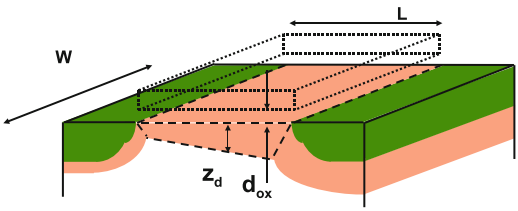


Fig. 11.24 A cross section through a MOS transistor indicating the depletion region

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, and punch-through implantation.[5] 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 ions is now approximated by Poisson statistics:

$$\sigma_c^2 = \mu_c \quad \mu_c = WLz_d N_x, \quad \Rightarrow \sigma_c = \sqrt{WLz_d N_x}. \tag{11.13}$$

The threshold variance can now be derived from Eq. 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

$$\sigma_{\text{single } V_T} = \sigma_c \frac{\partial(V_T)}{\partial(WLz_d N_x)}. \tag{11.14}$$

As matching usually occurs between pairs of transistors, the variance of the *difference between two transistors* is [31, 322]

$$\sigma_{\Delta V_T} = \sqrt{2}\sigma_{\text{single } V_T} = \frac{qd_{ox}\sqrt{2N_x z_d}}{\varepsilon_{ox}\sqrt{WL}} = \frac{A_{V_T}}{\sqrt{WL}} \propto \frac{d_{ox}\sqrt[4]{N_x}}{\sqrt{WL}}. \tag{11.15}$$

This function is commonly depicted as a linear relation between $\sigma_{\Delta V_T}$ and $1/\sqrt{\text{area}}$. Figure 11.25 shows an example of the measured dependence for $\sigma_{\Delta V_T}$ versus $1/\sqrt{\text{area}}$. The slope of the line equals the parameter A_{V_T}. In this plot the layout dimensions were used. For the smallest sizes the effective gate area is smaller due to under-diffusion and channel encroachment. In order to test the hypothesis that depletion charge is the dominant factor in threshold matching, Table 11.8 compares the A_{V_T} coefficients as measured and as calculated using the above formula. The quantity $N_x z_d$ was derived from process simulation which was tuned with accurate C/V measurements [298]. In these relatively straightforward 0.6 µm and 0.8 µm

[5] For ease of understanding only a uniformly distributed dopant is assumed; more complicated distributions must be numerically evaluated.

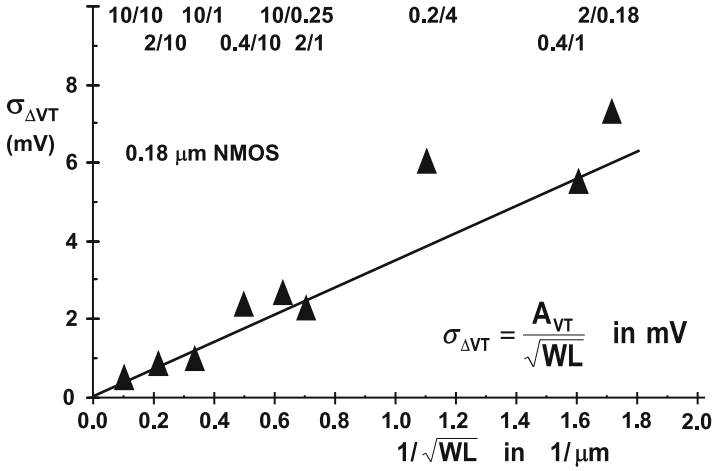


Fig. 11.25 The standard deviation of the NMOS threshold and the relative current factor versus the inverse square root of the area, for a 0.18 µm CMOS process

Table 11.8 Comparison of measured and calculated threshold mismatch coefficients

	A_{V_T} measured	A_{V_T} calculated
0.8 µm data		
NMOST (mVµm)	10.7	10.6
PMOST (mVµm)	18.0	18.6
0.6 µm data		
NMOST (mVµm)	11.0	7.4
PMOST (mVµm)	8.5	8.6

CMOS processes the fit is good for three out of four A_{V_T} coefficients. The deviation 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 [307].

The basic threshold-voltage mismatch model[6] has been extended by various authors. More geometry dependence factors can be included to address deep-submicron effects [323]. A fundamental limit for dopant fluctuation related mismatch was derived in [324]. Work reported in [305] indicates that there is no relation between deterministic variations and random dopant fluctuations.

[6] An often proposed "happy-hour" mismatch model explains threshold fluctuation as a form of petrified $1/f$ noise. Threshold mismatch is determined by the number of charged ions and even for the smallest size transistors there are still on average >100 ions. The same transistor shows none or just one or two electron trapping centers. The numbers don't fit. The trapping centers also follow a Poisson distribution and follow an area scaling law.

Andricciola and Tuinhout [325] shows that the threshold-voltage mismatch and the relative current factor mismatch are hardly affected by temperature.

In deep-submicron 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 hypothesis for the calculation of the threshold mismatch of Eq. 11.8. An additional term can be included for the variation due to the pocket implant:

$$\sigma^2_{\Delta V_\mathrm{T}} = \frac{A^2_{V_\mathrm{T}}}{WL} + \frac{B^2_{V_\mathrm{T}}}{f(W,L)}, \tag{11.16}$$

where the function $f(W,L)$ of the width and length still needs to be established [326].

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:

$$\frac{\sigma^2_{\Delta\beta}}{\beta^2} = \frac{\sigma^2_{\Delta W}}{W^2} + \frac{\sigma^2_{\Delta L}}{L^2} + \frac{\sigma^2_{\Delta C_\mathrm{ox}}}{C^2_\mathrm{ox}} + \frac{\sigma^2_{\Delta\mu n}}{\mu^2_n}. \tag{11.17}$$

The mismatch-generating processes for the gate oxide and the mobility are treated in accordance with Eq. 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 $\sigma^2(W) \propto 1/L$:

$$\frac{\sigma^2_{\Delta\beta}}{\beta^2} = \frac{A^2_W}{W^2L} + \frac{A^2_L}{WL^2} + \frac{A^2_\mu}{WL} + \frac{A^2_{C_\mathrm{ox}}}{WL} \approx \frac{A^2_\beta}{WL}, \tag{11.18}$$

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 [321] indicate some role for edge roughness. This role is in measurements hard to identify in the presence of large threshold mismatch. In [31] 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 Eq. 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 Eq. 11.18 is according to [325] much less sensitive to temperature.

Considering only the threshold and current factor variations,[7] 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 [11],

$$\sigma_{\Delta I}^2 = \left(\frac{\mathrm{d}I}{\mathrm{d}V_\mathrm{T}}\right)^2 \sigma_{\Delta V_\mathrm{T}}^2 + \left(\frac{\mathrm{d}I}{\mathrm{d}\beta}\right)^2 \sigma_{\Delta\beta}^2 . \tag{11.19}$$

For strong inversion this equation is combined with the simple square-law current model:

$$\left(\frac{\sigma_{\Delta I}}{I}\right)^2 = \left(\frac{2\sigma_{\Delta V_\mathrm{T}}}{V_\mathrm{GS} - V_\mathrm{T}}\right)^2 + \left(\frac{\sigma_{\Delta\beta}}{\beta}\right)^2 . \tag{11.20}$$

At higher gate voltages the denominator term $V_\mathrm{GS} - V_\mathrm{T}$ will reduce the contribution of the threshold mismatch in the total current. It is advisable to use high gate drive voltages to minimize current mismatch. In a circuit with constant current levels, the denominator term $V_\mathrm{GS} - V_\mathrm{T}$ of Eq. 11.20 will increase with rising temperatures. Under the assumption that the threshold mismatch and the relative beta mismatch remain constant with rising temperature [325], this will result in less the current mismatch. 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. Applying the Taylor approximation [11] gives the mean current and the variance of the resulting log-normal distribution for the current:

$$\mu_I = I(\sigma_{\Delta V_\mathrm{T}} = 0)\left(1 + \frac{1}{2}\left(\frac{q\sigma_{\Delta V_\mathrm{T}}}{mkT}\right)^2\right) \tag{11.21}$$

$$\left(\frac{\sigma_{\Delta I}}{I}\right)^2 = \left(\frac{q\sigma_{\Delta V_\mathrm{T}}}{mkT}\right)^2 . \tag{11.22}$$

Note that the mean value is larger than the nominal current without mismatch.

Figure 11.26 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 Eq. 11.20 dominates. At lower gate-source voltages the threshold-related term in this equation gains importance.

[7] The contribution of mobility reduction factor θ and source drain series resistance are next in line.

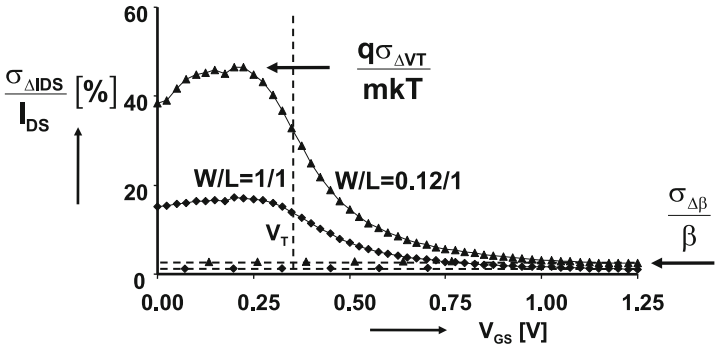


Fig. 11.26 The relative current mismatch for two 65-nm technology transistor geometries swept over the full voltage range. Measurements by N. Wils

In a 65-nm process $A_{V_T} = 3.5\,\text{mV}\mu\text{m}$. So $\sigma_{V_T} = 10\,\text{mV}$ for a 0.12/1 device. Equation 11.21 predicts a relative current mismatch of 40%, which is confirmed by the measurement in Fig. 11.26. Except for extremely low-current densities where the depletion layer width is shrinking, the observation can be made that the same value for $\sigma_{\Delta V_T}$ applies for both the strong and the weak inversion regimes. This example shows that the operating regime where the mismatch parameters of the transistor are extracted has a marginal effect on the accuracy[8] of the prediction in other regimes, as confirmed in, e.g., [318, 319, 327].

The standard deviation of the current difference between the 0.12/1 µm transistors reaches ≈40% of the current in the subthreshold regime. Obviously this would 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. 11.27 the threshold mismatch coefficient A_{V_T} is plotted as a function of the nominal oxide thickness. As predicted by Eq. 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 are caused by the compensating implants: the *N*- and PMOS transistor threshold

[8] Accuracy means that the standard deviation of a circuit parameter is within 10% of the prediction; see Sect. 11.4.6.

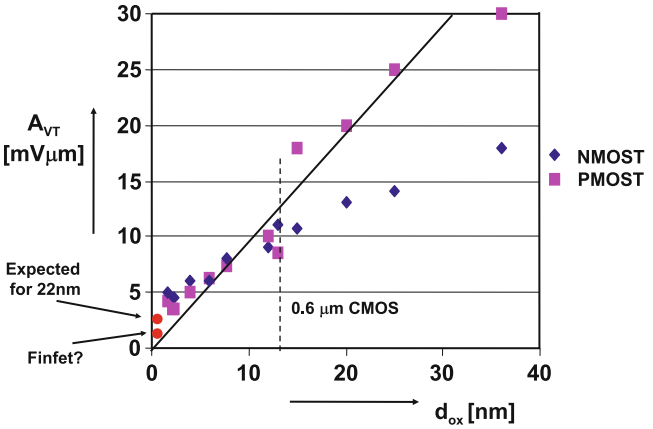


Fig. 11.27 Development of the threshold mismatch factor A_{V_T} for NMOS and PMOS transistors as a function of the nominal oxide thickness of the process. The processes span 28-nm up to 1.6 µm CMOS

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. 11.27 the diagonal line indicates an A_{V_T} factor increase of 1 mVµm 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 and 2% µm.

The first indications of 28-nm high-*k* factor metal-gate processes are favorable. With $A_{V_T} \approx 2$ mVµm the trend is continued; see too Fig. 8.29. Nevertheless some concern is present on the increased threshold gradients in these processes. Figure 11.28 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.

The analysis of matching in various processes allows to compare in Fig. 11.29 the MOS mismatch to the development of power supply voltage. A transistor with a size of $1.5L^2_{\text{min}}$ 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

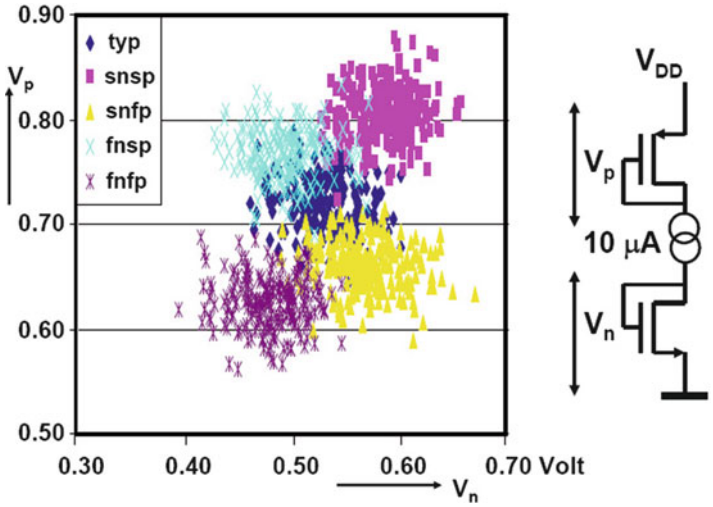


Fig. 11.28 Simulation of 200 0.2/0.1 P- and NMOS transistors in their 90-nm process corners. The notation "snfp" refers to slow NMOS and fast PMOS transistors. The variation due to mismatch is of equal importance as the process variation

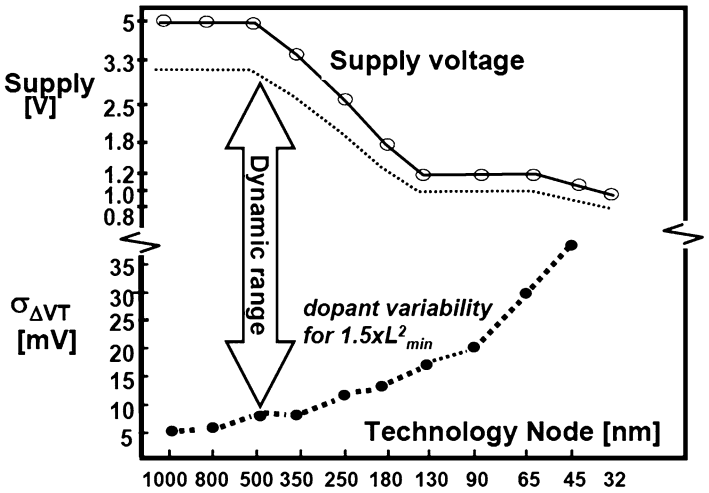


Fig. 11.29 Development of power supply voltage and the measured NMOS threshold matching of a transistor 1.5 times the minimum size through various process generations

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

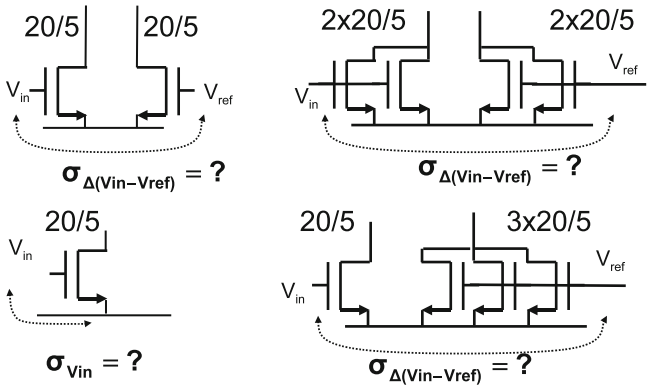


Fig. 11.30 Four mismatch situations

power supply voltage was lowered. On the other hand, the need to tailor the internal 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_{V_T} 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 submicron 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 submicron 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.

Example 11.5. In Fig. 11.30 four transistors circuits are given. Calculate the mismatch if $A_{V_T} = 10\,\text{mV}\mu\text{m}$.

Solution. The upper left circuit contains 2 transistors each 20/5 µm. As A_{V_T} has been defined for the difference between two transistors, the standard deviation can directly be calculated: $\sigma_{\text{Vin}-\text{Vref}} = A_{V_T}/\sqrt{WL} = 10/\sqrt{5 \times 20} = 1\,\text{mV}$. The top-right circuit in Fig. 11.30 contains twice the number of components; based on the observation that M parallel circuits will show a $1/\sqrt{M}$ reduction in standard deviation, the solution is found as $\sigma_{\text{Vin}-\text{Vref}} = 0.7\,\text{mV}$. The same result is obtained by summing up the total area, and use the standard formula.

The lower left transistor is not part of a pair. Nevertheless a standard deviation of its threshold voltage (or any derived circuit parameter) can be specified by considering that two of these transistor in series via their sources form the first circuit. As the standard deviations of two transistors are mutually uncorrelated, each must amount 0.7 mV in order to get 1 mV for a series connection.

Table 11.9 An overview of matching models and value ranges

MOS transistors	$\sigma_{\Delta V_\mathrm{T}} = \frac{A_{V_\mathrm{T}}}{\sqrt{WL}}$	$A_{V_\mathrm{T}} = 1\,\mathrm{mV\mu m/nm}$
MOS transistors	$\frac{\sigma_{\Delta\beta}}{\beta} = \frac{A_\beta}{\sqrt{WL}}$	$A_\beta = 1\text{–}2\%\,\mu\mathrm{m}$
Bipolar transistors (BJT)	$\sigma_{\Delta V_\mathrm{be}} = \frac{A_{V_\mathrm{be}}}{\sqrt{WL}}$	$A_\mathrm{be} = 0.3\,\mathrm{mV\mu m}$
Diffused/poly resistors	$\frac{\sigma_{\Delta R}}{R} = \frac{A_R}{\sqrt{WL}}$	$A_R = 0.5/5\%\,\mu\mathrm{m}$
Plate, fringe capacitors	$\frac{\sigma_{\Delta C}}{C} = \frac{A_R}{\sqrt{C}}$	$A_C = 0.3\%\sqrt{\mathrm{fF}}$

The last circuit uses the previous result. Now the standard deviation of a single transistor must be added to the parallel connection of three. With the previous result in mind, this is a series connection of a single device and three devices in parallel, so: $\sigma_{\mathrm{Vin-Vref}} = \sqrt{0.7^2 + 0.7^2/3} = 0.81\,\mathrm{mV}$.

11.4.5 Application to Other Components

In Table 11.9 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 [320] 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

$$\frac{\sigma^2_{\Delta R}}{R^2} = \frac{A^2_W}{W^2 L} + \frac{A^2_L}{W L^2} + \frac{A^2_\mu}{WL} \approx \frac{A^2_R}{WL}. \tag{11.23}$$

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 is the head-end connections [314, 315]. 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 [25] mismatch of capacitors was attributed to edge effects and to area nonuniformities. The first error would result in a line variation $\sigma_{\Delta C} \propto \sqrt{L}, \sqrt{W}$. The area nonuniformities 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 11.9 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. 6.8. 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 [31] the mismatch contributions of the back-bias factor and the mobility reduction coefficient have been added. Also other authors [317, 323] 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 [328]. For example 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 [317] 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 then reiterates 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 [329] 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 [329, Fig. 5]. This observation is in line with, e.g., [330].

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. 6.2.4 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. 6.8 the measured output voltages of a few hundred devices are 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 Eq. 11.15.

A second example is shown in Fig. 11.31. The differential nonlinearity 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 of the input pair of the comparators is the dominant contributor to DNL. 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.

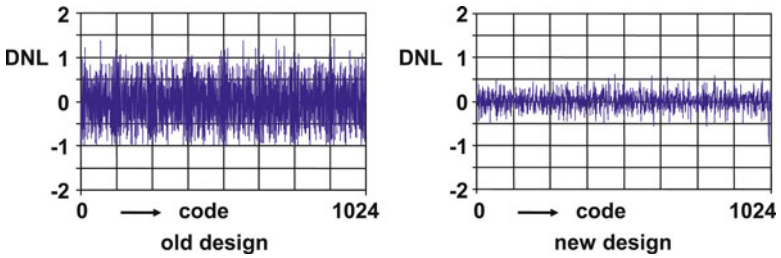


Fig. 11.31 The differential nonlinearity is a quality parameter of this 10-bit analog-to-digital converter

The consequences of mismatch on the yield of analog-to-digital converters are discussed in Fig. 8.27 [236]. 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., [331]. The idea that accuracy must be balanced against power can be easily understood by considering that the threshold voltage uncertainty in combination with the gate capacitance can be described as an energy term [332, 333] (Sect. 11.5.4).

Example 11.6. 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?

Solution. Inspecting Eq. 11.15 shows that the $\sigma_{\Delta V_T}$ is proportional to the oxide thickness. As $\sigma_{\Delta V_T}$ is of direct impact on the DNL, a reduction of the oxide thickness by a factor 2 will improve the DNL with a factor 2. A reduction of both the oxide thickness and the signal amplitude by a factor 2 (so-called constant-field scaling), does not change the DNL!

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. 11.32 and Table 11.10 [334]. 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 of both rows of inverters, a random skew will exist in the difference of arrival time between various samples of this circuit. In Table 11.10 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.

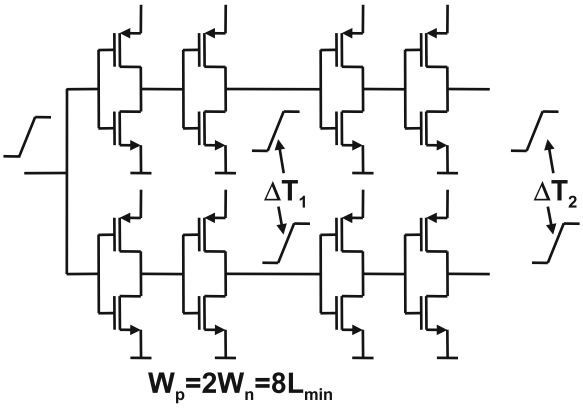


Fig. 11.32 An input pulse is applied to two chains of inverters. Due to mismatches in the transistors, there is a time difference at the two outputs. This example mimics a critical timing path

Table 11.10 The simulated standard deviation of the difference in arrival time of the two pulses in the inverter chain of Fig. 11.32

Process node	0.25 µm	0.18 µm	0.13 µm	90 nm	65 nm
Clock period (ns)	10	5	2	1	0.5
$\sigma_{\Delta T2}$ C_{load}=50 fF (ps)	16	21	38	68	88
$\sigma_{\Delta T2}$ C_{load}=50$\cdots$15 fF (ps)	16	16	22	33	32

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 [335]. 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.

Example 11.7. Three circuits in Fig. 11.33 experience threshold-voltage mismatch. Discuss how to reduce the effect of threshold mismatch for each case. Consider to increase or decrease the gate width or gate length.

Solution. The left circuit in Fig. 11.33 shows a typical input stage for an opamp or comparator. Crucial is the input-referred mismatch, which is caused by the threshold standard deviation. As this standard deviation is equal to $A_{V_T}/\sqrt{WL}$ the area must

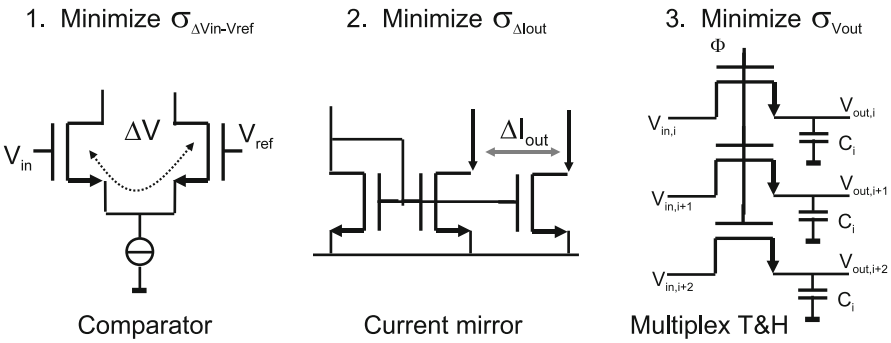


Fig. 11.33 Three circuits with mismatch

be increased to reduce the mismatch. In most cases a large input transconductance is desired and the preferred option is to increase the gate width.

In the middle circuit the standard deviation of the output current must be decreased. Now it is important to realize that

$$\sigma_{\Delta I} = g_m \sigma_{\Delta V_\mathrm{T}} = \sqrt{\frac{2W\beta_\square I}{L}} \frac{A_{V_\mathrm{T}}}{\sqrt{WL}}.$$

Rearranging terms makes clear that in first order increasing the gate width does not affect the result as the increase reduces the threshold mismatch but amplifies the remainder. So the solution is to increase the gate length: the mismatch is reduced and the transconductance reduced.

In the last circuit three parallel track-and-hold circuits are implemented. It is crucial that all three track-and-hold circuits show the same output voltage for equal input voltage. The pedestal step due to the channel charge injection on the hold capacitor must be small and should not fluctuate per section. The standard deviation of the channel charge is $\sigma_{Q\mathrm{chan}} = C_\mathrm{gate}\sigma_{\Delta V_\mathrm{T}} = C_\mathrm{ox} WL \times A_{V_\mathrm{T}}/\sqrt{WL}$. Evaluating the terms shows that $\sigma_{Q\mathrm{chan}}$ is minimum for the lowest gate area WL. So here the gate length is certainly minimized to reduce the charge uncertainty and obtain the highest speed. The gate width is chosen at a value where the distortion is acceptable.

11.5.3 Drift

In literature the term "drift" refers to various phenomena in electronic 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:

$$\Delta V_{\text{in}} = \frac{I_d}{g_m}\frac{\Delta\beta}{\beta}. \tag{11.24}$$

Andricciola and Tuinhout [325] shows 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 Eq. 2.120 an expression for the temperature dependent drift at the input of a differential pair due to current factor mismatch is found:

$$\text{Drift} = \frac{\mathrm{d}\Delta V_{\text{in}}}{\mathrm{d}T} = \frac{\alpha}{T}\frac{I_d}{g_m}\frac{\Delta\beta}{\beta} = \frac{\alpha}{T}\frac{V_{\text{GS}}-V_{\text{T}}}{2}\frac{\Delta\beta}{\beta} = \frac{\alpha\Delta V_{\text{in}}}{T}. \tag{11.25}$$

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 [336] for ΣΔ 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, and 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:

Capacitive load	Threshold variance
$C_{\text{gate}} = WLC_{\text{ox}}$	$\sigma_{\Delta V_{\text{T}}} = \dfrac{A_{V_{\text{T}}}}{\sqrt{WL}}.$ (11.26)

The voltage uncertainty on the gate capacitance can be described as an energy term [332, 333]:

$$E_{\sigma V_{\text{T}}} = C_{\text{gate}} \sigma^2_{\Delta V_{\text{T}}} = C_{\text{ox}} A^2_{V_{\text{T}}} = 3 \times 10^{-19} \quad \text{Joule} \tag{11.27}$$

which is independent of the transistor size and corresponds to about 100 kT at room temperature. A similar analysis holds for flash converters in Fig. 8.29. This energy can be seen as the energy required to toggle a latch pair of transistors in metastable 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.

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 electronic circuitry. In the 1970s and 1980s, the dominant form of packaging was the "dual-in-line" package (DIL or DIP); see Fig. 11.34 top. The die is attached to a lead frame and two rows at a pitch of 2.54 mm. Typical packages are made of ceramic or plastic material and count 14, 16, 24, 40, and 64 pins. The pins or leads are inserted in the printed-circuit board via through-holes (Fig. 11.35). Ceramic packages are expensive and mostly used in a prototype phase, while plastic-molded packages 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.

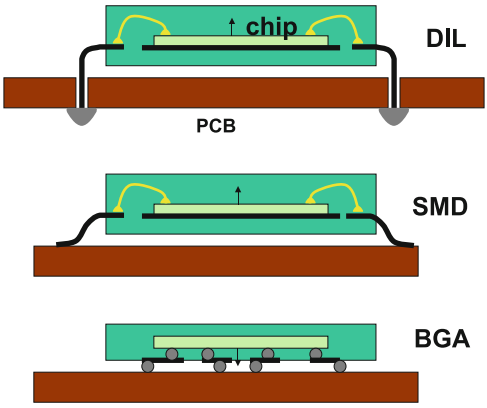


Fig. 11.34 Three types of packages: dual-in-line, surface-mounted device, and bump-bonded device

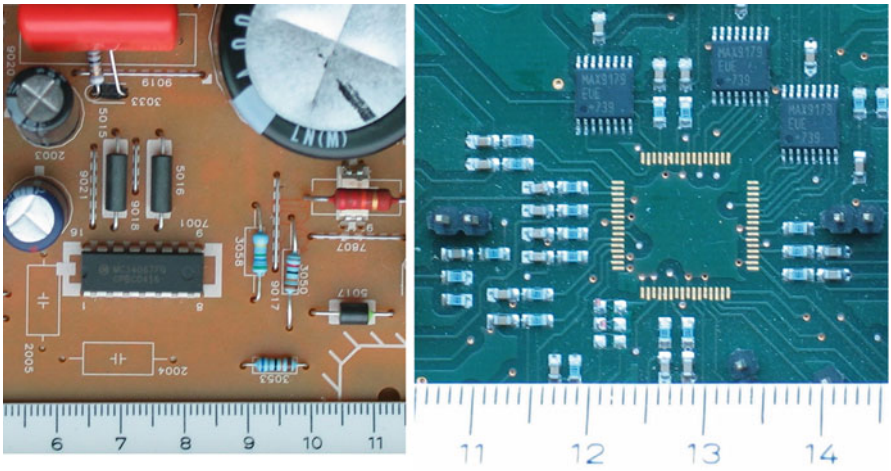


Fig. 11.35 *Left*: A printed-circuit board with through-hole components. *Right*: An example of surface-mounted PCB. In the *middle* the pads on which a 64-pin quad flat pack will be soldered are visible

- 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 11.11. Pin grid arrays use the entire bottom of the package for connecting hundreds of signals to the printed-circuit board.

Table 11.11 Commonly used packaging abbreviations

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

Often these terms are followed by the number of pins

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. 11.35(right). 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. 11.34 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. 11.34(bottom) the die is placed upside down and a special redistribution pattern is used to move the bond pads 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:

$$\theta_{JA} = \frac{\text{Junction temperature} - \text{Ambient temperature}}{\text{Power}}. \tag{11.28}$$

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 layout, the airflow, 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. 11.36. The bondpad consists of a top-level metal flap of dimensions between $70 \times 70\,\mu\text{m}^2$ for advanced bonding and $130 \times 130\,\mu\text{m}^2$ for somewhat older processes. Next to this connection electrode, protection structures are used to avoid electrostatic discharges (ESD) that could damage the connected circuitry. These

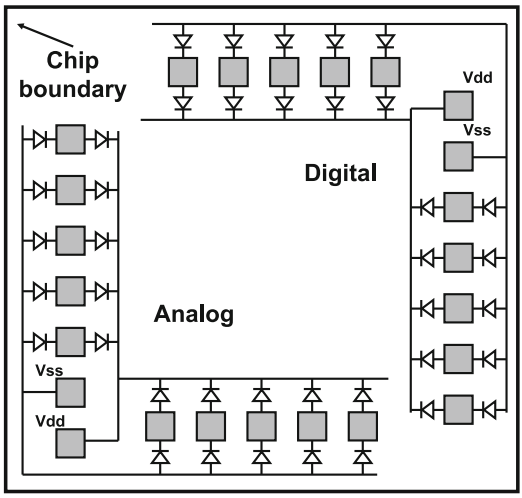


Fig. 11.36 A schematic layout of the bondpads for analog and digital and their protection devices

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 in the bond pad via a 1,500 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 nonlinear nature of the diodes in combination with a high source impedance can create distortion. For example, 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 11.37 and Table 11.12 show a simple equivalent circuit diagram and some data of the components.

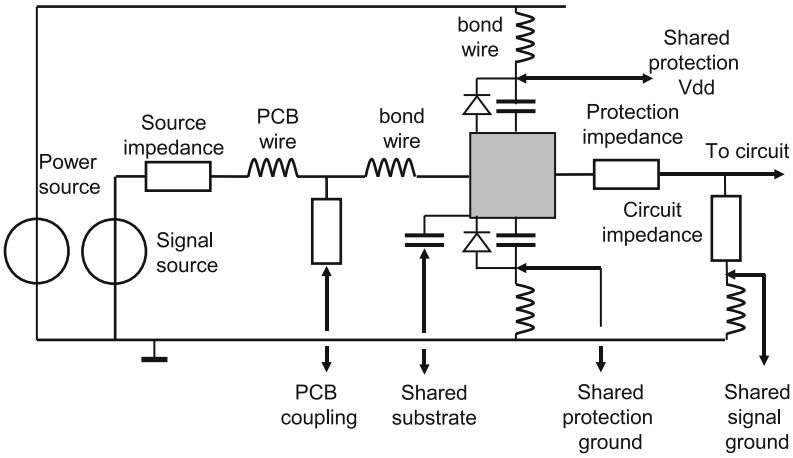


Fig. 11.37 Equivalent diagram showing the most important parasitic elements in connecting a signal to the circuit on the IC

Table 11.12 Some values for the components in the equivalent circuit diagram of Fig. 11.37

Bond wire	$L \approx 1$ nH/mm, $L = 1$ nH(CSP)-10 nH(DIL)
Bond wire	$R \approx 0.3$–$1\,\Omega$
Bondpad 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

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 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 and jitter.

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.

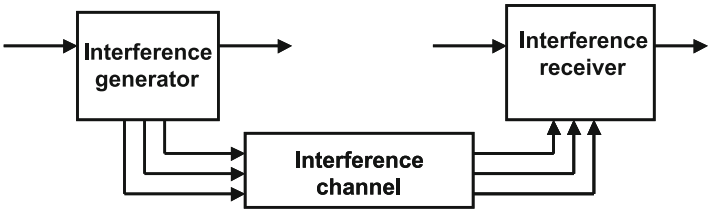


Fig. 11.38 A general model for describing substrate noise generation, transport, and effect

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 11.38 shows the generally accepted model for analyzing and describing interference.[9] 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 and a processor. Also analog cells can act as generator, e.g., an oscillator.

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
- Layout: symmetrical shapes, differential signaling, etc.
- Parasitic components: junctions and 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 11.39 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

[9]This research was carried out together with Sergei Kapora and Jacob Bakker. The model was suggested by Joost Briaire.

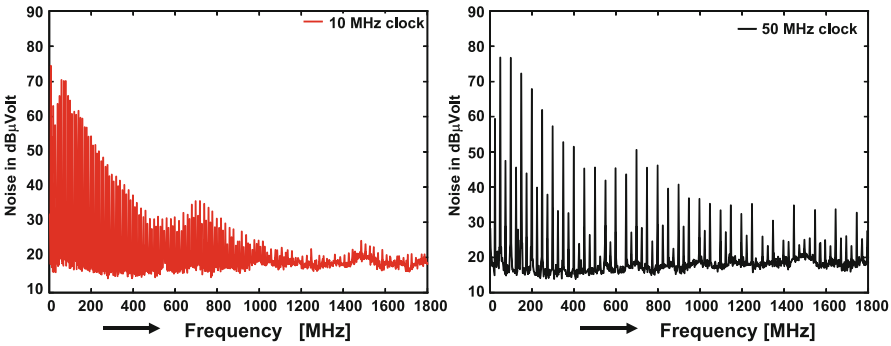


Fig. 11.39 A spectrum of the substrate potential measured close to a 0.5 mm^2 65-nm CMOS digital shift-register block. The clock runs at 10 and 50 MHz, and the input is a half-clock signal. The vertical axis is in dBμV: 60 dBμV equals 1 mV (courtesy: S. Kapora and J. Bakker, NXP)

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.

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, layout: symmetrical shapes, differential signaling, etc., parasitic components: junctions and coupling capacitors, substrate connectivity: location of substrate contacts and density, guard rings, deep well isolation, and other technological options, and power wiring: resistivity.

An error is easily made in substrate noise prevention. Figure 11.40 shows a transistor where the designer has attempted to connect the substrate to the ground

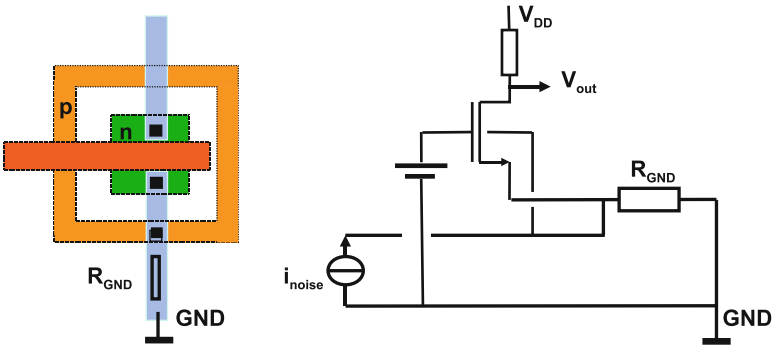


Fig. 11.40 *Left*: A layout of a substrate noise-sensitive NMOS transistor surrounded by a p-doped substrate connection. *Right*: The equivalent circuit diagram

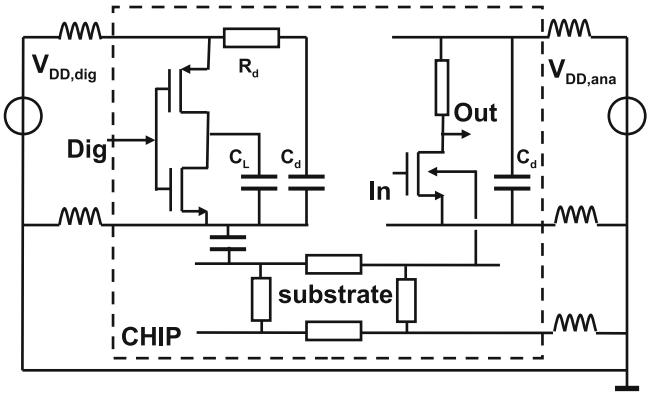


Fig. 11.41 The most simple form of coupling: on the *left* is an inverter whose switching noise propagates via the substrate to the analog amplifier

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 bond wire) 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 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 11.41 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 bond wire 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.
- 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 for external noise. And on the other hand the internal power supply is stabilized.
- Never connect a decoupling capacitor to "just a ground node." Think carefully between which voltage nodes decoupling must be applied. Use all free space for decoupling capacitors.

The path from noise generator to noise receiver is modeled in Fig. 11.42. 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.

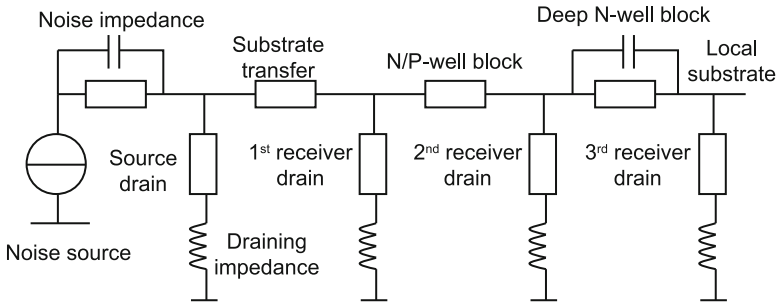


Fig. 11.42 An equivalent circuit diagram of the noise propagation route

- 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 bond pads influence the substrate noise in two ways: first the connections have series inductance and resistance, and secondly the bond pad forms a capacitive coupling to the substrate.

The values of the components in Fig. 11.42 depend on the technology and the layout. Although every layout is specific some orders of magnitude can be estimated. A typical path through the substrate is of the order of 200–1,000 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.

Exercises

11.1. Explain what the consequences are of power supply reduction for analog circuits.

11.2. What is the reason in Example 11.3 on page 499 that the magnitude of the distortion decreases at higher signal frequencies.

11.3. The perimeter of a square unit capacitor contributes between 8% and 16% depending on the batch processing (global variation). What is the accuracy of a 1:1.5 ratio, if this ratio is obtained by stretching one side of a capacitor.

11.4. An experimental current-steering digital-to-analog converter is based on a radix $e = 2.71\ldots$. Design four current sources with ratios: $1 : e : e^2 : e^3$. The mask grid allows only multiples of $0.1\,\mu\text{m}$ and the structure cannot exceed an area of $100\,\mu\text{m}^2$.

11.5. In Example 11.4 on page 506 the currents in the individual transistors is 1 mA. The square resistance of the metal-1 is $0.1\,\Omega$. Calculate the input-referred mismatch.

11.6. Modify the layout of Fig. 11.22 in such a way that the sources are all toward the center and the drains are on the border of the structure. Is this an advantageous modification?

11.7. An SRAM latch with $W/L = 0.2/0.1$ transistor sizes for all transistors in 65-nm CMOS is designed. How much differential voltage is needed to reach a 6-σ probability that the latch toggles correctly from its metastable position.

11.8. An array of 64 NMOS current sources $W/L = 10/5$ in 90-nm CMOS is designed. A drive voltage of $V_{GS} - V_T = 200\,\text{mV}$ is applied. Calculate the standard deviation of the output current.

11.9. There is an area of $10\,\mu\text{m}^2$ available to generate from a perfect voltage source a current with minimum spread. Make a choice of components in $0.18\,\mu\text{m}$ CMOS from Tables 2.16 and 11.8. Give an estimate of the achievable accuracy.

11.10. In the previous exercise there is also an accurate clock signal available, enabling the use of capacitors defined in Table 2.19. Does the picture change?

11.11. Discuss the limits to the proposed length and width changes in Example 11.7 on page 523.

11.12. Discuss series resistors in power supply lines. Are they useful for decoupling?

Chapter 12
System Aspects of Conversion

Abstract The analog-to-digital and the digital-to-analog converter are components in larger systems. This chapter discusses a number of points related to the system interaction. Aspects of signal processing strategy play an important role for the application of an analog-to-digital converter. A special point is the input interfacing. Several input structures are reconsidered on a system level.

The use of a figure of merit helps to create a uniform scheme for various converters with different resolutions, bandwidth, and powers. In a second step the figure of merit is used to estimate the power consumption in a system. Although the results of this approach requires to consider the results with large margins, it can help to estimate the validity of system partitioning.

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.). Table 12.1 lists various types of analog-to-digital converters and their application in electronic systems. The CMOS signal-processing chain formed by analog-to-digital converters, DSPs and ASICs, is controlled by CPUs or microcontroller 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 electromagnetic 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

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4_12,

© Springer Science+Business Media, LLC 2013

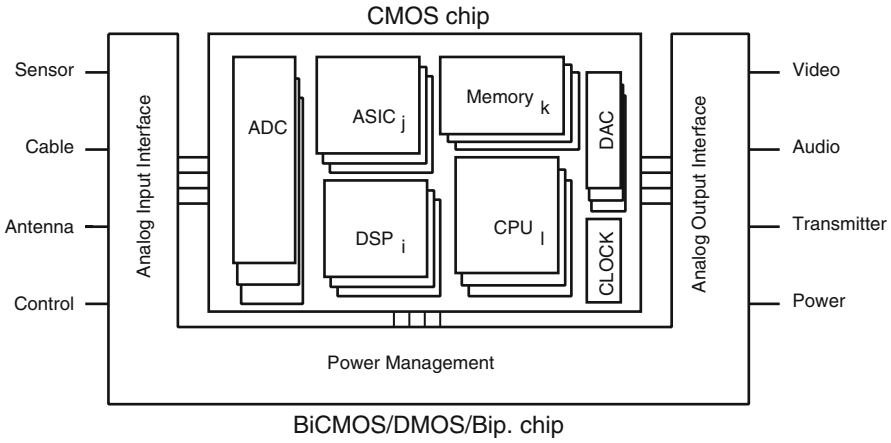


Fig. 12.1 General set-up of an electronic system-on-chip

Table 12.1 Some typical application areas for analog-to-digital conversion

Digital-to-analog architecture	Application	N, BW
Unary resistor string	In flash ADCs	6 bit, >1 GHz
Binary resistor R-2R	Voltage setting in ICs	6 bit, kHz
Current steering	Video, games, transmitter	10–15 bit, >1 GHz
Capacitor array	Sensors, low-power IC	8–10 bit, 10 MHz
Analog-to-digital architecture	**Application**	**N, BW**
Flash	Fast binary signals in communication and interface	6 bit, 3 GHz
SAR	Input ADC for processors	8–10 bit, 1–5 MHz
Multiplexed SAR	Wide-band data acquisition	8–10 bit, 1 GHz
Pipeline	Data acquisition for communication, medical	10–12 bit, 100 MHz
Dual slope	Harsh conditions, monitoring chemical processes	8–10 bit, 1–10 kHz
Counting	Image sensors	10–12 bit, 100 kHz
1-bit sigma–delta	Audio, seismic	16–22 bit, 300Hz–50 kHz
Multi-bit sigma–delta	Communication systems	10–12 bit, 2–100 MHz

dissipation. This requirement and the need for longer lifetime in portable equipment have 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 [20]. This has led to a significant power reduction in the digital domain. The technological advances

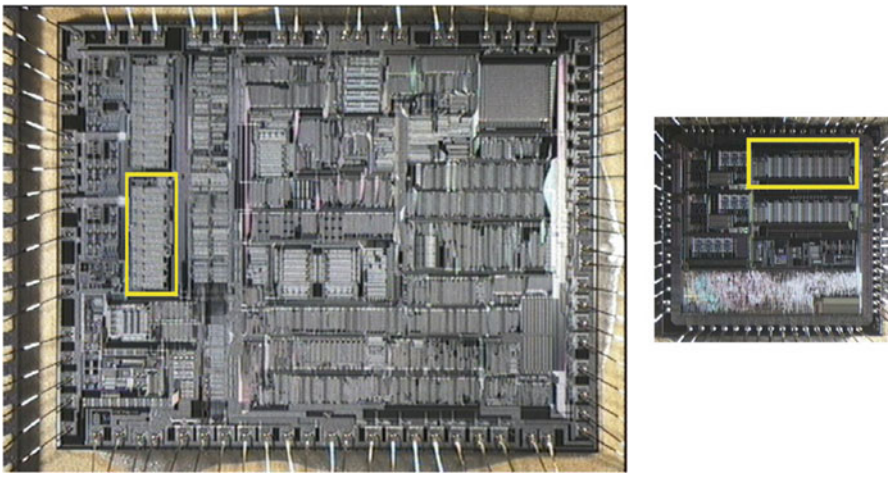


Fig. 12.2 Evolution of digital television processing, *left*: a die in 1 µm CMOS (1995), and *right*: a die with equivalent processing power in 0.35 µm CMOS (2002). The digital circuitry has shrunk by a factor of 10, while the analog-to-digital converter (*light box*) is approximately of the same size (courtesy: Philips Semiconductors Hamburg)

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 and their support circuits such as variable-gain amplifiers, references, timing circuits, and the postprocessing filters present a contribution typically 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.

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

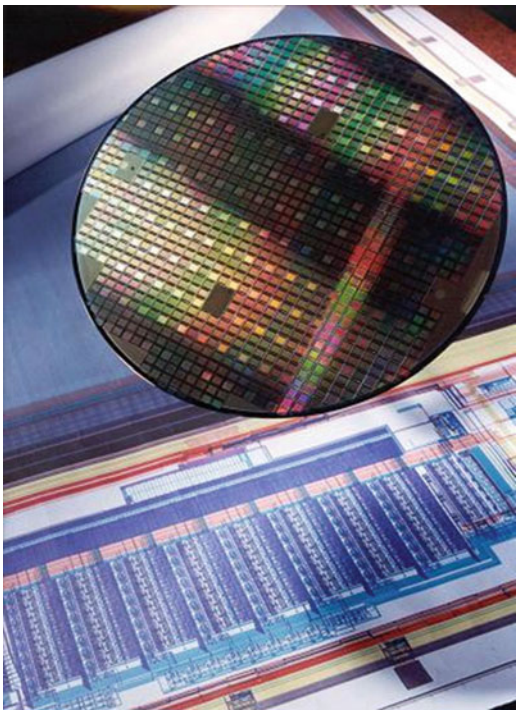


Fig. 12.3 An analog-to-digital converter (*on lay-out*) is an important part of this CMOS image sensor (*on wafer*). Photo is courtesy of the philips semiconductor imaging sensor group

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?
- 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 12.2 compares a few systems and their requirements on converters.

Table 12.2 Some system requirements on analog-to-digital conversion

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, base station	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

Here, some indicative data is given. Implementation choices can lead to shifts in requirements

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 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 amplitude, 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 specifications within one class of systems may look alike in Table 12.2, yet strong differences can exist. The video CCD image-sensor analog-to-digital converter requires a relatively high degree of accuracy at low frequencies in order to match the CCD dynamic range quality. Data transmission for binary video

signals needs some 6–7 effective bits, but requires proper conversion at the high frequencies. Even a straightforward application such as in a television can create requirements that 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 performance up to 10–12 MHz signal frequencies for (QAM) HDTV-like applications. Optimization balancing these demands is needed to avoid excessive power consumption.

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 important aspect of this optimization is flexibility. Unlike specification tolerances, flexibility demands the adaptation of the system to predetermined (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 and 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 (SNR), 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 ($f_s > 2$ BW). For high-bandwidth circuits the sampling frequency is in practice 20–50 % higher than the minimum rate required by the Nyquist theorem; see Fig. 3.9. 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 12.3. 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

Table 12.3 Balancing between the digital drawbacks 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

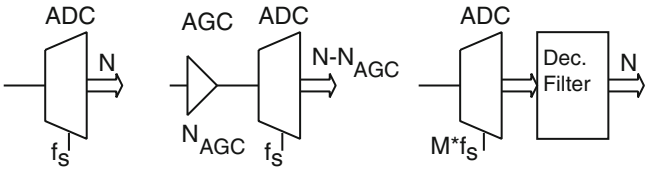


Fig. 12.4 (**a**) Standard analog-to-digital converter, (**b**) analog-to-digital converter with gain control, and (**c**) reduced resolution analog-to-digital converter with oversampling and decimation filter

chain. As 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 SNR 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. 12.4. 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 nonuniform 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. Figure 12.5 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, the USA, 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 (USA)). 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

[1]The arguments here presented apply to analog-to-digital conversion, but hold mostly also for digital-to-analog conversion.

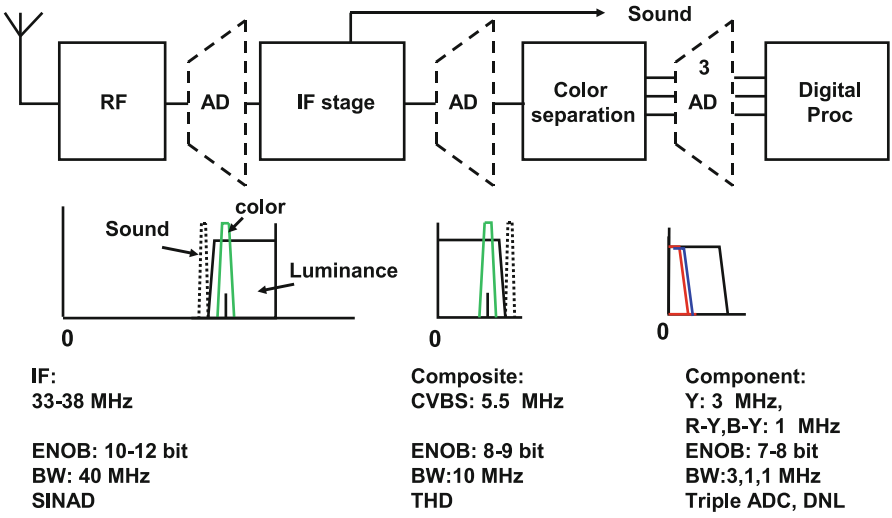


Fig. 12.5 In a television signal processing chain there are three potential positions for the analog-to-digital converter: in the IF domain, on composite signal level, or on component level. Spectra for all three positions are indicated and some global specifications of the converters

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 straightforward.

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 12.6 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 cutoff 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

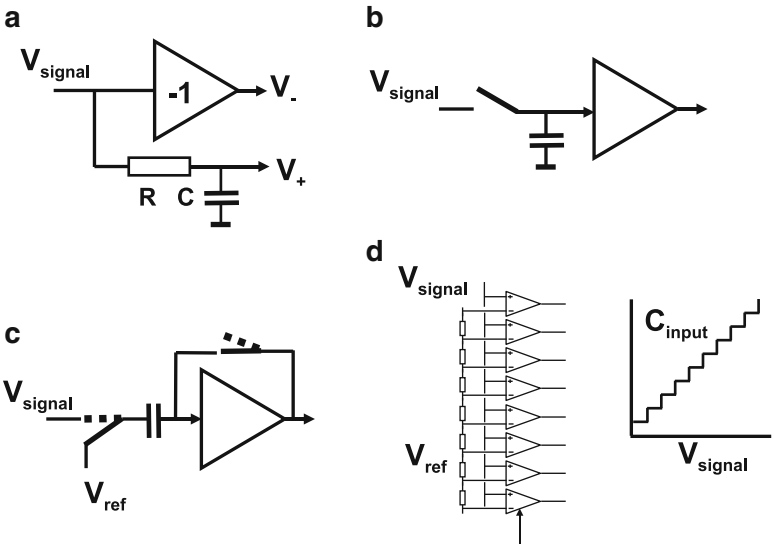


Fig. 12.6 Four examples of input structures for analog-to-digital converters

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 an 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 electromagnetic interference can enter the system via ground or reference connections.

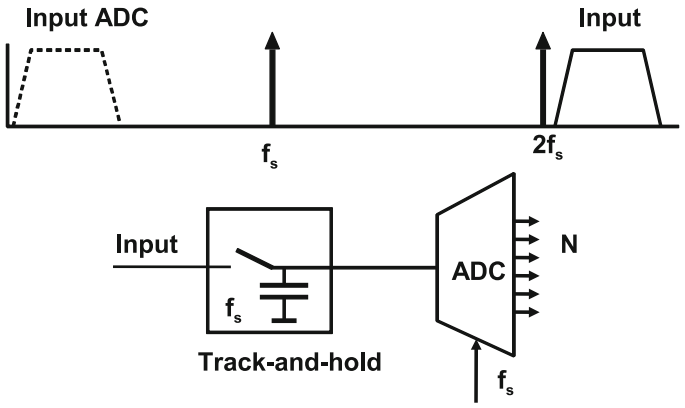


Fig. 12.7 Converter arrangement for IF conversion

- The jitter of the sampling clock should be in line with the specifications of the converter: treat the sample clock as if it were 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 3.9 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 demodulation (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).

12.1.4 Conversion of Modulated Signals

Figure 12.7 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.

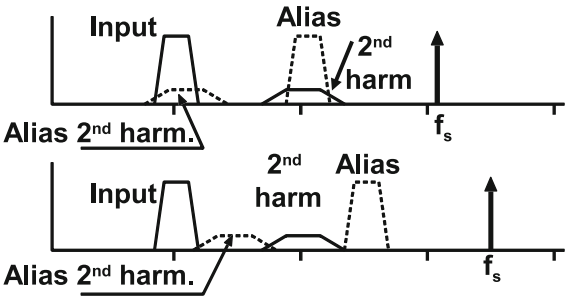


Fig. 12.8 Spectrum of a sampled modulation signal with distortion: (**a**) Sample rate equals 3.1 times the modulation frequency and (**b**) sample rate equals 3.5 times the modulation frequency

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 "subsampling." 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 difficulties: 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. 12.8). 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 necessitate 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 signal and system analysis.

12.2 Comparing Converters

Table 12.4 briefly summarizes some analog-to-digital converter characteristics and Fig. 12.9 shows various converter architectures in a resolution-bandwidth space. Figure 12.10 shows an overview of publications on flash, SAR, pipeline, and sigma–delta architectures. In recent years sigma–delta converters are gradually increasing their bandwidth into regions that were previously served by pipeline converters. If the pipeline and time-continuous sigma–delta converters published in the ISSCC in the years 2009–2012 are analyzed in more detail, then pipeline converters have a slightly lower F.o.M., but have been mostly designed in 90-nm CMOS or older. Also two BiCMOS converters are part of the comparison. Although the sigma–delta converters have a poorer power efficiency (probably even worse than reported

Table 12.4 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

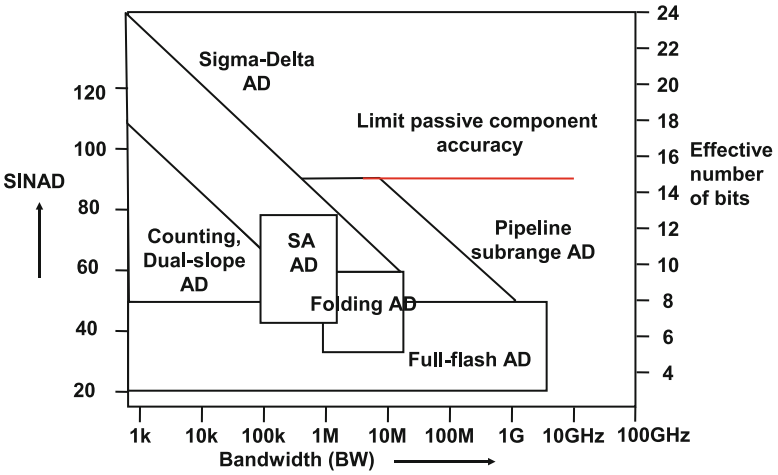


Fig. 12.9 Typical bandwidth and accuracies for various analog-to-digital converters

because mostly the decimation filter is not included), these converters are often designed in advanced CMOS nodes down to 40-nm CMOS and are more suited for Systems-on-Chip (SoCs).

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 [337, 338].

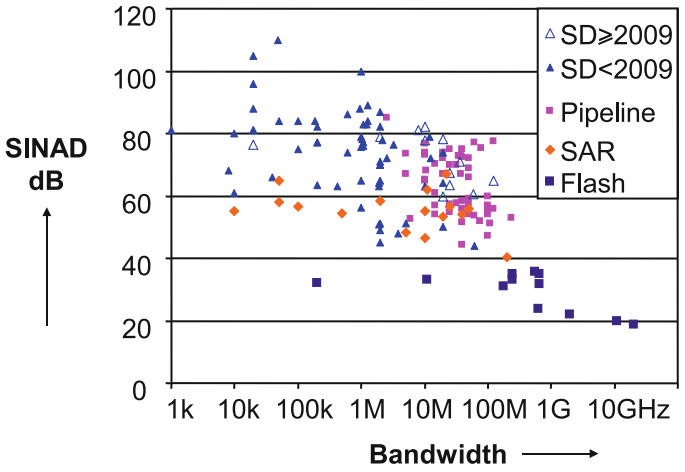


Fig. 12.10 Bandwidth and SINAD comparison for flash, SAR, pipeline, and sigma–delta converters as published on the international solid-state circuits Conference. Time-interleaved converters have been left out. More recent time-continuous sigma–delta converters enter bandwidths that were the domain of pipeline architectures (from: B. Murmann, "ADC Performance Survey 1997–2012," [Online]. Available:http://www.stanford.edu/~murmann/adcsurvey.html)

The starting point for a performance versus power efficiency comparison is the amount of signal power needed to overcome thermal noise at a certain SNR:

$$P_{\text{sig}} = 4kT\text{BW} \times \text{SNR} \tag{12.1}$$

$$\text{Power efficiency} = \frac{\text{Power consumed}}{P_{\text{sig}}} = \frac{\text{Power consumed}}{4kT\text{BWSNR}}. \tag{12.2}$$

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., [271, 336, 339].

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

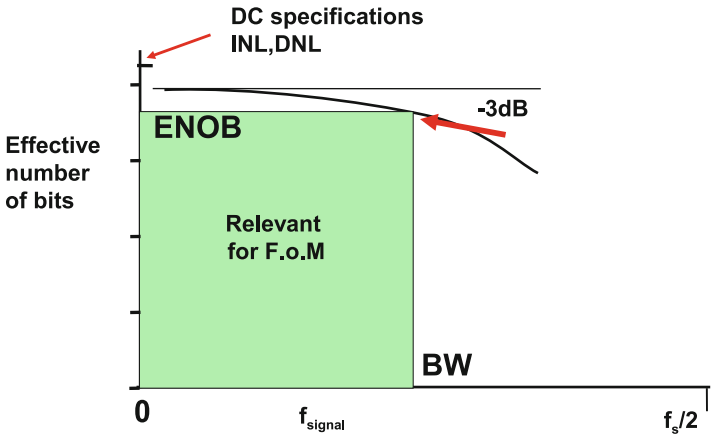


Fig. 12.11 Definition of the combination of bandwidth and resolution for the figure of merit.

measuring the efficiency of the conversion. The ENOB is a substitute for the SNR, Eq. 5.13, $\text{SNR} = \frac{3}{2}2^{2\text{ENOB}}$, leading to

$$\text{Power efficiency} = \frac{\text{Power consumed}}{4kT\text{BW}\frac{3}{2}2^{2\text{ENOB}}}. \tag{12.3}$$

However, using this substitution in the theoretical relation of Eq. 12.2 does not fit well to the available data on analog-to-digital converters. Comparing many converters as in Figs. 12.12 and 12.13, indicates that a better fit is obtained with

$$\text{F.o.M.} = \frac{\text{Power}}{2^{\text{ENOB}} \times \text{Minimum}(2\text{BW}, f_s)} \quad \text{[Energy per level]}. \tag{12.4}$$

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 $2 \times \text{BW}$ 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.

The dominant difference between the theoretical F.o.M. and the practically chosen F.o.M. is in the exponent of the SNR 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.

Example 12.1. Determine the F.o.M. for the ADC reported in [184], Fig. 12.14.

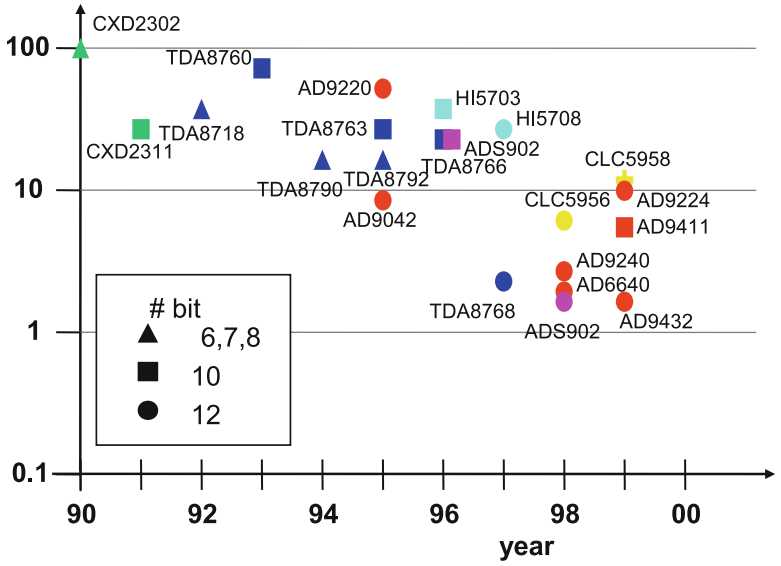


Fig. 12.12 Figure of merit for various industrial analog-to-digital converters in the time frame 1990–2000)

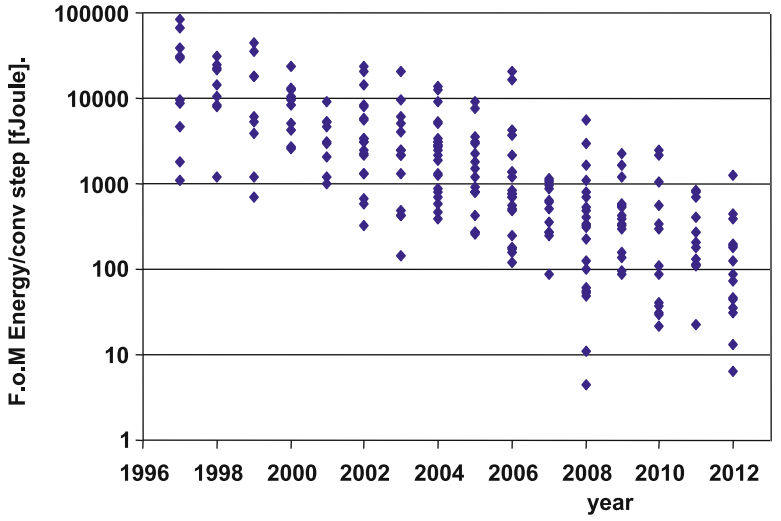


Fig. 12.13 Figure of merit for various analog-to-digital converters as published on the international solid-state circuits conference (from: B. Murmann, "ADC Performance Survey 1997–2012," [Online]. Available:http://www.stanford.edu/~murmann/adcsurvey.html)

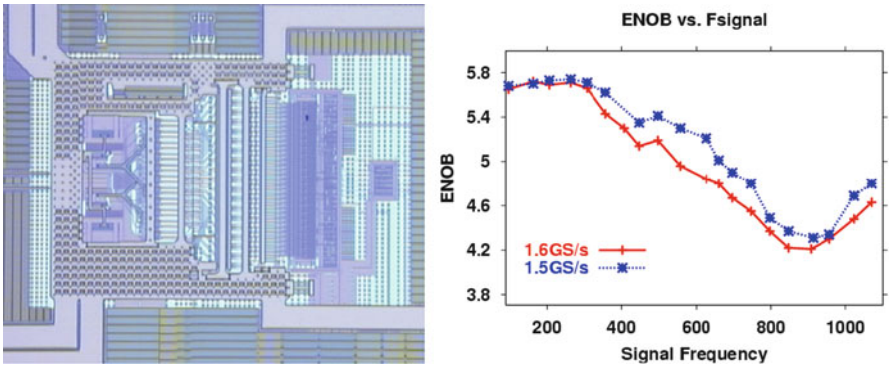


Fig. 12.14 Photograph and performance plot of a 6-bit 328-mW 1.6-Gs/s flash ADC

Table 12.5 Power efficiency of some ISSCC published analog-to-digital conveters

Year	Author	Architecture	N bit	f_s Ms/s	Power W	BW MHz	ENOB bit	F.o.M. fJ/conv
2006	Schvan [180]	Mux flash	6	24,000	1.2	12,000	3.5	4,400
2009	Brooks [215]	Pipe	12	50	0.0045	25	10.0	90
2006	Geelen [206]	Pipe	10	100	0.035	50	9.3	550
2012	Chai [198]	Pipe	10	200	0.0054	100	9.5	46
2006	Shimizu [194]	Sub-range	12	40	0.03	20	10.5	520
2007	Hsu [234]	Interleaved	11	800	0.35	400	8.7	1,100
2011	Doris [233]	SAR interleaved	10	2,600	0.48	1,300	7.6	849
2007	Craninckx[218]	SAR	9	50	0.00029	10	7.4	65
2008	Elzakker [220]	SAR	10	1	0.0000019	0.5	8.5	4
2007	Hesener [223]	SAR	14	40	0.66	0.96	13.8	24,200
2012	Harpe [221]	SAR	10	4	0.000017	2	9.7	6
2009	Naraghi [228]	Linear	9	1	0.000014	0.5	7.9	100
2007	Christen [267]	$\Sigma\Delta$	12	240	0.021	10	10.2	900
2006	Schreier [286]	BP $\Sigma\Delta$	15	264	0.375	8	12.3	4,500
2011	Bolatkale [277]	TC $\Sigma\Delta$	11	4,000	0.256	125	10.5	705
2012	Shettigar [276]	TC $\Sigma\Delta$	14	3,600	0.015	3.6	11.5	73

Some care must be taken in comparing the numbers as the measurements conditions are not identical

Solution. The reported power is 0.328 W. A suitable data point on the plot of the ENOB versus the input frequency is 5.4 bit at 400 MHz, yielding

$$\text{F.o.M.} = \frac{\text{Power}}{2^{\text{ENOB}} \times \text{Minimum}(2\text{BW}, f_s)}$$

$$\text{Energy per level} = \frac{0.328}{2^{5.4} \times 2 \times 400 \times 10^6} = 9.7 \times 10^{-12}.$$

Some more examples are given in Table 12.5.

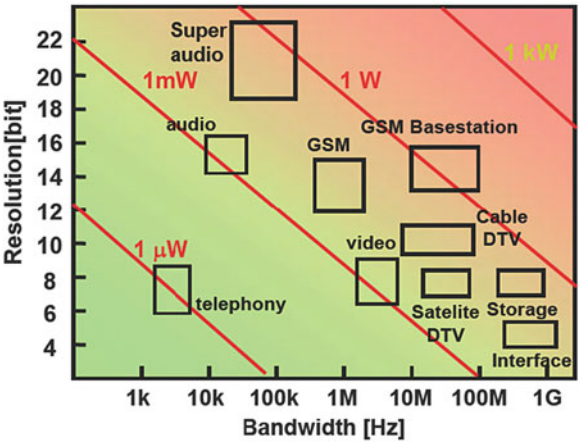


Fig. 12.15 In this bandwidth/resolution field the global specifications of some consumer and communication analog-to-digital interfaces are shown. Also a power estimation based on an F.o.M. of 1 pJ/conv step is indicated.

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, 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.

The 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:

$$\text{Estimated power} = \text{F.o.M.} \times 2\text{BW} \times 2^{\text{ENOB}}. \tag{12.5}$$

Figure 12.15 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 (8 b/3 kHz) to (12 b/200 kHz) but also costs three orders of magnitude in power consumption.

Several F.o.Ms exist to compare high-performance current-steering digital-to-analog converters; see Sect. 7.3.1. In [340] a relation between peak-peak signal

[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 meager result.

swing V_{pp}, signal frequency f, corresponding spurious-free dynamic range ratio SFDRR, and total consumed power P is proposed as F.o.M.:

$$\text{F.o.M.} = \frac{V_{pp} f \text{SFDRR}}{P} \tag{12.6}$$

where the SFDRR equals $10^{\text{SFDR}/20}$. A comparison with the F.o.M. for analog-to-digital converters in Eq. 12.4 shows that an additional term V_{pp} accounts for the power (and design problems) needed for generating the analog signal in a digital-to-analog converter.

Example 12.2. Give a power estimate for a 100 MHz bandwidth, 14-bit analog-to-digital converter.

Solution. With an F.o.M.= 1 pJ/conv step, the estimated power is F.o.M. $\times 2^{14} \times 2 \times 10^8 = 3.2$ W

12.3 Limits of Conversion

The previous metrics also allow some thoughts on the potential limits of analog-to-digital conversion. Figure 12.16 combines a number of limits that have been discussed. Thermal noise is in any system the fundamental lower limit. In Fig. 12.16 the thermal noise is plotted as formulated in the denominator of Eq. 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.

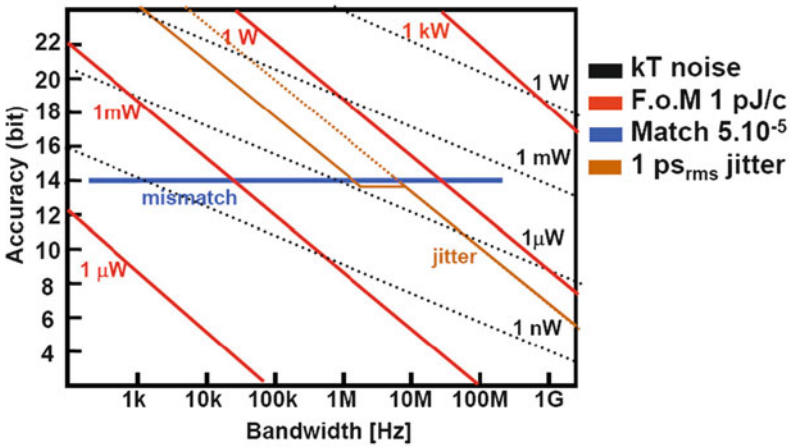


Fig. 12.16 A few limits to analog-to-digital conversion: noise, matching, and jitter are compared to the F.o.M. values

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 required. Or, component matching is simply removed from the comparison by means of (1-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. 7.30).

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

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.

12.2. Why does it make sense to have the signal amplitude in the Figure of Merit for digital-to-analog converters and not for analog-to-digital converters?

12.3. Choose one of the papers listed in the web site from the IEEE Journal of Solid-State Circuits, e.g., [99, 147, 158, 193, 202–205, 207, 209, 212, 216, 226, 241, 242] and [269, 280, 282, 284, 292, 293]. Write a 4–5-page comment on the chosen paper, dealing with:

- What system is this converter intended for, what is the critical specification point? If this is unclear from the paper, what do you think is the application area and the relevant specification?
- What novelty/improvement is claimed?
- Analyze the details of the block and circuit diagrams. Try to understand the function of every component.
- What basic type of ADC/DAC is used, and which variation is applied?
- Is that variation relevant for the specification, and why?
- What part of the AD is relevant for the critical specification? Analyze the algorithm.
- Is there an alternative way of obtaining this improvement?
- How does the technology affect the performance? Are specific issues indicated (e.g., what effect has a higher threshold, how important is matching)?
- What happens if the power supply voltage changes to 80%?

- Examine the measurement and simulation results: do these results give sufficient information to judge the effectiveness of the proposed improvement? Is important information missing from the measurements? Can you explain the frequency components?
- Determine the F.o.M.
- And anything else that you consider important in this paper. Not all points can be equally simply accessed from every paper, still try to find a suitable answer.

References

1. Seitzer D, Pretzl G, Hamdy NA (1983) Electronic analog-to-digital conversion. Wiley Interscience, New York. ISBN: 0-471-90198-9
2. van de Plassche R (1994) Integrated analog-to-digital and digital-to-analog converters. Kluwer Academic Publishers, the Netherlands. ISBN: 0-7923-9436-4 (2nd edition ISBN: 1-4020-7500-6, the Netherlands, 2003)
3. Razavi B (1994) Principles of data conversion system design. Wiley-IEEE Press, USA. ISBN: 978-0-7803-1093-3
4. Jespers P (2001) Integrated converters D-to-A and A-to-D architectures, analysis and simulation. Oxford Press, USA. ISBN: 0-19-856446-5
5. Maloberti F (2007) Data converters. Springer, Berlin. ISBN: 0-38-732485-2
6. Carmichael RD (1931) Smith ER Mathematical tables and formulas. Dover publications, New York. ISBN: 486-60111-0
7. Abramovic M, Stegun IA (eds) (1965) Handbook of mathematical functions. Dover publications, New York. ISBN: 0-486-61272-4
8. Beyer WH (1987) CRC standard mathematical tables, 28th edn. CRC press, Boca Raton. ISBN: 0-8493-0628-0
9. Sansen W (1999) Distortion in elementary transistor circuits. IEEE Trans Circ Syst II 46: 315–325
10. Rey WJJ (1983) Introduction to robust and quasi-robust statistical methods. Springer, Berlin. ISBN: 0-387-12866-2
11. Papoulis A (1965) Probability, random variables, and stochastic processes, student edn. McGrawHill, New York (4th edition ISBN: 0-073-66011-6, McGrawHill 2001)
12. Weast RC (ed) (1984) CRC handbook of chemstry and physics, 64th edn. CRC Press, Boca Raton. ISBN: 0-8493-0464-4
13. van der Pauw LJ (1958) A method of measuring specific resistivity and hall effect of discs of arbitrary shape. Philips Res Rep 13:1–9
14. Sze SM (1981) Physics of semiconductor devices, 2nd edn. Wiley, New York (3rd edition ISBN: 978-0-471-14323-9, 2006)
15. Black JR (1969) Electromigration: a brief survey and some recent results. IEEE Trans Electron Devices ED-16:338–347
16. Ogawa ET, Lee K-D, Blaschke VA, Ho PS (2002) Electromigration reliability issues in dual-damascene Cu interconnections. IEEE Trans Reliab 51:403–419
17. van der Ziel A (1986) Noise in solid-state devices and circuits. Wiley-lnterscience, New York. ISBN: 0-471-832340
18. Feynman RP, Leighton RB, Sands M (1977) The Feynman lectures on physics, 6th edn, vol 1,2, and 3. Addison Wesley Publishing Company, California. ISBN: 0-201-02010-6-H

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4,

© Springer Science+Business Media, LLC 2013

19. Rosa EB (1908) The self and mutual inductances of linear conductors. Bull Bur of Stand 4:301–344
20. ITRS The national technology roadmap for semiconductors, technology needs. 1994–2011. Updates: http://www.itrs.net.
21. Seto JYW (1975) The electrical properties of polycrystalline silicon films. J Appl Phys 46:5247–5254
22. de Graaff HC, Klaassen FM (1990) Compact transistor modeling for circuit design. Springer, Wien . ISBN: 3-211-82136-8
23. Ghandhi S (1957) Darlington's compound connection for transistors. IRE Trans Circ Theor 4:291–292 (see also US patent 2-663-806)
24. McCreary JL (1981) Matching properties, and voltage and temperature dependence of MOS capacitors. IEEE J Solid-State Circ 16:608–616
25. Shyu J-B, Temes GC, Yao K (1982) Random errors in MOS capacitors. IEEE J Solid-State Circ 17:1070–1076
26. Aparicio R (2002) Capacity limits and matching properties of integrated capacitors. IEEE J Solid-State Circ 27:384–393
27. Bely M et al (2007) Capacitive integrated circuit structure. US patent 7-170-178
28. Wei C, Barrington RF, Mautz JR, Sarkar TK (1984) Multiconductor transmission lines in multilayered dielectric media. IEEE Trans Microw Theor Technol 32:439–450
29. Wallinga H, Bult K (1989) Design and analysis of CMOS analog signal processing circuits by means of a graphcal MOST model. IEEE J Solid-State Circ 24:672–680
30. Vertregt M (2006) The analog challenge of nanometer CMOS. In: Technical digest international electron devices meeting, IEDM Digest of Technical Papers, pp 1–8
31. Pelgrom MJM, Duinmaijer ACJ, Welbers APG (1989) Matching properties of MOS transistors. IEEE J Solid-State Circ 24:1433–1440
32. Woerlee PH, Knitel MJ, van Langevelde R, Klaassen DBM, Tiemeijer LF, Scholten AJ, Zegers-van Duijnhoven ATA (2001) RF-CMOS performance trends. IEEE Trans Electron Devices 48:1776–1782
33. Scholten AJ, Tiemeijer LF, De Vreede PWH, Klaassen DBM (1999) A large signal non-quasi-static MOS model for RF circuit simulation. In: Technical digest international electron devices meeting, pp 163–166
34. Meindl JD (1995) Low power microelectronics: retrospect and prospect. Proc of the IEEE 83:619–635
35. Klaassen FM, Hes W (1986) On the temperature coefficient of the MOSFET threshold voltage. Solid-State Electron 29:787–789
36. Scholten AJ, Tiemeijer LF, van Langevelde R, Havens RJ, Zegers-van Duijnhoven ATA, Venezia VC (2003) Noise modeling for RF CMOS circuit simulation. IEEE Trans Electron Devices 50:618–632
37. Brews JR (2006) MOSFET hand analysis using BSIM. IEEE Circ Devices Mag 21:28–36
38. Gildenblat G, Li X, Wu W, Wang H, Jha A, van Langevelde R, Smit GDJ, Scholten AJ, Klaassen DBM (2006) PSP: an advanced surface-potential-based MOSFET model for circuit simulation. IEEE Trans Electron Devices 53:1979–1993
39. Enz CC, Krummenacher F, Vittoz EA (1995) An analytical MOS transistor model valid in all regions of operations and dedicated to low- voltage and low-current applications. Analog Integr Circ Signal Process J 8:83–114
40. Sakurai T, Newton AR (1990) Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas. IEEE J Solid-State Circ 25:584–594
41. Lee MSL, Tenbroek BM, Redman-White W, Benson J, Uren MJ (2001) A physically based compact model of partially depleted MOSFETs for analog circuit stimulation. IEEE J Solid-State Circ 36:110–121
42. Middlebrook RD (2006) The general feedback theorem: a final solution for feedback systems. IEEE Microw Mag 7:50–63
43. Unbehauen R (1972) Synthese elektrischer Netzwerke. Oldenbourg Verlag, German

44. Sallen RP, Key EL (1955) A practical method of designing RC active filters. IRE Trans Circ Theor 2 CT-2:74–85
45. Nauta B (1992) Analog CMOS filters for very high frequencies. Kluwer Academic Publishers, Dordrecht. ISBN 0792392728
46. Berndt DF, Dutta Roy SC (1969) Inductor simulation with a single unity gain amplifier. IEEE J Solid State Circ 4:161–162
47. Martin K, Sedra A (1981) Effects of the op amp finite gain and bandwidth on the performance of switched-capacitor filters. IEEE Trans Circ Syst CAS-28:822–829
48. Huang Q, Sansen W (1987) Design techniques for improved capacitor area efficiency in switched-capacitor biquads. IEEE Trans Circ Syst CAS-34:1590–1599
49. Allen PE, Sanchez-Sinencio E (1984) Switched capacitor circuits. Van Nostrand Reinhold, New York. ISBN: 0-4422-0873-1
50. Allstot D, Black W (1983) Technological design consideration for monolithic MOS switched-capacitor filtering systems. Proc IEEE 71:967–986
51. Gregorian R, Temes GC (1986) Analog MOS integrated circuits for signal processing. Wiley, New York. ISBN: 0-471-09797-7
52. Johns D, Martin KW (1997) Analog integrated circuit design. Wiley, New York. ISBN: 0-471-14448-7
53. Schreier R, Silva J, Steensgaard J, Temes GC (2005) Design-oriented estimation of thermal noise in switched-capacitor circuits. IEEE Trans Circ Syst I 15:2358–2368
54. Gray PR, Meyer RG (1993) Analysis and design of analog integrated circuits, 3rd edn. Wiley, New York (4th edn, Wiley, New York, ISBN: 0-471-32168-0, 2001)
55. Allen P, Holberg D (1987) CMOS analog circuit design. Holt, Rinehart and Winston Inc, New York
56. Rijns JJF (1996) CMOS low-distortion high-frequency variable-gain amplifier. IEEE J Solid-State Circ 31:1029–1034
57. Krummenacher F, Joehl N (2009) A differential-ramp based 65 dB-linear VGA technique in 65 nm CMOS. IEEE J Solid-State Circ 44:2503–2514
58. Gilbert B (1968) A precise four-quadrant multiplier with subnanosecond response. IEEE J Solid-State Circ 3:365–373
59. Elwan H, Tekin A, Pedrotti K (1988) A 4-MHz CMOS continuous-time filter with on-chip automatic tuning. IEEE J. Solid-State Circ 23:750–758
60. Bult K, Geelen GJGM (1990) A fast-settling CMOS op amp for SC circuits with 90-dB DC gain. IEEE J Solid-State Circ 25:1379–1384
61. Sackinger E, Guggenbuhl W (1990) A high-swing, high-impedance MOS cascode circuit. IEEE J Solid-State Circ 25:289–298
62. Kamath BYT, Meyer RG, Gray PR (1974) Relationship between frequency response and settling time of operational amplifiers. IEEE J Solid-State Circ 9:347–352
63. Solomon J (1974) The monolitic op amp: a tutorial study. IEEE J Solid-State Circ 9:314–332
64. Tsividis Y, Gray P (1976) An integrated NMOS operational amplifier with internal compensation. IEEE J Solid-State Circ 11:748–754
65. Tsividis Y (1978) Design consideration in single-channel MOS analog lntegraied circuits–a tutorial. IEEE J Solid-State Circ 13:383–391
66. Gray PR, Meyer R (1982) MOS operational amplifier design a tutorial overview. IEEE J Solid-state Circ 17:969–982
67. Redman-White W (1997) 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 J Solid-State Circ 32:701–712
68. Ahuja B (1983) An improved frequency compensation technique for CMOS operational amplifiers. IEEE J Solid-state Circ 18:629–633
69. Cherry EM, Hooper DE (1963) The design of wide-band transistor feedback amplifiers. Proc IEEE 110(2):375–389
70. Hermans C, Steyaert MSJ (2006) A high-speed 850-nm optical receiver front-end in 0.18μm CMOS. IEEE J Solid-State Circ 41:1606–1614

71. Leeson DB (1966) A simple model of feedback oscillator noise spectrum. Proc IEEE 54:329–330
72. Demir A (2006) Computing timing jitter from phase noise spectra for oscillators and phase-locked loops with white and $1/f$ noise. IEEE Trans Circ Syst I 53:1869–1884
73. Hajimiri A, Lee T (1998) A general theory of phase noise in electrical oscillators. IEEE J Solid-State Circ 33:179–194
74. Razavi B (1996) A study of phase noise in CMOS oscillators. IEEE J Solid-State Circ 31:331–343
75. van der Tang JD (2002) High frequency oscillator design for integrated transceivers. Ph.D. thesis, Technical University Eindhoven
76. Huang Q (2000) Phase noise to carrier ratio in LC oscillators. IEEE Trans Circ Syst I: Fundam Theor Appl 47:965–980
77. Vittoz E, Degrauwe M, Bitz S (1988) High-performance crystal oscillator circuits: theory and application. IEEE J Solid-state Circ 23:774–783
78. Thommen W (1999) An improved low-power crystal oscillator. In: 25th European solid-state circuits conference, pp 146–149
79. Santos JT, Meyer RG A one-pin crystal oscillator for VLSI circuits. IEEE J Solid-State Circ 19:228–236 (1984)
80. Geraedts P, van Tuijl E, Klumperink E, Wienk G, Nauta B (2008) A 90 μW 12 MHz relaxation oscillator with a -162dB FOM. In: International solid-state circuits conference, digest of technical papers, pp 348–349
81. Sebastiano F, Breems L, Makinwa K, Drago S, Leenaerts D, Nauta B (2009) A low-voltage mobility-based frequency reference for crystal-less ULP radios. IEEE J Solid-State Circ 44:2002–2009
82. Gao X, Klumperink EAM, Bohsali M, Nauta B (2009) A low noise sub-sampling PLL in which divider noise is eliminated and PD/CP noise is not multiplied by N^2. IEEE J Solid-State Circ 44:3253–3263
83. Rabiner LR, Gold B (1975) Theory and application of digital signal processing. Prentice-Hall, Englewood Cliffs. ISBN: 0-139-141014
84. van den Enden AWM, Verhoeckx NAM (1989) Discrete time signal processing, an introduction. Prentice Hall, New Jersey. ISBN: 0-132-167557
85. Nyquist H (1928) Certain topics in telegraph transmission theory. Trans AIEE 47:617–644 (Reprinted in Proc IEEE 90:280–305, 2002)
86. Shannon CE (1948) A mathematical theory of communication. Bell Syst Technol J 27: 379–423 and 623–656
87. Shannon CE (1949) Communication in the presence of noise. Proc IRE 37:10–21 (Reprinted in Proc IEEE 86:447–457, 1998)
88. Unser M (2000) Sampling-50 years after Shannon. Proc IEEE 88:569–587
89. Candes E, Romberg J, Tao T (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans Inform Theor 52:489–509
90. Shinagawa M, Akazawa Y, Wakimoto T (1990) Jitter analysis of high-speed sampling systems. IEEE J Solid-State Circ 25:220–224
91. McClellan JH, Parks TW, Rabiner LR (1973) A computer program for designing optimum FIR linear phase digital filters. IEEE Trans Audio Electroacoustics 21:506–526
92. Séquin CH, Tompsett MF (1975) Charge transfer devices, supplement 8 to advances in electronics and electron physics. Academic, New York
93. Crols J, Steyaert M (1994) Switched-opamp: an approach to realize full CMOS switched-capacitor circuits at very low power supply voltages. IEEE J Solid-State Circ 29:936–942
94. Baschirotto A (1997) Castello R A 1-V 1.8-MHz CMOS switched-opamp SC filter with rail-to-rail output swing. IEEE J Solid-State Circ 32:1979–1986
95. Keramat A, Tao Z (2000) 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, pp 48–51

96. Dickson J (1976) On-chip high-voltage generation MNOS integrated circuits using an improved voltage multiplier technique. IEEE J Solid-State Circ 11:374–378
97. Knepper RW (1978) Dynamic depletion mode: An E/D mosfet circuit method for improved performance. IEEE J Solid-State Circ 13:542–548
98. Abo AM, Gray PR (1999) A 1.5-V, 10-bit, 14.3-MS/s CMOS pipe-line analog-to-digital converter. IEEE J Solid-State Circ 34:599–606
99. Limotyrakis S, Kulchycki SD, Su DK, Wooley BA (2005) A 150-MS/s 8-b 71-mW CMOS time-interleaved ADC. IEEE J Solid-State Circ 40:1057–1067
100. Song BA, Tompsett MF, Lakshmikumar KR (1988) A 12-bit 1 -MS / s capacitor error-averaging pipelined A/D converter. IEEE J Solid-State Circ 23:1324–1333
101. Yang W, Kelly D, Mehr I, Sayuk MT, Singer L (2001) A 3-V 340-mW 14-b 75-MS/s CMOS ADC with 85-dB SFDR at Nyquist input. IEEE J Solid-State Circ 36:1931–1936
102. Gregoire BR, Moon U (2008) An Over-60 dB true rail-to-rail performance using correlated level shifting and an opamp with only 30 dB loop gain. IEEE J Solid-State Circ 43:2620–2630
103. Wakimoto T, Akazawa Y (1993) Circuits to reduce distortion in the diode-bridge track-and-hold. IEEE J Solid-State Circ 28:384–387
104. Vorenkamp P, Verdaasdonk JPM (1992) Fully bipolar, 120-MS/s 10-b track-and-hold circuit. IEEE J Solid-State Circ 27:988–992
105. Harley Reeves A (1942) Electric signaling system. US Patent 2-272-070, 3 Feb 1942. Also French Patent 852-183 issued 1938, and British Patent 538-860 issued 1939
106. IEEE Std 1057-1994 (1994) IEEE standard for digitizing waveform recorders
107. IEEE 1241–2000 Standard for terminology and test methods for analog-to-digital converters. IEEE Std1241, 2000, ISBN: 0-7381-2724-8, revision 2007
108. Tilden SJ, Linnenbrink TE, Green PJ (1999) Overview of IEEE-STD-1241 standard for terminology and test methods for analog-to-digital converters. In: Instrumentation and measurement technology conference, pp 1498–1503
109. Bennett WR (1948) Spectra of quantized signals. Bell Syst Technol J 27:446–472
110. Blachman N (1985) The intermodulation and distortion due to quantization of sinusoids. IEEE Trans Acoustics Speech Signal Process ASSP 33:1417–1426
111. Oude Alink MS, Kokkeler ABJ, Klumperink EAM, Rovers KC, Smit G, Nauta B (2009) Spurious-free dynamic range of a uniform quantizer. IEEE Trans Circ Syst II: Express Briefs 56:434–438
112. Lloyd S (1982) Least squares quantization in PCM. IEEE Trans Inform Theor 28:129–137 (transcript from 1957 paper)
113. Max J (1960) Quantizing for minimum distortion. IRE Trans Inform Theor 6:7–12
114. Wannamaker RA, Lipshitz SP, Vanderkooy J, Wright JN (2000) A theory of nonsubtractive dither. IEEE Trans on Signal Process 48:499–516
115. Hilbiber D (1964) A new semiconductor voltage standard. In: IEEE International Solid-State Circuits Conference, Digest of Technical Papers, pp 32–33
116. Widlar RJ (1971) New developments in IC voltage regulators. IEEE J Solid-State Circ 6:2–7
117. Kuijk KE (1973) A precision reference voltage source. IEEE J Solid-State Circ 8:222–226
118. Brokaw AP (1974) A simple three-terminal IC bandgap reference. IEEE J Solid-State Circ 9:388–393
119. Pertijs MAP, Huijsing JH (2006) Precision temperature sensors in CMOS technology. Springer, New York, ISBN: 140205257X
120. Song BS, Gray PR (1983) A precision curvature-compensated CMOS bandgap reference. IEEE J Solid-State Circ 18:634–643
121. Banba H, Shiga H, Umezawa A, Miyaba T, Tanzawa T, Atsumi S, Sakui K (1999) A CMOS Band-gap reference circuit with sub 1-V operation. IEEE J Solid-State Circ 34:670–674
122. Petrescu V, Pelgrom MJM, Veendrick HJM, Pavithran P, Wieling J (2006) Monitors for a signal integrity measurement system. In: 32nd European solid-state circuits conference, pp 122–125
123. Petrescu V, Pelgrom MJM, Veendrick HJM, Pavithran P, Wieling J (2006) A signal-integrity self-test concept for debugging nanometer CMOS ICs. In: IEEE international solid-state circuits conference, digest of technical papers, pp 544–545 (2006)

124. Fayomi CJB, Wirth GI, Achigui HF, Matsuzawa A (2010) Sub 1 V CMOS bandgap reference design techniques: a survey. Analog Integr Circ Signal Process 62:141–157
125. Annema A-J (1999) Low-power bandgap references featuring DTMOSTs. IEEE J Solid-State Circ 34:949–955
126. Sansen WM, Op't Eynde F, Steyaert M (1988) A CMOS temperature compensated current reference. IEEE J Solid-State Circ 23:821–823
127. Blauschild RA, Tucci PA, Muller RS, Meyer RG (1978) A new NMOS temperature-stable voltage reference. IEEE J Solid-State Circ 13:767–774
128. Song H-J, Kim C-K (1993) A temperature-stabilized SOI voltage reference based on threshold voltage difference between enhancement and depletion NMOSFET's. IEEE J Solid-State Circ 28:671–677
129. van de Plassche RJ (1976) Dynamic element matching for high-accuracy monolithic D/A converters. IEEE J Solid-State Circ 21:795–800
130. Schoeff JA (1979) An inherently monotonic 12 bit DAC. IEEE J Solid-State Circ 24:904–911
131. Naylor JR (1983) A complete high-speed voltage output 16-bit monolithic DAC. IEEE J Solid-State Circ 28:729–735
132. Schouwenaars HJ, Dijkmans EC, Kup BMJ, van Tuijl EJM (1986) A monolithic dual 16-bit D/A converter. IEEE J Solid-State Circ 21:424–429
133. Groeneveld DWJ, Schouwenaars HJ, Termeer HAH, Bastiaansen CAA (1989) A self-calibration technique for monolithic high-resolution D/A converters. IEEE J Solid-State Circ 24:1517–1522
134. Schouwenaars HJ, Groeneveld DWJ, Termeer HAH (1988) A low-power stereo 16-bit CMOS D/A converter for digital audio. IEEE J Solid-State Circ 23:1290–1297
135. Lin C-H, Bult K (1998) A 10-b 250-M sample/s CMOS DAC in 1 mm^2. In: IEEE International solid-state circuits conference, digest of technical papers, pp 214–215
136. Van Den Bosch A, Borremans M, Steyaert M, Sansen W (2001) A 12 b 500 MS/s current-steering CMOS D/A converter. In: IEEE international solid-state circuits conference, digest of technical papers, pp 366–367
137. Van Den Bosch A, Borremans M, Steyaert M, Sansen W (2001) A 10-bit 1-GSample/s Nyquist current-steering CMOS D/A converter. IEEE J Solid-State Circ 36:315–324
138. Bastiaansen CAA, Groeneveld DWJ, Schouwenaars HJ, Termeer HAH (1991) A 10-b 40-MHz 0.8-μm CMOS current-output D/A converter. IEEE J Solid-State Circ 26:917–921
139. Doris K, Briaire J, Leenaerts D, Vertregt M, van Roermund A (2005) A 12b 500MS/s DAC with $>$ 70 dB SFDR up to 120MHz in 0.18μm CMOS. In: IEEE international solid-state circuits conference, digest of technical papers, pp 116–588
140. Van der Plas GAM, Vandenbussche J, Sansen W, Steyaert MSJ, Gielen GGE (1999) A 14-bit intrinsic accuracy Q^2 random walk CMOS DAC. IEEE J Solid-State Circ 34:1708–1718
141. Jewett B, Liu J, Poulton K (2005) A 1.2GS/s 15b DAC for precision signal generation. In: IEEE international solid-state circuits conference, digest of technical papers, pp 110–111
142. Park S, Kim G, Park S-C, Kim W (2002) A digital-to-analog converter based on differential-quad switching. IEEE J Solid-State Circ 37:1335–1338
143. Engel G, Kuo S, Rose S (2012) A 14b 3/6GHz current-steering RF DAC in 0.18m CMOS with 66dB ACLR at 2.9GHz. In: International solid-state circuits conference, digest of technical papers, pp 458–449
144. Schafferer B, Adams R (2004) A 3V CMOS 400mW 14b 1.4GS/s DAC for multi-carrier applications. In: IEEE international solid-state circuits conference, digest of technical papers, pp 360–361
145. Schofield W, Mercer D, Onge LS (2003) A 16b 400MS/s DAC with <-80dBc IMD to 300MHz and <-160dBm/Hz noise power spectral density. In: IEEE international solid-state circuits conference, digest of technical papers, pp 126–127
146. Su DK, Wooley BA (1993) A CMOS oversampling D/A converter with a current-mode semidigital reconstruction filter. IEEE J Solid-State Circ 28:1224–1233
147. Barkin DB, Lin ACY, Su DK, Wooley BA (2004) A CMOS oversampling bandpass cascaded D/A converter with digital FIR and current-mode semi-digital filtering. IEEE J Solid-State Circ 39:585–593

148. Suarez RE, Gray PR, Hodges DA (1975) All-MOS charge-redistribution analog-to-digital conversion techniques, II. IEEE J Solid-State Circ 10:379–385
149. Song B-S, Lee S-H, Tompsett MF (1990) A 10-b 15-MHz CMOS recycling two-step A/D converter. IEEE J Solid-State Circ 25:1328–1338
150. Philips K, van den Homberg J, Dijkmans C (1999) 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, pp 154–155
151. Fan Q, Huijsing JH, Makinwa KAA (2012) A 21 nV/$\sqrt{\text{Hz}}$ Hz chopper-stabilized multi-path current-feedback instrumentation amplifier with 2 μV offset. IEEE J Solid-State Circ 47: 464-475
152. Blanken PG, Menten SEJ (2002) A 10 μV-offset 8 kHz bandwidth 4th-order chopped $\Sigma\Delta$ A/D converter for battery management. In: IEEE international solid-state circuits conference, digest of technical papers, pp 388–389
153. Carley L (1989) A noise-shaping coder topology for 15+ bit converters. IEEE J Solid-State Circ 24:267–273
154. Nys OJAP, Henderson RK (1996) An analysis of dynamic element matching techniques in sigma-delta modulation. In: IEEE international symposium on circuits and systems, pp 231–234
155. Henderson RK, Nys OJAP (1996) Dynamic element matching techniques with arbitrary noise shaping function. In: IEEE international symposium on circuits and systems, pp. 293–296
156. Story MJ (1992) 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
157. Maloberti A (1991) 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)
158. Miller MR, Petrie CS (2003) A multibit sigma-delta ADC for multimode receivers. IEEE J Solid-State Circ 38:475–482
159. Risbo L, Hezar R, Kelleci B, Kiper H, Fares M (2011) A 108dB-DR 120dB-THD and 0.5Vrms output audio DAC with inter-symbol-interference-shaping algorithm in 45nm CMOS. In: IEEE international solid-state circuits conference, digest of technical papers, pp 484–485
160. Dingwall AGF, Zazzu V (1985) An 8-MHz CMOS subranging 8-bit A/D converter. IEEE J Solid-State Circ 20:1138–1143
161. Abrial A, Bouvier J, Fournier J, Senn P, Viellard M (1988) A 27-MHz digital-to-analog video processor. IEEE J Solid-State Circ 23:1358–1369
162. Pelgrom MJM (1990) A 10b 50MHz CMOS D/A converter with 75Ω buffer. IEEE J Solid-State Circ 25:1347–1352
163. Miki T, Nakamura Y, Nakaya M, Asai S, Akasaka Y, Horiba Y (1986) An 80-MHz 8-bit CMOS D/A converter. IEEE J Solid-State Circ 21:983–988
164. Pelgrom MJM, Roorda M (1988) An algorithmic 15 bit CMOS digital-to-analog converter. IEEE J Solid-State Circ 23:1402–1405
165. Matsumoto H, Watanabe K (1986) Switched-capacitor algorithmic digital-to-analog converters. IEEE Trans Circ Syst 33:721–724
166. Fiedler HL, Hoefflinger B, Demmer W, Draheim P (1981) A 5-bit building block for 20 MHz A/D converters. IEEE J Solid-State Circ 26:151–155
167. Wu J-T, Wooley BA (1988) A 100-MHz pipelined CMOS comparator. IEEE J Solid-State Circ 23:1379–1385
168. Nauta B, Venes AGW (1995) A 70 Ms/s 110 mW 8-b CMOS folding and interpolating A/D converter. IEEE J Solid-State Circ 30:1302–1308
169. Yin G, Op't Eynde F, Sansen W (1992) A high-speed CMOS comparator with 8-b resolution. IEEE J Solid-State Circ 37:208–211
170. Venes AGW, van de Plassche RJ (1996) An 80-MHz, 80-mW, 8-b CMOS folding A/D converter with distributed track-and-hold preprocessing. IEEE J Solid-State Circ 31: 1846–1853

171. Ellersick W, Chih-Kong KY, Horowitz M, Dally W (1999) 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, pp 49–52
172. Montanaro J et al (1996) A 160 MHz, 32b, 0.5W CMOS RISC microprocessor. IEEE J Solid-State Circ 31:1703–1714
173. Verbruggen B, Craninckx J, Kuijk M, Wambacq P, Van der Plas G (2008) A 2.2mW 5b 1.75GS/s folding flash ADC in 90nm digital CMOS. In: IEEE international solid-state circuits conference, digest of technical papers, pp 252–611
174. Schinkel D, Mensink E, Klumperink E, van Tuijl E, Nauta B (2007) A double-tail latch-type voltage sense amplifier with 18ps setup+hold time. In: IEEE international solid-state circuits conference, digest of technical papers, pp 314–315
175. Fukushima N, Yamada T, Kumazawa N, Hasegawa Y, Soneda M (1989) A CMOS 40MHz 8b 105mW two-step ADC. In: International solid-state circuits conference, digest of technical papers, pp 14–15
176. Atherton JH, Simmonds HT (1992) An offset reduction technique for use with CMOS integrated comparators and amplifiers. IEEE J Solid-State Circ 27:1168–1175
177. Wong K-LJ, Yang C-KK (2004) Offset compensation in comparators with minimum input-referred supply noise. IEEE J Solid-State Circ 37:837–840
178. Kusumoto K, Matsuzawa A, Murata K (1993) A 10-b 20-MHz 30-mW pipelined interpolating CMOS ADC. IEEE J Solid-State Circ 28:1200–1206
179. Haas M, Draxelmayr D, Kuttner F, Zojer B (1990) A monolithic triple 8-bit CMOS video coder. IEEE Trans Consum Electron 36:722–729
180. Schvan P, Pollex D, Wang S-C, Falt C, Ben-Hamida N (2006) A 22GS/s 5b ADC in 0.13m SiGe BiCMOS. In: International solid-state circuits conference, digest of technical papers, pp 572–573
181. Reyhani H, Quinlan P (1994) A 5 V, 6-b, 80 Ms/s BiCMOS flash ADC. IEEE J Solid-State Circ 29:873–878
182. Vorenkamp P, Verdaasdonk JPM (1992) A 10b 50MHz pipelined ADC. In: International solid-state circuits conference, digest of technical papers, pp 32–33
183. Kattmann K, Barrow J (1991) A technique for reducing differential nonlinearity errors in flash A/D converters. In: International solid-state circuits conference, digest of technical papers, pp 170–171
184. Scholtens PCS, Vertregt M (2002) A 6-b 1.6-Gsample/s flash ADC in 0.18μm CMOS using averaging termination. IEEE J Solid-State Circ 37:1599–1609
185. Bult K, Buchwald A (1997) An embedded 240-mW 10-b 50-MS/s CMOS ADC in 1-mm^2. IEEE J Solid-State Circ 32:1887–1895
186. Uyttenhove K, Vandenbussche J, Lauwers E, Gielen GGE, Steyaert MSJ (2003) Design techniques and implementation of an 8-bit 200-MS/s interpolating/averaging CMOS A/D converter. IEEE J Solid-State Circ 38:483–494
187. Van De Grift REJ, Rutten IWJM, van der Veen M (1987) An 8-bit video ADC incorporating folding and interpolation techniques. IEEE J Solid-State Circ 22:944–953
188. Vorenkamp P, Roovers R (1997) A 12-b, 60-MS/s cascaded folding and interpolating ADC. IEEE J Solid-State Circ 32:1876–1886
189. Van De Plassche RJ, van der Grift REJ (1979) A high-speed 7 bit A/D converter. IEEE J Solid-State Circ 14:938–943
190. Hoogzaad G, Roovers R (1999) A 65-mW, 10-bit, 40-MS/s BiCMOS Nyquist ADC in 0.8 mm^2. IEEE J Solid-State Circ 34:1796–1802
191. Choe MJ, Song B-S, Bacrania K (2000) A 13b 40MS/s CMOS pipelined folding ADC with background offset trimming. In: International solid-state circuits conference, digest of technical papers, pp 36–37
192. Lewis SH, Gray PR (1987) A pipelined 5-MS/s 9-bit analog-to-digital converter. IEEE J Solid-State Circ 22:954–961
193. van der Ploeg H, Vertregt M, Lammers M (2006) A 15-bit 30-MS/s 145-mW three-step ADC for imaging applications. IEEE J Solid-State Circ 41:1572–1577

194. Shimizu Y, Murayama S, Kudoh K, Yatsuda H, Ogawa A (2006) A 30mW 12b 40MS/s subranging ADC with a high-gain offset-canceling positive-feedback amplifier in 90nm digital CMOS. In: International solid-state circuits conference, digest of technical papers, pp 802–803
195. Moreland C, Murden F, Elliott M, Young J, Hensley M, Stop R (2000) A 14b 100MS/s subranging ADC. IEEE J Solid-State Circ 35:1791–1798
196. McCharles R, Hodges D (1978) Charge circuits for analog LSI. IEEE Trans Circ Syst 25:490–497
197. Cho TB, Gray PR (1995) A 10 b, 20 Msample/s, 35 mW pipeline A/D converter. IEEE J Solid-State Circ 30:166–172
198. Chai Y, Wu J-T (2012) A 5.37mW 10b 200MS/s dual-path pipelined ADC. In: International solid-state circuits conference, digest of technical papers, pp 462–463
199. Li PW, Chin MJ, Gray PR, Castello R (1984) A ratio-independent algorithmic analog-to-digital conversion technique. IEEE J Solid-State Circ 19:828–836
200. Karanicolas AN, Lee H-S, Barcrania KL (1993) A 15-b 1-MS/s digitally self-calibrated pipeline ADC. IEEE J Solid-State Circ 28:1207–1215
201. Nagaraj K, Fetterman HS, Anidjar J, Lewis SH, Renninger RG (1997) A 250-mW, 8-b, 52-MSs/s parallel-pipelined A/D converter with reduced number of amplifiers. IEEE J Solid-State Circ 32:312–320
202. Chiu Y, Gray PR, Nikolic B (2004) A 14-b 12-MS/s CMOS pipeline ADC with over 100-dB SFDR. IEEE J Solid-State Circ 39:2139–2151
203. Wang X, Hurst PJ, Lewis SH (2004) A 12-bit 20-MS/s pipelined analog-to-digital converter with nested digital background calibration. IEEE J Solid-State Circ 39:1799–1808
204. Murmann B, Boser BE (2003) A 12-bit 75-MS/s pipelined ADC using open-loop residue amplification. IEEE J Solid-State Circ 38:2040–2050
205. Iroaga E, Murmann B (2007) A 12-Bit 75-MS/s pipelined ADC using incomplete settling. IEEE J Solid-State Circ 42:748–756
206. Geelen G, Paulus E, Simanjuntak D, Pastoor H, Verlinden R (2006) A 90nm CMOS 1.2V 10b power and speed programmable pipelined ADC with 0.5pJ/conversion-step. In: International solid-state circuits conference, digest of technical papers, 214–215
207. Bardsley S, Dillon C, Kummaraguntla R, Lane C, Ali AMA, Rigsbee B, Combs D (2006) A 100-dB SFDR 80-MSPS 14-Bit 0.35-μm BiCMOS Pipeline ADC. IEEE J Solid-State Circ 41:2144–2153
208. Lee BG, Min BM, Manganaro G, Valvano JW (2008) A 14b 100 MS/s pipelined ADC with a merged active S/H and first MDAC. In: International solid-state circuits conference, digest of technical papers, pp 248–249
209. van de Vel H, Buter B, Ploeg Hvd, Vertregt M, Geelen G, Paulus E (2009) A 1.2 V 250-mW 14-b 100 MS/s digitally calibrated pipeline ADC in 90-nm CMOS. IEEE J Solid-State Circ 44:1047–1056
210. Lee CC, Flynn MP (2011) A SAR-assisted two-stage pipeline ADC. IEEE J Solid-State Circ 46:859–869
211. Verbruggen B, Iriguchi M, Craninckx J (2012) A 1.7mW 11b 250MS/s 2 interleaved fully dynamic pipelined SAR ADC in 40nm digital CMOS. In: International solid-state circuits conference, digest of technical papers, pp 466–467
212. Min BM, Kim P, Bowman FW, Boisvert DM, Aude AJ (2003) A 69-mW 10-bit 80-MS/s pipelined CMOS ADC. IEEE J Solid-State Circ 38:2031–2039
213. Mehr I, Singer L (2000) A 55-mW 10-bit 40-MS/s Nyquist-rate CMOS ADC. IEEE J Solid-State Circ 35:318–323
214. Sepke T, Fiorenza JK, Sodini CG, Holloway P, Lee H-S (2006) Comparator-based switched-capacitor circuits for scaled CMOS technologies. In: International solid-state circuits conference, digest of technical papers, pp 812–821
215. Brooks L, Lee H-S (2009) A 12b 50MS/s fully differential zero-crossing-based ADC without CMFB. In: International solid-state circuits conference, digest of technical papers, pp 166–167

216. Wang H, Wang X, Hurst PJ, Lewis SH (2009) Nested digital background calibration of a 12-bit pipelined ADC without an input SHA. IEEE J Solid-State Circ 44:2780–2789
217. McCreary JL, Gray PR (1975) All-MOS charge redistribution analog-to-digital conversion techniques I. IEEE J Solid-State Circ 10:371–379
218. Craninckx J, Van der Plas G (2007) A 65fJ/conversion-step 0-to-50MS/s 0-to-0.7mW 9b cHARGE-sHaring SAR ADC in 90nm digital CMOS. In: International solid-state circuits conference, digest of technical papers, pp 246–247
219. Agnes A, Bonizzoni E, Malcovati P, Maloberti F (2008) A 9.4-ENOB 1V 3.8μW 100kS/s SAR ADC with time-domain comparator. In: International solid-state circuits conference, digest of technical papers, pp 246–247
220. van Elzakker M, van Tuijl E, Geraedts P, Schinkel D, Klumperink E, Nauta B (2008) A 1.9μW 4.4fJ/conversion-step 10b 1MS/s charge-redistribution ADC. In: International solid-state circuits conference, digest of technical papers, pp 244–245
221. Harpe P, Zhang Y, Dolmans G, Philips K, de Groot H (2012) A 7-to-10b 0-to-4MS/s flexible SAR ADC with 6.5-to-16fJ/conversion-step. In: International solid-state circuits conference, digest of technical papers, pp 472–473
222. Kuttner F (2002) A 1.2V 10 b 20 MS/s non-binary successive approximation ADC in 0.13μm CMOS. In: International solid-state circuits conference, digest of technical papers, pp 176–177
223. Hesener M, Ficher T, Hanneberg A, Herbison D, Kuttner F, Wenskel H (2007) A 14b 4OMS/s redundant SAR ADC with 480MHz in 0.13pm CMOS. In: International solid-state circuits conference, digest of technical papers, pp 248–249
224. Shih C, Gray PR (1986) Reference refreshing cyclic analog-to-digital and digital-to-analog converters. IEEE J Solid-State Circ 21:544–554
225. Ginetti B, Jespers P, Vandemeulebroecke A (1992) A CMOS 13-b cyclic. A/D converter. IEEE J Solid-State Circ 27:957–964
226. Mase M, Kawahito S, Sasaki M, Wakamori Y, Furuta M (2005) A wide dynamic range CMOS image sensor with multiple exposure-time signal outputs and 12-bit column-parallel cyclic A/D converters. IEEE J Solid-State Circ 40:2787–2795
227. Snoeij MF, Theuwissen AJP, Makinwa KAA, Huijsing JH (2007) Multiple-ramp column-parallel ADC architectures for CMOS image sensors. IEEE J Solid-State Circ 42:2986–2977
228. Naraghi S, Courcy M, Flynn MP (2009) A 9b 14 μW 0.06mm^2 PPM ADC in 90nm digital CMOS. In: IEEE international solid-state circuits conference digest of technical papers, pp 168–169
229. Howard BK (1955) Binary quantizer. US patent 2-715-678
230. van der Ploeg H, Hoogzaad G, Termeer HAH, Vertregt M, Roovers RLJ (2001) A 2.5V, 12b, 54MS/s, 0.25um CMOS ADC. In: International solid-state circuits conference, digest of technical papers, pp 132–133
231. Pelgrom MJM, Jochijms A, Heijns H (1987) A CCD delay line for video applications. IEEE Trans Consum Electron 33:603–609
232. Kurosawa N, Kobayashi H, Maruyama K, Sugawara H, Kobayashi K (2001) Explicit analysis of channel mismatch effects in time-interleaved ADC systems. IEEE Trans Circ Syst I: Fundam Theor Appl 48:261–271
233. Doris K, Janssen E, Nani C, Zanikopoulos A, Wiede Gvd (2011) A 480 mW 2.6 GS/s 10b time-interleaved ADC With 48.5 dB SNDR up to Nyquist in 65 nm CMOS. IEEE J Solid-State Circ 46:2821–2833
234. Hsu C-C, Huang F-C, Shih C-Y, Huang C-C, Lin Y-H, Lee C-C, Razavi B (2007) An 11b 800MS/s time-interleaved ADC with digital background calibration. In: International solid-state circuits conference, digest of technical papers, pp 164–165
235. Vertregt M, Dijkstra MB, Rens ACv, Pelgrom MJM (1993) A Versatile digital CMOS video delay line with embedded ADC, DAC and RAM. In: 19th European solid-state circuits conference, pp 226–229
236. Pelgrom MJM, Rens ACv, Vertregt M, Dijkstra MB (1994) A 25-Ms/s 8-bit CMOS A/D converter for embedded application. IEEE J Solid-State Circ 29:879–886

237. Murray B, Menting H (1992) A highly integrated D2MAC decoder. In: IEEE international conference on consumer electronics, digest of technical papers, pp 56–57
238. Mark JW, Todd TD (1981) A nonuniform sampling approach to data compression. IEEE Trans Commun 29:24–32
239. Allier E, Goulier J, Sicard G, Dezzani A, Andre E, Renaudin M (2005) A 120nm low power asynchronous ADC. In: International symposium on low-power electronics and design, pp 60–65
240. Trakimas M, Sonkusale SR (2011) An adaptive resolution asynchronous adc architecture for data compression in energy constrained sensing applications. IEEE Trans Circ Syst I 58: 921–934
241. Lin C-S, Liu B-D (2003) A new successive approximation architecture for low-power low-cost CMOS A/D converter. IEEE J Solid-State Circ 38:54–62
242. Chen S-WM, Brodersen RW (2006) A 6-bit 600-MS/s 5.3-mW asynchronous ADC in 0.13-μm CMOS. IEEE J Solid-State Circ 41:2669–2680
243. Pernillo J, Flynn MP (2011) A 1.5-GS/s flash ADC With 57.7-dB SFDR and 6.4-bit ENOB in 90 nm digital CMOS. IEEE Trans Circ Syst II: Express Briefs 58:837–841
244. Jansson J-P, Mantyniemi A, Kostamovaara J (2006) A CMOS time-to-digital converter with better than 10 ps single-shot precision. IEEE J Solid-State Circ 41:1286–1296
245. Chen P, Liu S-L, Wu J (2000) A CMOS pulse-shrinking delay element for time interval measurement. IEEE Trans Circ Syst 47:954–958
246. Rahkonen TE, Kostamovaara JT (1993) The use of stabilized CMOS delay lines for the digitization of short time intervals. IEEE J Solid-State Circ 28:887–894
247. van der Ploeg H (1997) The Nonius analog-to-digital converter. In: Internal Philips research report/ University Twente B.Sc. report, supervisor M. Pelgrom
248. Groza VZ (2001) High-resolution floating-point ADC. IEEE Trans Instrumentation and Measurement 50:1812–1829
249. de Jager F (1952) Delta modulation, a method of PCM transmission using the 1-unit code. Philips Res Rep 7:442–466
250. Cutler C (1960) Transmission systems employing quantization. US Patent 2-927-962
251. Widrow B (1956) A study of rough amplitude quantization by means of Nyquist sampling theory. IRE Trans Circ Theor CT-3:266–276
252. Inose H, Yasuda Y, Murakami J (1962) A telemetering system by code modulation- $\Delta\Sigma$ modulation. IRE Trans Space Electron Telemetry SET-8:204–209 (Proc IEEE 51:1524–1535, 1963)
253. Candy JC, Temes GC (eds) (1992) Oversampling delta-sigma data converters: theory, design and simulation. IEEE, New York
254. Norsworthy SR, Schreier R, Temes GC (eds) (1997) Delta-sigma data converters: theory, design, and simulation. IEEE Press, Piscataway. ISBN: 0-7803-1045-4
255. Schreier R, Temes GC (2004) Understanding delta-sigma data converters. Wiley, New York. ISBN: 0-471-46585-2
256. Naus PJA, Dijkmans EC (1991) Multi-bit oversampled $\Sigma\Delta$ A/D converters as front-end for CD players. IEEE J Solid-State Circ 26:905–909
257. Kup BMJ, Dijkmans EC, Naus PJA, Sneep J (1991) A bit-stream digital-to-analog converter with 18-b resolution. IEEE J Solid-State Circ 26:1757–1763
258. Naus PJA, Dijkmans EC, Stikvoort EF, Holland DJ, Bradinal W (1987) A CMOS stereo 16-bit D/A converter for digital audio. IEEE J Solid-State Circ 22:390–395
259. Adams R, Nguyen KQ (1998) A 113-dB SNR oversampling DAC with segmented noise-shaped scrambling. IEEE J Solid-State Circ 33:1871–1878
260. Naus PJA, Dijkmans EC (1988) Low signal level distortion in sigma-delta modulators. In: 84th convention of the audio engineering society, Paris
261. Vleugels K, Rabii S, Wooley BA (2001) A 2.5 v sigma-delta modulator for broadband communications applications. IEEE J Solid-State Circ 36:1887–1889
262. Hart A, Voinigescu SP (2009) A 1 GHz bandwidth low-pass sigma-delta ADC with 20-50 GHz adjustable sampling rate. IEEE J Solid-State Circ 44:1401–1414

263. Gambini S, Rabaey J (2007) A 100-kS/s 65-dB DR sigma-delta ADC with 0.65V supply voltage. In: 33th European solid state circuits conference, pp 202–205
264. Lee WL, Sodini CG (1987) A topology for higher order interpolative coders. In: Proceedings of the IEEE international symposium on circuits and systems, pp 459–462
265. Chao KC-H, Nadeem S, Lee WL, Sodini CG (1990) A higher order topology for interpolative modulators for oversampling A/D converters. IEEE Trans Circ Syst 37:309–318
266. Williams III LA, Wooley BA (1994) A third-order sigma-delta modulator with extended dynamic range. IEEE J Solid-State Circ 29:193–202
267. Christen T, Burger T, Huang Q (2007) A 0.13μm CMOS EDGE/UMTS/WLAN tri-mode ADC with -92dB THD. In: International solid-state circuits conference, digest of technical papers, pp 240–241
268. Breems LJ, Rutten R, van Veldhoven RHM, van der Weide G (2007) A 56 mW continuous-time quadrature cascaded $\Sigma\Delta$ modulator with 77 dB DR in a near zero-IF 20 MHz Band. IEEE J Solid-State Circ 42:2696–2705
269. Kulchycki SD, Trofin R, Vleugels K, Wooley BA (2008) A 77-dB dynamic range, 7.5-MHz hybrid continuous-time/discrete-time cascaded sigma delta modulator. IEEE J Solid-State Circ 43:796–804
270. Breems LJ, Dijkmans EC, Huijsing JH (2001) A quadrature data-dependent DEM algorithm to improve image rejection of a complex $\Sigma\Delta$ modulator. IEEE J Solid-State Circ 36:1879–1886
271. van der Zwan EJ, Dijkmans EC (1996) A 0.2 mW CMOS $\Sigma\Delta$ modulator for speech coding with 80 dB dynamic range. IEEE J Solid-State Circ 31:1873–1880
272. Keller M, Buhmann A, Sauerbrey J, Ortmanns M, Manoli Y (2008) A comparative study on excess-loop-delay compensation techniques for continuous-time sigmadelta modulators. IEEE Trans Circ Syst I 55:3480–3487
273. Philips KJP (2005) $\Sigma\Delta$ AD conversion for signal conditioning. Ph.D. thesis Technical University Eindhoven
274. Philips K, Nuijten PACM, Roovers RLJ, van Roermund AHM, Chavero FM, Pallares MT, Torralba A (2004) A continuous-time SD ADC with increased immunity to interferers. IEEE J Solid-State Circ 39:2170–2178
275. Nguyen K, Adams R, Sweetland K, Chen H (2005) A 106-dB SNR hybrid oversampling analog-to-digital converter for digital audio. IEEE J Solid-State Circ 40:2408–2415
276. Shettigar P, Pavan S (2012) A 15mW 3.6GS/s CT-SD ADC with 36MHz bandwidth and 83dB DR in 90nm CMOS. In: International solid-state circuits conference, digest of technical papers, pp 156–157
277. Bolatkale M, Breems LJ, Rutten R, Makinwa KAA (2011) A 4 GHz continuous-time ADC with 70 dB DR and 74 dBFS THD in 125 MHz BW. IEEE J Solid-State Circ 46:2857–2868
278. Shibata1 H, Schreier R, Yang W, Shaikh A, Paterson D, Caldwell1 T, Alldred D, Lai PW (2012) A DC-to-1GHz tunable RF SD ADC achieving DR = 74dB and BW = 150MHz at f0 = 450MHz Using 550mW. In: IEEE international solid-state circuits conference, digest of technical papers, pp 150–151
279. Adams RW (1986) Design and implementation of an audio 18-bit analog-to-digital converter using oversampling techniques. J. Audio Eng Soc 34:153–166
280. van Veldhoven RHM, Minnis BJ, Hegt HA, van Roermund AHM (2002) A 3.3-mW $\Sigma\Delta$ modulator for UMTS in 0.18-μm CMOS with 70-dB dynamic range in 2-MHz bandwidth. IEEE J Solid-State Circ 37:1645–1652
281. Lipshitz SP, Vanderkooy J (2001) Why 1-bit sigma-delta conversion is unsuitable for high-quality applications. In: Paper 5395 in 110th convention of the audio engineering society, Amsterdam
282. Silva PGR, Breems LJ, Makinwa K, Roovers R, Huijsing JH (2007) An IF-to-baseband $\Sigma\Delta$ modulator for AM/FM/IBOC radio receivers with a 118 dB dynamic range. IEEE J Solid-State Circ 42:1076–1089
283. 292 Ouzounov SF, Roza E, Hegt JA, van de Weide G, van Roermund AHM (2006) Analysis and design of high-performance asynchronous sigma delta modulators with binary quantizer. IEEE J Solid-State Circ 41:588–596

284. Park H, Nam KY, Su DK, Vleugels K, Wooley BA (2009) A 0.7-V 870-μW digital-audio CMOS sigma-delta modulator. IEEE J Solid-State Circ 44:1078–1088
285. Engelen J, van de Plassche R, Stikoort E, Venes A (1999) A sixth-order continuous-time bandpass sigma-delta modulator for digital radio IF. IEEE J Solid-State Circ 34:1753–1764
286. Schreier R, Abaskharoun N, Shibata H, Mehr I, Rose S, Paterson D (2006) A 375mW quadrature bandpass delta sigma ADC with 90dB DR and 8.5MHz BW at 44MHz. In: International solid-state circuits conference, digest of technical papers, pp 141–142
287. Ryckaert J et al (2009) A 2.4GHz Low-Power Sixth-Order RF bandpass SD Converter in CMOS. IEEE J Solid-State Circ 44:2873–2880
288. 289 Harrison J, Nesselroth M, Mamuad R, Behzad A, Adams A, Avery S (2012) An LC bandpass SD ADC with 70dB SNDR over 20MHz bandwidth using CMOS DACs. In: International solid-state circuits conference, digest of technical papers, pp 146–147
289. Chae H, Jeong J, Manganaro G, Flynn M (2012) A 12mW low-power continuous-time bandpass SD modulator with 58dB SNDR and 24MHz bandwidth at 200MHz IF. In: International solid-state circuits conference, digest of technical papers, pp 148–149
290. van der Zwan EJ, Philips K, Bastiaansen CAA (2000) A 10.7-MHz IF-to-baseband $\Sigma\Delta$ A/D conversion system for AM/FM radio receivers. IEEE J Solid-State Circ 35:1810–1819
291. Bergveld HJ, van Kaam KMM, Leenaerts DMW, Philips KJP, Vaassen AWP, Wetzker G (2004) A low-power highly-digitized receiver for 2.4-GHz-band GFSK applications. In: IEEE radio frequency integrated circuits (RFIC) symposium, pp 347–350
292. 283 van Veldhoven RHM (2003) A triple-mode continuous-time $\Sigma\Delta$ modulator with switched-capacitor feedback DAC for a GSM-EDGE/ CDMA2000/UMTS receiver. IEEE J Solid-State Circ 38:2069–2076
293. Kavusi S, Kakavand H, El Gamal A (2006) On incremental sigma-delta modulation with optimal filtering. IEEE Trans Circ Syst I: Fundam Theor Appl 53:1004–1015
294. Irons FH (2000) The noise power ratio theory and adc testing. IEEE Trans Instrum Meas 49:659–665
295. Milor LS (1998) A tutorial introduction to research on analog and mixed-signal circuit testing. IEEE Trans Circ Syst-II 45:1389–1407
296. Veendrick H (1998) Deep submicron CMOS ICs. Kluwer, Deventer. ISBN: 90-557-612-81 (Revised edition: H. Veendrick, Nanometer CMOS ICs, Springer, Deventer ISBN: 978-1-4020-8332-7, 2008)
297. Langevelde Rv (2003) RF performance and modeling of CMOS devices. In: Educational sessions workbook of custom integrated circuits conference
298. Pelgrom MJM, Vertregt M (1997) CMOS technology for mixed signal ICs. Solid-State Electron 41:967–974
299. Gregor RW (1992) On the relationship between topography and transistor matching in an analog CMOS technology. IEEE Trans Electron Devices 39:275–282
300. Stathis JH, Zafar S (2006) The negative bias temperature instability in MOS devices: a review. Microelectron Reliab 46:270–286
301. Hook TB, Brown J, Cottrell P, Adler E, Hoyniak D, Johnson J, Mann R (2003) Lateral ion implant straggle and mask proximity effect. IEEE Trans Electron Devices 50:1946–1951
302. Drennan P, Kniffin M, Locascio D (2006) Implications of proximity effects for analog design. In: IEEE custom integrated circuits conference, pp 169–176
303. Bianchi RA, Bouche G, Roux-dit-Buisson O (2002) Accurate modeling of trench isolation induced mechanical stress effects on MOSFET electrical performance. In: Technical digest international electron devices meeting, pp 117–120
304. Su K-Wi et al (2003) A scaleable model for STI mechanical stress effect on layout dependence of MOS electrical characteristics. In: Proceedings of the IEEE custom integrated circuits conference, pp 245–248
305. Wils N, Tuinhout HP, Meijer M (2009) Characterization of STI edge effects on CMOS variability. IEEE Trans Semiconductor Manufacturing 22:59–65
306. Ge L, Adams V, Loiko K, Tekleab D, Bo X-Z, Foisy M, Kolagunta V, Veeraraghavan S (2007) Modeling and simulation of poly-space effects in uniaxially-strained etch stop layer stressors. In: IEEE international SOI conference, pp 25–26

307. Tuinhout HP, Pelgrom MJM, Penning de Vries R, Vertregt M (1996) Effects of metal coverage on MOSFET matching. In: Technical digest international electron devices meeting, pp 735–739
308. Tuinhout HP, Bretveld A, Peters WCM (2004) Measuring the span of stress asymmetries on high-precision matched devices. In: International conference on microelectronic test structures, pp 117–122
309. Tuinhout HP, Vertregt M (2001) Characterization of systematic MOSFET current factor mismatch caused by metal CMP dummy structures. IEEE Trans Semiconductor Manufacturing 14:302–310
310. Pelgrom MJM, Vertregt M, Tuinhout HP Matching of MOS transistors. MEAD course material, 1998–2009
311. Lam M-F, Tammineedi A, Geiger R (2001) Current mirror layout strategies for enhancing matching performance. Analog Integr Circ Signal Process 28:9–26
312. Tuinhout HP, Elzinga H, Brugman JT, Postma F (1994) Accurate capacitor matching measurements using floating-gate test structures. In: IEEE international conference on microelectronic test structures, pp 133–137
313. Tuinhout HP, van Rossem F, Wils N (2009) High-precision on-wafer backend capacitor mismatch measurements using a benchtop semiconductor characterization system. In: IEEE international conference on microelectronic test structures, pp 3–8
314. Drennan PG (1999) Diffused resistor mismatch modeling and characterization. In: Bipolar/BiCMOS circuits and technology meeting, pp 27–30
315. Tuinhout HP, Hoogzaad G, Vertregt M, Roovers R, Erdmann C (2002) Design and characterisation of a high precision resistor ladder test structure. In: IEEE international conference on microelectronic test structures, pp 223–228
316. Lakshmikumar KR, Hadaway RA, Copeland MA (1986) Characterization and modeling of mismatch in MOS transistors for precision analog design. IEEE J Solid-State Circ 21: 1057–1066
317. Michael C, Ismail M (1992) Statistical modeling of device mismatch for analog MOS integrated circuits. IEEE J Solid-State Circ 27:154–166
318. Forti F, Wright ME (1994) Measurement of MOS current mismatch in the weak inversion region. IEEE J Solid-State Circ 29:138–142
319. Croon JA, Sansen W, Maes HE (2005) Matching properties of deep sub-micron MOS transistors. Springer, Dordrecht, The Netherlands. ISBN: 0-387-24314-3
320. Tuinhout HP (2003) Improving BiCMOS technologies using BJT parametric mismatch characterisation. In: Bipolar/BiCMOS circuits and technology meeting, pp 163–170
321. Brown AR, Roy G, Asenov A (2007) Poly-Si-gate-related variability in decananometer MOSFETs with conventional architecture. IEEE Trans Electron Devices 54:3056–3063
322. Mizuno T, Okamura J, Toriumi A (1994) Experimental study of threshold voltage fluctuation due to statistical variation of channel dopant number in MOSFETs. IEEE Trans Electron Devices 41:2216–2221
323. Bastos J, Steyaert M, Roovers R, Kinget P, Sansen W, Graindourze B, Pergoot A, Janssens E (1995) Mismatch characterization of small size MOS transistors. In: IEEE international conference on microelectronic test structures, pp 271–276
324. Stolk PA, Widdershoven FP, Klaassen DBM (1998) Modeling statistical dopant fluctuations in MOS transistors. IEEE Trans Electron Devices 45:1960–1971
325. Andricciola P, Tuinhout HP (2009) The temperature dependence of mismatch in deep-submicrometer bulk MOSFETs. IEEE Electron Device Lett 30:690–692
326. Hook TB, Johnson JB, Cathignol A, Cros A, Ghibaudo G (2011) Comment on (channel length and threshold voltage dependence of a transistor mismatch in a 32-nm HKMG technology). IEEE Trans Electron Devices 58:1255–1256
327. Croon JA, Tuinhout HP, Difrenza R, Knol J, Moonen AJ, Decoutere S, Maes HE, Sansen W (2002) A comparison of extraction techniques for threshold voltage mismatch. In: IEEE international conference on microelectronic test structures, pp 235–240

328. Tuinhout HP (2005) Electrical characterisation of matched pairs for evaluation of integrated circuit technologies. Ph.D. Thesis Delft University of Technology. (http://repository.tudelft.nl/file/82893/025295)
329. Takeuchi K, Hane M (2008) Statistical compact model parameter extraction by direct fitting to variations. IEEE Trans Electron Devices 55:1487–1493
330. Cheng B, Roy S, Asenov A (2007) Statistical compact model parameter extraction strategy for intrinsic parameter fluctuation. In: Grasser T, Selberherr S (eds) Simulation on semiconductor processes and devices. Springer, Wien, pp 301–304
331. Vertregt M, Scholtens PCS (2004) Assessment of the merits of CMOS technology scaling for analog circuit design. In: 30th European solid-state circuits conference, pp 57–64
332. Pelgrom MJM (1994) Low-power high-speed A/D conversion. In: 20th European solid-state circuits conference, low-power workshop
333. Kinget P, Steyaert M (1996) Impact of transistor mismatch on the speed accuracy power trade-off. In: Custom integrated circuits conference
334. Pelgrom MJM, Tuinhout HP, Vertregt M (1998) Transistor matching in analog CMOS applications. In: International electron devices meeting, pp 915–918
335. Doorn TS, ter Maten EJW, Croon JA, Di Bucchianico A, Wittich O (2008) Importance sampling Monte Carlo simulations for accurate estimation of SRAM yield. In: 34th European solid-state circuits conference, pp 230–233
336. Dijkstra E, Nys O, Blumenkrantz E (1995) Low power oversampled A/D converters. In: van de Plassche RJ (ed) Advances in analog circuit design. Kluwer Academic Publishers, Norwell, p 89
337. Pelgrom MJM (1998) Low-power CMOS data conversion. In: Sanchez-Sinencio E, Andreou A (eds) Low-voltage low-power integrated circutis and systems. IEEE Press, New York. ISBN: 0-7803-3446-9
338. Walden RH (1999) Analog-to-digital converter survey and analysis. IEEE J Select Areas in Commun 17:539–550
339. Vittoz E (1995) Low power low-voltage limitations and prospects in analog design. In: van de Plassche RJ (ed) Advances in analog circuit design. Kluwer Academic Publishers, Norwell, p 3
340. Giotta D, Pessl P, Clara M, Klatzer W, Gaggl R (2004) Low-power 14-bit current steering DAC for ADSL applications in 0.13 μm CMOS. In: European solid-state circuits conference, pp 163–166

About the Author

Marcel Pelgrom received his B.EE, M.Sc, and PhD from Twente University, Enschede, The Netherlands. In 1979, he joined Philips Research Laboratories, where his research has covered topics such as charge coupled devices, MOS matching properties, analog-to-digital conversion, digital image correlation, and various analog building block techniques. He has headed several project teams and was a team leader for high-speed analog-to-digital conversion. From 1996 until 2003, he was a department head for mixed-signal electronics. In addition to various activities concerning industry-academic relations, he is a research fellow in research on the edge of design and technology. In 2003, he spent a sabbatical in Stanford University where he served as a consulting professor. Since 2007, he has been a member of the technical staff at NXP Semiconductors. Dr. Pelgrom was an IEEE Distinguished Lecturer and has written over 40 publications and seven book chapters, and he holds 35 US patents. He currently lectures at Delft University and Twente University and for MEAD Inc.

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4,

© Springer Science+Business Media, LLC 2013

Index

M.J.M. Pelgrom, *Analog-to-Digital Conversion*, DOI 10.1007/978-1-4614-1371-4,
© Springer Science+Business Media, LLC 2013