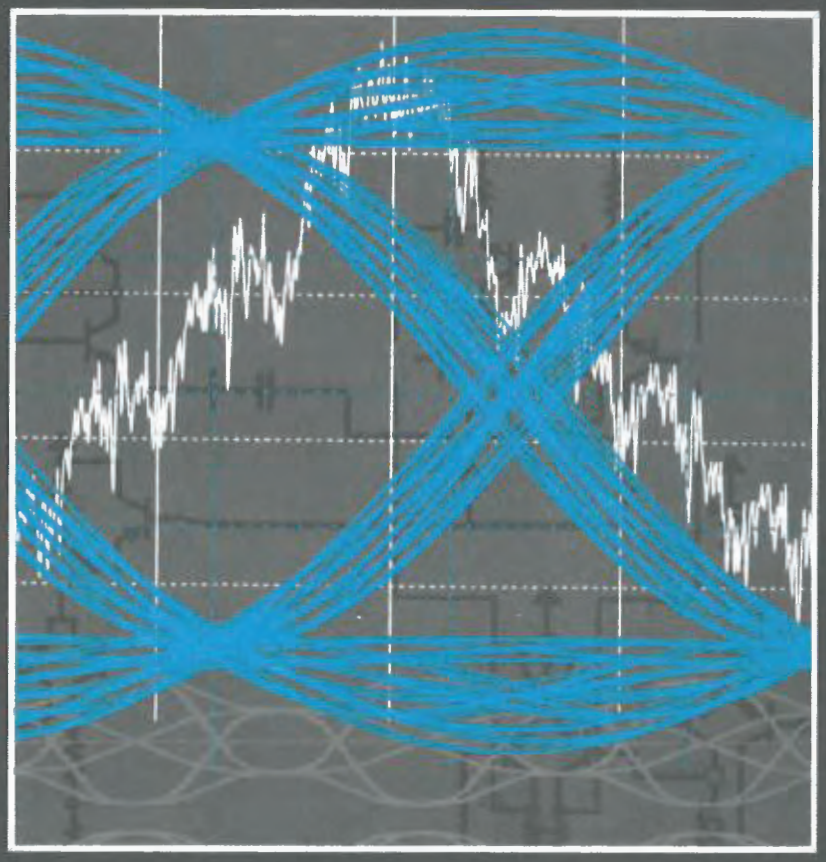
MOBILE COMMUNICATIONS SERIES

RF AND MICROWAVE CIRCUIT DESIGN FOR WIRELESS COMMUNICATIONS



LAWRENCE E. LARSON, EDITOR

Artech House Publishers BOSTON • LONDON

IPR2023-00697
Petitioners Lenovo (United States) Inc. and Motorola Mobility LLC - Ex.1017

h House Mobile Communications Library,
back of this book.

RF AND MICROWAVE CIRCUIT DESIGN FOR WIRELESS COMMUNICATIONS

Lawrence E. Larson
editor

Artech House
Boston • London

Library of Congress Cataloging-in-Publication Data
RF and microwave circuit design for wireless communications / Lawrence E. Larson, editor.
p. cm.
Includes bibliographical references and index.
ISBN 0-89006-818-6 (alk. paper)
1.Wireless communication systems. 2. Radio circuits. I. Larson, Lawrence E..
TK5103.2.R4 1996
621.384'12—dc21
97-49987
CIP

British Library Cataloguing in Publication Data
RF and microwave circuit design for wireless communications
1. Radio circuits—Design 2. Microwave sircuits—Design 3. Microwave communication systems 4. Wireless communication systems
I. Title
621.3'8412

ISBN 0-89006-818-6

© 1997 ARTECH HOUSE, INC.
685 Canton Street
Norwood, MA 02062

All rights reserved. Printed and bound in the United States of America. No part of this bo may be reproduced or utilized in any form or by any means, electronic or mechanical, inclu ing photocopying, recording, or by any information storage and retrieval system, witho permission in writing from the publisher.

All terms mentioned in this book that are known to be trademarks or service marks ha been appropriately capitalized. Artech House cannot attest to the accuracy of this inform tion. Use of a term in this book should not be regarded as affecting the validity of any trad mark or service mark.

International Standard Book Number: 0-89006-818-6
Library of Congress Catalog Card Number: 97-49987

10 9 8 7

ion Data
-less communications / Lawrence E.

index.

2. Radio circuits. I. Larson, Lawrence E..

97-49987
CIP

Data
:ss communications
›wave sircuits—Design 3. Microwave
-mmunication systems

the United States of America. No part of this book
ı or by any means, electronic or mechanical, includ-
information storage and retrieval system, without

are known to be trademarks or service marks have
ouse cannot attest to the accuracy of this informa-
ɔt be regarded as affecting the validity of any trade-

39006-818-6
er: 97-49987

▼▼▼

CONTENTS

Preface xi

Chapter 1 An Overview 1
1.1 Introduction 1
1.2 Markets and Frequencies for Wireless Communications 2
1.3 Circuit Technology Options for Wireless RF/Microwave Circuit Implementation 8
1.3.1 Semiconductor Devices 8
1.3.2 Passive Devices 10
1.4 Recent Examples of RF and Microwave Wireless Communications Circuits 10
1.5 Organization of the Book 13
References 15

Chapter 2 Introduction to Wireless Communications Applications and Circuit Design 17
2.1 Introduction 17
2.2 Current and Nascent Consumer Wireless Communications Standards and Applications 18
2.3 Radio Transceiver Technology Requirements 22

2.4 RF Component Requirements for Transceivers 27
2.4.1 Duplexer Requirements 28
2.4.2 Amplifier Requirements 28
2.4.3 Mixer Requirements 29
2.4.4 Filter Requirements 29
2.4.5 Frequency Synthesizer Requirements 30
2.5 Systematic Analysis of Transceiver Design 30
2.6 Digital Communications Requirements 33
2.7 Electromagnetic Compatibility in Wireless Devices 38
2.8 Conclusions 40
References 40

Chapter 3 Low-Power Radio Frequency ICs for Portable Communications 43
3.1 Introduction 43
3.2 Key Signal-Processing Issues in Wireless Transceivers 45
3.3 ICs in Broadcast Radio Receivers 51
3.4 ICs in Wireless Paging Receivers 57
3.5 ICs in Cellular Telephone Transceivers 60
3.6 Digital Cellular and Cordless Telephones 67
3.7 Direct-Conversion Transceivers 73
3.7.1 Direct-Conversion Single-Sideband Synthesizers 73
3.7.2 Direct-Conversion Receivers for GSM Digital Cellular Telephones 75
3.7.3 Local Oscillators with Quadrature Outputs 76
3.8 ICs for Spread-Spectrum Wireless Transceivers 78
3.9 ICs to Enable Future Transceivers 81
3.10 Conclusions 87
Acknowledgment 88
References 89

Chapter 4 Digital Modulation And Demodulation 99
4.1 Why Digital Modulation? 99
4.2 Phase Modulation 101
4.2.1 Binary Phase-Shift Keying (BPSK) 101
4.2.2 Quadriphase (Or Quadrature) Phase-Shift Keying (QPSK) 115
4.2.3 Offset Or Staggered Phase-Shift Keying (OQPSK) 134
4.2.4 8-PSK and 16-PSK 140
4.2.5 Nyquist Filtering as Applied to PSK Modulations 145
4.3 Frequency Modulation 155
4.3.1 Minimum Shift Keying (MSK) 156
4.3.2 Frequency Shift Keying (FSK) 166

4.4 Related Modulations 173
4.4.1 Gaussian-Filtered MSK (GMSK) 175
4.4.2 $\pi/4$-Shifted Differentially Encoded PSK ($\pi/4$-DQPSK) 193
4.4.3 Feher's QPSK (FQPSK) 204
4.5 Examples of Integrated Circuit Digital Communications Circuits 208
4.5.1 Design of Modulators and Demodulators for Wireless Communications Systems 214
4.6 Summary 219
References 222

Chapter 5 Mixers for Wireless Applications 225
5.1 Introduction 225
5.2 Properties of Mixers 226
5.2.1 Frequency Mixing 226
5.2.2 Practical Characteristics of Mixers 227
5.3 Semiconductor Devices for Mixers 231
5.3.1 Schottky-Barrier Diodes 232
5.3.2 Field-Effect Transistors 235
5.4 Frequency Mixing 242
5.5 Diode Mixers 244
5.5.1 Diode Mixer Characteristics 244
5.5.2 Single-Diode Mixers 245
5.5.3 Baluns and Hybrids for Wireless Mixers 248
5.5.4 Singly Balanced Diode Mixers 255
5.5.5 Doubly Balanced Diode Mixers 258
5.5.6 Image-Rejection Mixers 262
5.6 Active FET Mixers 263
5.6.1 Single-Device Active FET Mixers 263
5.6.2 Single-Device Mixer Circuits 266
5.6.3 Dual-Gate FET Mixer 268
5.6.4 Singly Balanced FET Mixers 269
5.6.5 Doubly Balanced FET Mixers 272
5.7 Passive FET Mixers 273
5.7.1 Single-Device FET Resistive Mixers 274
5.7.2 Balanced FET Resistive Mixers 276
5.7.3 FET Resistive Ring Mixer 277
5.7.4 Modeling FETs for Resistive Mixer Analysis 279
5.7.5 Mixer Comparison Table 280
5.8 Harmonic-Balance Simulation of Mixers 280
References 283

Chapter 6 RF and Microwave Frequency Synthesizer Design 285
6.1 Introduction 285
6.2 Frequency Synthesizer Fundamentals 287
6.2.1 Frequency Synthesizer Figures of Merit 289
6.3 Voltage Controlled Oscillators 292
6.3.1 Introduction 292
6.3.2 General Oscillator Design 292
6.3.3 Colpitts and Hartley Oscillators 295
6.3.4 Clapp Oscillator 299
6.3.5 Crystal Oscillators 300
6.3.6 Oscillator Configurations for Microwave Frequencies 302
6.3.7 Dielectric Resonator Oscillators (DROs) 304
6.3.8 Practical Considerations for Oscillator Design 308
6.3.9 Voltage Control of Oscillators 309
6.3.10 Noise in Oscillators 310
6.3.11 Ring Oscillator VCO Designs 313
6.3.12 Multivibrator VCO Designs 317
6.4 Phase Lock-Loop Design 320
6.4.1 PLL Design Fundamentals 320
6.4.2 PLL Applications 326
6.4.3 Implementation of PLLs 329
6.5 Direct Digital Synthesis (DDS) 338
6.5.1 DDS Architectures and Design Considerations 340
References 343

Chapter 7 Amplifiers for Wireless Communications 345
7.1 Introduction 345
7.2 Wideband Amplifier Techniques 346
7.2.1 Gain-Bandwidth Product 346
7.2.2 Gain-Bandwidth of Practical Circuits 353
7.2.3 Wideband Matched-Impedance Amplifiers 359
7.3 Low-Noise Amplifiers 365
7.3.1 Guidelines for Low-Noise Amplifier Design 365
7.3.2 Practical Design of a Low-Noise Amplifier 367
7.4 Variable-Gain Amplifiers 369
7.4.1 Dynamic Range 370
7.4.2 Emitter-Driven and Base-Driven Gain-Control 371
7.4.3 Basic Circuit Configurations 374
7.4.4 Variable-Gain Circuit Configurations 378
7.5 Power Amplifiers 382
7.5.1 Class A Power Amplifier Design 384
7.5.2 Higher Efficiency Power Amplifiers 388

nthesizer Design 285
285
287
of Merit 289
292
292
292
ors 295
299
300
Microwave Frequencies 302
ors (DROs) 304
Oscillator Design 308
s 309
310
313
317
320
320
326
329
338
n Considerations 340
343
ications 345
345
346
346
Circuits 353
ce Amplifiers 359
365
nplifier Design 365
oise Amplifier 367
369
370
ven Gain-Control 371
374
gurations 378
382
sign 384
plifiers 388

7.5.3 Maximum Power Gain and Stability 391
7.5.4 Practical Design Issues 392
References 394
Glossary 397
About the Editor 401
Index 403

▼▼▼

PREFACE

The rapid development of commercial markets for analog and digital wireless communications in the last few years has been a blessing to many of the technical professionals hit hard by the defense build-down of the late 1980s and early 1990s. A new market for high-frequency expertise has risen out of the ashes of what was considered a moribund field only a few years ago.

However, the technology required to meet these new markets is very different from that of the older disciplines of DOD RF/microwave electronics—with its emphasis on "performance at any price"—or analog radio engineering—with its emphasis on discrete analog solutions operating at relatively high power levels. This gap between the traditional techniques of RF communication engineering and the more modern disciplines required for wireless communications presents a real difficulty for the practicing engineer. This book was developed with the hope of bridging the gap between the traditional art of radio and RF engineering and the requirements of today's modern wireless communications systems.

Modern wireless communications systems require a thorough understanding of RF and microwave circuit techniques *in addition* to a background in digital communications techniques and familiarity with existing and emerging wireless communications protocol standards. As a result, the first half of this book presents an overview of wireless communications standards, digital communications techniques, and RF system issues related to wireless communications and low-power RF and

microwave electronics. The second half of the book covers key circuit areas relate to the implementation of wireless communications systems.

This book is organized with the purpose of guiding the reader to the mos essential elements of RF and microwave circuit design for wireless communications and implicitly assumes a knowledge of the basics of active circuit design and commu nications theory. In any field advancing as rapidly as this one, there is a danger tha a narrowly focused book will become rapidly dated. In order to avoid that, eac chapter contains a clear exposition of the fundamental aspects of the topic, i addition to more topical material related specifically to wireless communication circuits. It is our hope that the fundamental portions of the book will retain thei usefulness well after some of the more topical material has lost its freshness.

In addition, the book emphasizes *monolithic* solutions to microwave/RF circui design problems. This is because the industry-wide trend is to move to higher an higher levels of integration in order to meet the aggressive cost targets associate with the commercial/consumer market. It is our belief that even if the designer is n actually designing the monolithic circuits, the design of RF systems *using* monolithi circuits will be greatly enhanced by an understanding of monolithic circuit technique However, our intent is not to exclude hybrid design techniques, and they are als addressed where appropriate, especially in the chapters on mixer and amplifi design.

Chapter 1 presents an overview of the general area of wireless communication the market, frequencies, and technologies required to address it. The service addressed by the wireless industry include the ubiquitous cellular telephone an pager, in addition to the emerging markets of PCS, satellite-based telephony an data transmission, low-cost microwave links ,and wireless local area networks. I a field that is changing as rapidly as this, some of the information is evolving eve as we finish production of the book, but we have tried to make this material as u to date and thorough as possible.

Chapter 2 of the book presents an overview of wireless communications stan dards and the RF system design considerations associated with their implementatio These new standards cover voice and data transmission over a variety of frequenc bands in the United States as well as Europe and Asia. The chapter makes the poi that traditional RF design approaches often fall short of meeting the requiremen of modern digital wireless communications products, and presents some illustrativ examples where different design methodologies are required to meet these ne requirements.

Chapter 3 of this book presents an overview of the general field of wirele RF integrated circuit design, with particular emphasis on the historical developme of radio receiver architectures, receiver design, and low-power techniques. Th chapter develops the argument that modern communications system designs wi rely on direct downconversion and other novel architectural approaches to reali

the low-power/low-cost requirements of handheld consumer communications equipment.

Chapter 4 presents a detailed overview of digital RF modulation techniques and their circuit implementations. This material is normally covered in digital communications textbooks. Here, we have developed a one-chapter summary of the key results of that field, with particular emphasis on those modulation schemes that are currently being implemented in modern communications systems. The chapter also adopts the practical approach of outlining the methods of implementing those modulation techniques with a variety of circuit approaches.

Chapter 5 presents a thorough review of the design and fabrication of RF and microwave mixers. Once again, this field is often covered in an entire text. We have chosen here to cover those areas of mixer design particular to the field of wireless communications and have emphasized practical circuit approaches to mixer design, rather than theoretical depth.

Chapter 6 presents a complete overview of frequency synthesizer designs for wireless communications, with particular emphasis on VCO, PLL, and DDS implementations. The circuits often form the "heart" of the communications system in the sense that they provide the modulation and demodulation frequencies for tuning the transmitted and received signals. As a result, their performance and cost are crucial to the realization of acceptable system performance.

Finally, Chapter 7 presents a complete overview of the field of RF and microwave amplifier design. Although the design of amplifiers is a relatively mature field, a variety of new circuit architectures have been developed in the last few years that address the particular requirements presented by the wireless communications field. These requirements include an extremely high dynamic range in conjunction with low-power operation, as well as high efficiency in power output stages in conjunction with low-voltage operation. In addition, monolithic solutions are now being pursued where a hybrid approach would have been standard only a few years earlier.

The development of this book would not have been possible without the efforts of many of my friends and colleagues. In particular, I would like to especially thank all of the authors of the various chapters in the book. They worked tirelessly to produce manuscripts of exquisite quality with a very tight schedule. I would especially like to thank Dr. Paul Greiling of the Hughes Research Labs for his unflagging support of my work over the many years of our association. The long support and friendship of Professor Gabor Temes of Oregon State University is also greatly appreciated. I would also like to thank Drs. Michael Case, David Rensch, Loi Nguyen, Adele Schmitz, Mehran Matloubian, Perry MacDonald, Julia Brown, Ron Lundgren, William Stanchina, Robert Walden, Conilee Kirkpatrick, Cynthia Baringer, William Hooper, and Joseph Jensen of the Hughes Research Laboratories for many enlightening discussions in the area of high-frequency electronics. I would also like to acknowledge the support and encouragement of Mr. Lawrence Blue, Thomas Jackson, Olu George, Steven Rosenbaum, Adam Rachlin, and Ron Coburn from

Hughes Network Systems. I would also like to acknowledge many valuable discussions with Dr. Fred Blum of Sequel. I am indebted to the management of Hughes Electronics—particularly Drs. Arthur Chester, Richard Abrams, and Richard Reynolds of the Hughes Research Laboratories—who have supported my graduate education and career at Hughes.

Finally, I would like to acknowledge my wife Carole, my daughter Hannah, and my son Erik. Their love, support, and patience made the entire project possible.

Lawrence Larson
January 1996

CHAPTER 1

AN OVERVIEW

Lawrence E. Larson
Head, Telecommunications Technology Department
Hughes Research Laboratories
Malibu, CA

1.1 INTRODUCTION

The rapid development of wireless communications products over the last ten years has led to an explosion in interest in improved circuit design approaches in the radio frequency (RF) and microwave area. However, this field has developed so rapidly that traditional communications and analog circuit design textbooks have struggled to keep up with all of the advances in the field [1, 2]. The purpose of this book is to attempt to bridge the gap between traditional communications/circuit design textbooks, which develop the fundamental principles behind a variety of circuit design approaches, and the most up-to-date advances in the wireless communications field.

In particular, the development of the field is characterized by a shift from purely analog techniques to those employing digital or hybrid digital/analog approaches. The advantage of digital communications techniques are well established both theoretically and practically, and include more efficient use of bandwidth, greater immunity to impairment due to noise, and lower power dissipation [3]. In addition, digital communications techniques can exploit the relentless advance of digital integrated circuit (IC) technology to realize improved functionality at lower

power dissipation and lower cost. The migration of RF and microwave technology from analog to digital is briefly summarized in Table 1.1.

Since the widespread introduction of cellular telephone service during the 1980s, there has been rapid growth of those systems as well as mobile paging devices and a variety of wireless "data communications" services, under the broad heading of personal communication services (PCS). In addition to these terrestrial communications systems, the field of satellite-based video, telephone, and data communications services has grown at least as rapidly. The electromagnetic frequencies spanned by all of these systems range from several hundred megahertz to well over 60 GHz. The goal of all of this is to—as George Heilmeier put it in the October 1992 issue of *IEEE Communications Magazine*—"be able to access and communicate...in any medium, or combination of media—voice, data, image, video, or multimedia—anytime, anywhere, in a timely, cost-effective way" [4]. The growth of the market for mobile PCS products is expected to be rapid, as illustrated by the Table 1.2, which estimates market growth for a variety of mobile wireless services [5].

This chapter will briefly summarize the products and services that are anticipated to be developed in the coming years in this diverse field and conclude with some examples of the RF and microwave circuit design approaches that are being planned to meet these challenges.

1.2 MARKETS AND FREQUENCIES FOR WIRELESS COMMUNICATIONS

The majority of interest and activity for wireless communications systems is occurring at frequencies below 2 GHz, although there is increasing interest at frequencies above 2 GHz as well. The reasons that the majority of wireless systems are currently

Table 1.1
Comparison of Traditional Analog Communications RF/Microwave Circuit Approaches to Modern Digital Communications RF/Microwave Approaches

Criterion	*Traditional Analog Techniques*	*Modern Digital Techniques*
Level of integration	Hybrid (Low)	Monolithic (high)
Access method	FDMA	TDMA, CDMA
Down/up conversion	Dual	Direct
Power dissipation	High	Low

Table 1.2
Market Projections (in Millions of Users) for PCS Products

Product	*1991*	*1992*	*1993*	*1994*	*1995*	*1996*	*1997*	*1998*
Cellular	7.6	10.5	13.6	16.5	20.0	24.0	28.0	32.0
PCS	0.0	0.0	0.1	0.5	1.0	1.5	2.1	2.6
Paging	11.5	13.2	15.2	17.2	19.2	22.0	24.0	26.0
Mobile data	0.25	0.5	1.2	2.1	3.0	4.3	5.7	7.4
Mobile computers	0.0	0.0	0.023	0.075	0.255	0.795	1.515	2.595
PDAs	0.0	0.0	0.002	0.022	0.094	0.289	0.889	1.795

Source: [5].

implemented at frequencies below 2 GHz are both technical and historic. One technical reason for the use of frequencies below 2 GHz is that signal attenuation in the atmosphere and within buildings rises rapidly with frequency. Another reason is that the active circuits that are used to transmit and modulate the signals have historically had little gain at higher frequencies. This latter limitation is changing as advances in IC technology push transistor operating speeds well above 10 GHz. Figure 1.1 is a plot summarizing most of the popular available frequencies available for wireless communications throughout the world [6].

In the United States, the allocation and use of the electromagnetic spectrum is governed by the Federal Communications Commission (FCC). Historically, the spectrum was allocated to users based on a "comparative hearing" process, where the "best qualified" applicant was determined based on a lengthy and expensive

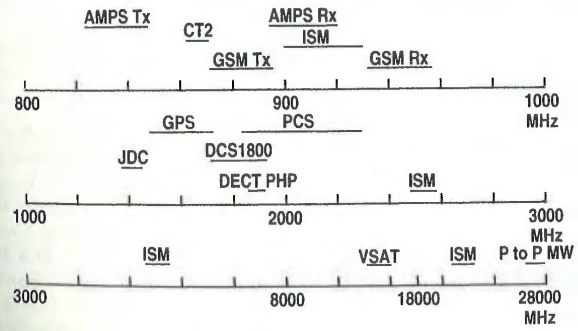


Figure 1.1 Summary of popular wireless transmission frequencies [6].

hearing process [7]. This process had its origins in the 1934 Communications Act, which established the FCC and defined the fundamental law that governed radio and television communications. However, Congress opened up the process of spectrum allocation with the 1993 Omnibus Budget Reconciliation Act, which permitted the FCC to grant spectrum licenses based on a competitive bidding process (an "auction"). The various wireless services encompassed by the FCC spectrum allocation include terrestrial telephony, PCS, satellite communications, and wireless LANs. Each of these areas is briefly discussed below.

Terrestrial Mobile Telephony

Chapter 2 of this book presents an excellent overview of spectrum allocation and data transmission standards for analog cellular, digital cellular, data, and specialized mobile radios (SMRs) (see Tables 2.1 through 2.4). In the United States, the existing advanced mobile phone service (AMPS) analog and IS-54 digital standard covers the frequency range from 824 MHz to 894 MHz. The global system for mobile communication (GSM) is the European digital telephone standard and covers the frequency range from 890 MHz to 960 MHz. GSM is also fast becoming a worldwide standard, with networks in South Africa, Singapore, and Australia to name but a few. Other standards have been proposed as well for future generations of telephones. For example, the Japanese Personal Handy Phone (PHP) is intended to operate at 1.9 GHz.

It is expected that, over the next several years, these telephone services will migrate from analog frequency division multiple access (FDMA) schemes to digital schemes. The resulting increase in cell capacity due to the more efficient bandwidth modulation of digital techniques will provide an added impetus to the implementation of these digital techniques.

PCS

Because of the relative infancy of the markets and services, frequency allocation and access methods for personal communications services (PCS) are highly fragmented. There is also some confusion about the services that will be provided by PCS, although the initial application of PCS in the United States is primarily for mobile telephony. Figure 1.2 shows the frequency allocations for PCS in the United States as of mid-1994, which extend from 900 MHz to well over 2 GHz, and are broken further into narrowband, broadband, and unlicensed services [7]. Narrowband PCS is primarily used for advanced paging services, while broadband PCS is intended for licensed voice, data, or video transmissions. Unlicensed PCS is intended for low-power, limited range devices, primarily for "premises communications."

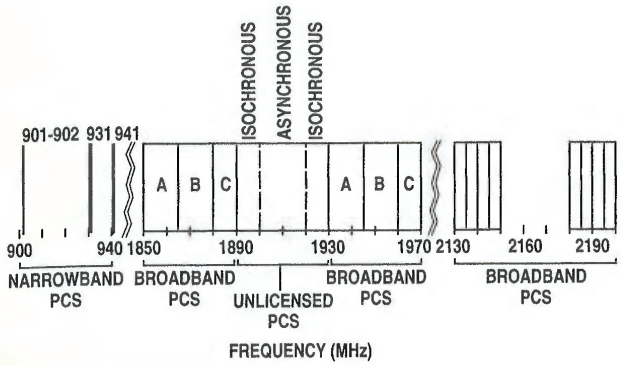


Figure 1.2 United States allocation of PCS frequencies [7].

Satellite Communications

Satellite communications is characterized by traditional voice, video, and data transmission using geosynchronous satellites (mostly at C- and Ku-band frequencies) and fixed earth terminals, and by a variety of proposed *mobile* satellite services (MSS) for handheld voice and data communication. Mobile satellite services are subdivided into geosynchronous (GEO), medium Earth orbit (MEO), and low Earth orbit (LEO), with the latter being subdivided into "little" LEO and "big" LEO. Big LEOs are intended to provide worldwide, or nearly worldwide, voice coverage, as well as limited high-speed data handling capabilities. Little LEOs are intended to provide lower bandwidth, less costly services to local niche markets using inexpensive satellites.

Geosynchronous mobile satellite services include Inmarsat, which is for general voice, data, and video services, and operates at a variety of frequencies from 1 to 6 GHz. OmniTRACs is another GEO MSS, which is used extensively for two-way truck fleet identification and operates in the 12/14-GHz band. In addition, the American Mobile Satellite Corporation has also proposed a GEO MSS system.

Over a dozen LEO systems have been proposed for MSS in the last several years [8]. They include the well-known Motorola Iridium system, as well as the Aries, Ellipso, and Globalstar systems. The Iridium system is a big LEO, and is planned to utilize 66 satellites in LEO, with the capability for global coverage. The Odyssey system, proposed by TRW, is a MEO System of 12 satellites.

Wireless LANs

The impetus for the development of wireless local area networks (LANs) comes from the increasing popularity of laptop computers and personal digital assistants

(PDAs), as well as the high cost of rewiring buildings for the inevitable next generation of wired LANs [9]. Labor and material costs can reach $1,000 per node for installation of copper wire networks, with an additional $200 to $1,000 for each upgrade [10].

So, there are enormous potential advantages if cost-effective, wide-bandwidth LANs can be developed. These networks are generally required to have all of the advantages of traditionally wired networks (high bandwidth, low latency, security, ease of configuration) in addition to several characteristics specific to the wireless environment:

- License-free operation (preferred use of nonlicensed bands);
- Long battery life (low power dissipation);
- Low interference (spread-spectrum techniques);
- Ease of configuration (good propagation characteristics within buildings and between buildings).

The most active frequency bands today for the development of wireless in-building LANs are the so-called industrial, scientific, and medical (ISM) bands at 902 to 928 MHz, 2,400 to 2,483.5 MHz, and 5,725 to 5,850 MHz. These bands were allocated by the FCC in the 1980s for unlicensed transmissions, as long as spread-spectrum transmission is employed and radiated power is less than 1W.

Each band has advantages and disadvantages for wireless LANs. The 2,400-MHz band is available in almost all countries, whereas the 900-MHz and 5,700-MHz bands are strictly limited to U.S. operation at this time. In addition, the 2,400-MHz band is prone to interference from microwave ovens. The 900-MHz band is the most mature technologically, but has a tremendous amount of potential interference from cordless telephones, wireless speakers, and so forth. The 5,700-MHz band has the widest potential bandwidth of any of the bands, but its in-building propagation characteristics are poor, and the analog technologies required to realize the system (GaAs or Si/SiGe) are the least mature. Table 1.3 summarizes the advantages and disadvantages of each of the ISM bands for wireless LAN systems [11].

The 2.4- to 2.5-GHz band has been under particularly active development, with several commercially available chip sets developed for these specific frequencies. The IEEE 802.11 Committee has been meeting for several years with the goal of developing broad hardware and software standards for the use of these frequencies. The overall goal is to support data rates in excess of 1 Mbps and to maintain compatibility with a wide variety of environments, from small rooms to buildings, campuses, and industrial manufacturing facilities. Both direct sequence spread spectrum (DSSS) and frequency hopped spread spectrum (FHSS) are provided for under these standards.

Table 1.3
Wireless LAN Band Alternatives: ISM Bands

Criterion	*915 MHz*	*2.4 GHz*	*5.8 GHz*
Frequencies	902–928 MHz	2.40–2.4835 GHz	5.725–5.825 GHz
Bandwidth	26 MHz	83.5 MHz	125 MHz
FCC licensing	No (Part 15)	No (Part 15)	No (Part 15)
Availability	US/Canada	Worldwide	US/Canada
Development cost	Low (Si)	Low/medium (Si, GaAs)	High (GaAs)
Transmission range	Greatest (power-dependent)	95% of 915 MHz	80% of 915 MHz
Current status	Crowded	Light use	Almost no use
Outlook	High growth	Medium growth	Low growth
Interference sources (U.S.)	Primary users	Primary users	Primary users
	Many LANs	Few LANs	Very few LANs
	Many non-SS	Few non-SS	Very few non-SS
		Microwave ovens	
Additional interference sources (worldwide)	Cellular telephones		Some radar

Motorola has taken a different approach with its Altair System, which operates at an 18-GHz frequency band granted solely to it by the FCC [12]. The advantages of this frequency band include a wide bandwidth as well as lack of potential interference by more common microwave radiators. Recent versions of the system boast a performance of > 10 Mbps, and implement Ethernet and TCP/IP protocols. However, the high-frequency operation of the system also increases its cost.

Higher frequencies than these are being utilized for outdoor "point-to-point" communications systems, with applications for cellular, private, and LAN applications. The most common of these links are from 10 to 38 GHz and have peak transmission rates of 10 to 20 Mbps. An example of a typical 23-GHz product from P-COM is shown in Figure 1.3.

Finally, even higher frequencies are being investigated for applications where wider data bandwidths are required (> 100 Mbps). In particular, the 60-GHz frequency band has been actively pursued for years by a number of groups [13]. The RACE Project in Europe is developing the hardware and software technology for broadband networks at 62 to 63 and 65 to 66 GHz. In addition to the broad-bandwidth capacity at this frequency, the high attenuation of the atmosphere at these frequencies makes frequency reuse more attractive than at lower frequencies [14].

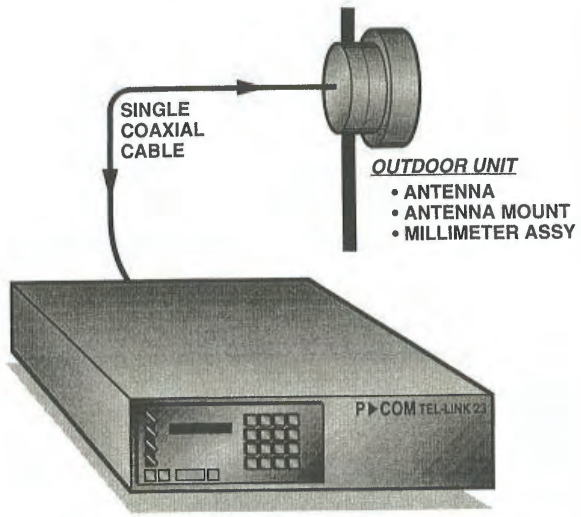


Figure 1.3 A 23-GHz wireless data product from P-COM.

1.3 CIRCUIT TECHNOLOGY OPTIONS FOR WIRELESS RF/MICROWAVE CIRCUIT IMPLEMENTATION

This section will briefly outline the circuit technology options for the implementation of wireless communications circuits.

1.3.1 Semiconductor Devices

The choice of the optimum semiconductor technology for wireless applications is a hotly contested issue, with various proponents arguing passionately for their positions [15]. In the last few years, wireless components have been announced or

proposed using silicon bipolar technology, complementary metal-oxide semiconductor (CMOS), Si/SiGe heterojunction bipolar transistors (HBTs), as well as GaAs MESFETs and HBTs [16].

The optimum technology choice is ultimately an economic rather than a technical issue, and whoever can realize components with the lowest cost and quickest time to market is destined to win the argument. However, the economic arguments in favor of one technology or another are even more contentious than the technical arguments, so we will confine ourselves here to the technical merits of the various technologies and let the marketplace sort out the eventual "winners."

The key transistor technology figures of merit for RF and microwave design of wireless communication products are:

1. Short-circuit current gain bandwidth (f_T) (important for both analog and digital circuits);
2. Maximum power gain bandwidth (f_{MAX}) (important for both analog and digital circuits);
3. Minimum noise figure (LNAs, wide dynamic range front ends, mixers);
4. Maximum power-added efficiency (power amplifiers);
5. Linearity (AGC amplifiers, front ends);
6. Yield (relates to price as well as performance).

These various figures of merit can be compared for a number of competing semiconductor technologies and appear in Table 1.4 [17].

This table demonstrates the confusing picture that confronts the designer of wireless communications systems. Unlike digital IC technology, where CMOS is clearly dominant, no one technology dominates (or is likely to dominate) the imple-

Table 1.4
Performance of Various Semiconductor Technologies Addressing Wireless Communications Applications

Criterion	*Si BJT*	*Si/SiGe HBT*	*GaAs MESFET*	*GaAs HBT*	*BiCMOS BJT*
B_{VCEO}/B_{VDS} (V)	4	4	8	15	6
f_T (GHz)	32	55	50	30	13
f_{MAX} (GHz)	35	55	60	70	11
G_{MAX} @ 2 GHz (dB)	24	28	20	19	17
Noise figure @ 2 GHz (dB)	1.0	0.5	0.3	1.5	1.0
IP3/P-1 dB (dB)	9	9	12	16	9
Power-added efficiency @ 3V (%)	–	70	60 @ 5V	70	40

Source: [17].

mentation of wireless communications ICs. In many cases, GaAs technology is used for efficient power amplification, and Si bipolar technology is used for RF signal-processing functions. More recently, Si/SiGe bipolar technology has been introduced as a promising technique for implementing both power amplification and signal processing with the same technology.

1.3.2 Passive Devices

Though often neglected, passive devices often determine the overall size, topology, and performance of a wireless communications transceiver. These elements include inductors, filters, resonators, and capacitors. However, as the size and cost of active devices continue to shrink, the importance of improvements in passive devices performance and size become more urgent. Many of these elements are difficult if not impossible to integrate monolithically.

One of the most important developments in the field of wireless communications was the development of surface acoustic wave (SAW) devices. These devices are typically bandpass filters with very sharp and stable cutoff characteristics. Their operation is based on the interference of mechanical surface waves on a piezoelectric substrate, which have been launched through piezoelectric transducers. The one drawback of SAW devices is their relatively high insertion loss (typically > 20 dB), which creates the need for pre- and postamplification around the circuit. Some low-loss SAW filters have been developed, but they are still in the experimental stage and are not yet commercially available. A drawing of a typical SAW filter appears in Figure 1.4.

In order to eliminate some of the disadvantages of SAW filters (principally the high insertion loss), a number of laboratories have been exploring alternative techniques for the implementation of passive elements. Figure 1.5 shows a dielectric monoblock filter intended for 800-MHz cordless telephone applications [18]. This design consists of a pair of coupled quarter-wave sections, mounted inside a high-dielectric coaxial waveguide. The ceramic material is $BaO\text{-}PbO\text{-}Nd_2O_3\text{-}TiO_2$, with a relative dielectric constant of 90. The filter has a center frequency of 886 MHz, a bandwidth of 2 MHz, and an insertion loss of 2.1 dB. The filter exhibited an out-of-band rejection of 44 dB at 44 MHz from the carrier.

Other important passive elements include dielectric resonators (discussed in Chapter 6), as well as surface-mount transformers and inductors (discussed in Chapter 5), capacitors, and resistors.

1.4 RECENT EXAMPLES OF RF AND MICROWAVE WIRELESS COMMUNICATIONS CIRCUITS

One of the keys to the recent development of wireless communications devices is the increased level of integration of RF and microwave devices. An RF function that

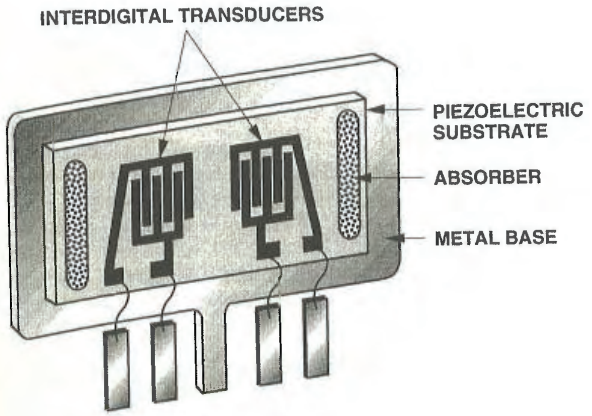


Figure 1.4 SAW filter used for communications transceivers.

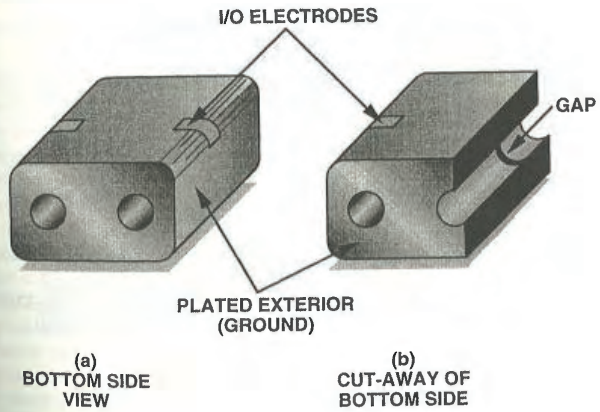


Figure 1.5 Dielectric monoblock filter used for cordless telephone applications [18].

would historically have been implemented with discrete devices and laboriously tuned by hand on the test bench is today implemented in a single monolithic die. Equipment manufacturers are able to reduce the size, weight, cost, and power consumption of the resulting circuits. A wide variety of monolithic circuits of this type have been demonstrated recently.

A good example of a highly integrated transmitter IC for digital cellular applications is demonstrated by Hewlett Packard in the block diagram of Figure 1.6 [19]. This circuit was designed for compatibility with the North American Digital Cellular (NADC) System at 824 to 849 MHz, as well as the GSM European System at 890 to 915 MHz and the Japan Digital Cellular (JDC) System at 940 to 956 MHz.

The circuit was realized in an RF silicon bipolar process, with a 15 GHz f_T and 30 GHz f_{MAX}. Other features of the process include gold metallization, a thick field oxide, and high-density MIS capacitors. Two of the most interesting features of the design are the use of on-chip LC bandpass filters for mixer product suppression and the use of external off-chip inductive tuning to peak the frequency response in the desired band.

ATT recently announced an integrated circuit specifically targeted for the GSM market [20]. A block diagram of the chip set is shown in Figure 1.7. The chip includes two fixed-frequency synthesizers (for transmit and receive IFs), an RF mixer, single IF downconversion with quadrature demodulation and variable gain, as well as a variety of digital control circuitry. The circuit was implemented in a 12-GHz silicon bipolar technology, with 1.5-μm feature size. The addition of a power amplifier and low-noise amplifier are all that are needed to realize the RF portions of a complete GSM transceiver.

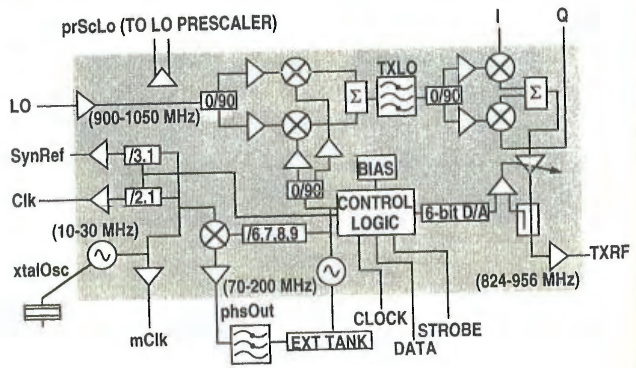


Figure 1.6 Integrated transmitter IC for digital cellular telephone application [19].

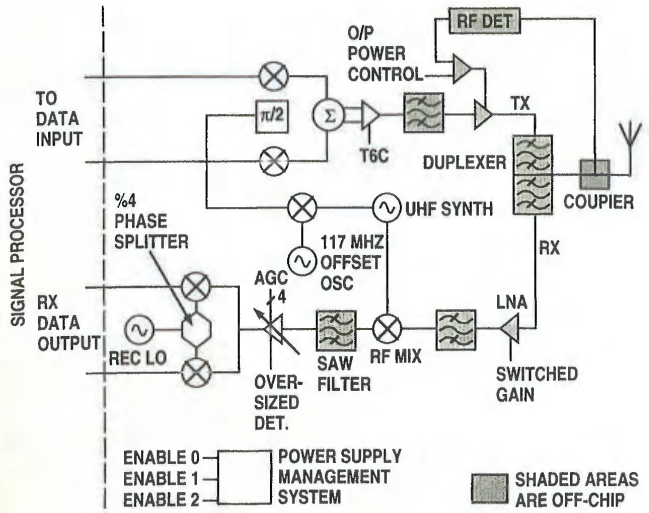


Figure 1.7 GSM chip set recently announced by ATT [20].

1.5 ORGANIZATION OF THE BOOK

This book is organized with the purpose of guiding the reader to the most essential elements of RF and microwave circuit design for wireless communications, and implicitly assumes a knowledge of the basics of active circuit design and communications theory. In any field advancing as rapidly as this one, there is a danger that a narrowly focused book will become rapidly dated. In order to avoid that, each chapter contains a clear exposition of the fundamental aspects of the topic in addition to more topical material related specifically to wireless communications circuits. It is our hope that the fundamental portions of the book will retain their usefulness well after some of the material has lost its freshness.

In addition, the book emphasizes *monolithic* solutions to microwave/RF circuit design problems. This is because the industry-wide trend is to move to higher and higher levels of integration in order to meet the aggressive cost targets associated with the commercial/consumer market. It is our belief that even if the designer is not actually designing the monolithic circuits, the design of RF systems *using* monolithic circuits will be greatly enhanced by an understanding of monolithic circuit techniques. However, our intent is not to exclude hybrid design techniques, and they are also

addressed where appropriate, especially in the chapters on mixer and amplifier design.

Chapter 2 of the book presents an overview of wireless communications standards and the associated RF system design considerations associated with their implementation. These new standards cover voice and data transmission over a variety of frequency bands, in the United States as well as Europe and Asia. The chapter makes the point that traditional RF design approaches often fall short of meeting the requirements of modern digital wireless communications products, and presents some illustrative examples where different design methodologies are required to meet these new requirements.

Chapter 3 of this book presents an overview of the general field of wireless RF IC design, with particular emphasis on the historical development of radio receiver architectures, receiver design, and low-power techniques. This chapter develops the argument that modern communications system designs will rely on direct downconversion and other novel architectural approaches to realize the low-power/low-cost requirements of handheld consumer communications equipment.

Chapter 4 presents a detailed overview of digital RF modulation techniques and their circuit implementations. This material is normally covered in digital communications textbooks. Here, we have developed a one-chapter summary of the key results of that field, with particular emphasis on those modulation schemes that are currently being implemented in modern communications systems. The chapter also adopts the practical approach of outlining the methods of implementing those modulation techniques with a variety of circuit approaches.

Chapter 5 presents a thorough review of the design and fabrication of RF and microwave mixers. Once again, this field is often covered in an entire text. We have chosen here to cover those areas of mixer design particular to the field of wireless communications, and have emphasized practical circuit approaches to mixer design rather than theoretical depth.

Chapter 6 presents a complete overview of frequency synthesizer designs for wireless communications, with particular emphasis on voltage controlled oscillators (VCOs), phase lock loops (PLLs), and direct digital synthesis (DDS) implementations. The circuits often form the "heart" of the communications system, in the sense that they provide the modulation and demodulation frequencies for tuning the transmitted and received signals. As a result, their performance and cost are crucial to the realization of acceptable system performance.

Finally, Chapter 7 presents a complete overview of the field of RF and microwave amplifier design. Although the design of amplifiers is a relatively mature field, a variety of new circuit architectures have been developed in the last few years that address the particular requirements presented by the wireless communications field. These requirements include an extremely high dynamic range in conjunction with low-power operation as well as high efficiency in power output stages in conjunction with low-voltage operation. In addition, monolithic solutions are now being pursued where a hybrid approach would have been standard only a few years earlier.

References

[1] H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, New York, NY: John Wiley & Sons, 1980.

[2] W. H. Hayward, *Introduction to Radio Frequency Design*, Englewood Cliffs, NJ: Prentice-Hall, 1982.

[3] U. L. Rohde, and T.T.N. Bucher, *Communications Receivers: Principles & Design*, New York, NY: McGraw-Hill, 1988.

[4] G. H. Heilmeier, "Personal Communications: Quo Vadis," in *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1992, pp. 24–26.

[5] Source: The Yankee Group, 1993. From R. Schneiderman, *Wireless Personal Communications—The Future of Talk:* IEEE Press, 1994.

[6] C. Kerramac et al., "New Application Opportunities for SiGe HBTs," *Microwave Journal*, Oct. 1994, pp. 22–37.

[7] N. Colmenares, "The FCC on Personal Wireless," *IEEE Spectrum*, May 1994, pp. 39–46.

[8] J. Lodge, "Mobile Satellite Communications: Toward Global Personal Communications," *IEEE Communications Mag.*, Vol. 29, No. 11, Nov. 1991, pp. 24–30.

[9] T. Freeburg, "Enabling technologies for wireless in-building network communications—four technical challenges, four solutions," *IEEE Communications*, April 1991, pp. 58–64.

[10] J. Lida, "MIS likes unshielded wiring, but worries about support," *MIS Week*, July 1989, p. 31.

[11] D. Bantz, and F. Bauchot, "Wireless LAN Design alternatives," *IEEE Network*, March/April 1994, pp. 43–52.

[12] G. Berline, and E. Perratore, "Portable, affordable, secure: wireless, LANs," *PC Magazine*, Feb. 11, 1992, pp. 291–314.

[13] M. Chelouche, and A. Plattner, "Mobile broadband system (MBS): Trends and impact on 60 GHz band MMIC development," *Electronics and Communications Engineering Journal*, June 1993, pp. 187–197.

[14] R. G. Blake, "The use of millimeterwave for broadband local distribution," *Proc. 17th European Microwave Conf.*, Sept. 1987, pp. 293–298.

[15] L. M. Burns, "Applications for GaAs and Silicon ICs in next generation wireless communications systems," *Proc. of 1994 IEEE GaAs IC Symp.*, pp. 155–158.

[16] J. Neal, "HBT Technology adds power to CDMA chip set," *Microwaves and RF*, Dec. 1994, pp. 205–207.

[17] C. Kermarrec et al., "New application opportunities for SiGe HBTs," *Microwave Journal*, Oct. 1994, pp. 22–36.

[18] M. Matsumoto et al., "A miniaturized dielectric monoblock band-pass filter for 800 MHz band cordless telephone system," *1994 MTT-S Symposium*, San Diego, CA, May 23–27, 1994.

[19] K. Negus et al., "Highly integrated transmitter RFIC with monolithic narrowband tuning for digital cellular handsets," *1994 IEEE Int'l Solid-State Circuits Conf.*, Feb. 1994, pp. 38–39.

[20] T. Stetzler et al., "A 2.7V to 4.5V Single Chip GSM Transceiver RF Integrated Circuit," *1995 IEEE International Solid-State Circuits Conf.*, Feb. 1995, pp. 150–151.

CHAPTER 2

INTRODUCTION TO WIRELESS COMMUNICATIONS APPLICATIONS AND CIRCUIT DESIGN

Jonathon Y. C. Cheah
Solectek Corporation
San Diego, CA

2.1 INTRODUCTION

Stimulated by the increased consumer and commercial users demand for wireless communications applications, the advancement of sophistication in wireless circuit design has progressed at unprecedented speed in recent years. The last ten years have witnessed a remarkable miniaturization of portable equipment, a similar extension of battery life, and almost an order of magnitude reduction in retail prices. One ubiquitous example of this phenomenon is the North American AMPS handheld cellular telephone. Other examples include handheld portable telephones, paging devices, and wireless local area networks (LANs).

Although not well known among everyday wireless communication consumers, *commercial* wireless communications technology has also experienced a similar impetus in its advancement, both in miniaturization and cost. Commercial very small aperture (satellite ground) terminal (VSAT) is a good example. Before the advent of VSAT technology some ten years ago, the cost of satellite ground terminals for

commercial data, voice, and video circuits was at least a few hundred thousand dollars per station, which were customarily housed in specially dedicated buildings. Today, it is not uncommon to see VSATs mounted unobtrusively on top of gas station kiosks and other commercial buildings, at a cost of less than ten thousand dollars per station.

In the portable wireless communication industry, it is well known that the design of the RF transceiver is usually the key element that determines the cost, the size, and the useful battery life of the equipment, as well as how the equipment is used. In *digital* wireless communications, the requirements imposed on the RF transceiver design demand RF engineers to understand and be conversant in digital communications techniques.

For example, in digital communications systems, oscillator phase noise is often specified not only in the traditional dBc/Hz at a fixed-frequency offset profile, but also in noise variance at a determined probability density function. This specification allows the designer to optimize the digital modulation performance while simultaneously optimizing the design in terms of cost and size of the device. Likewise, filters are required to meet phase distortion and group delay requirements, in addition to a prescribed gain versus frequency response.

This chapter will examine some of the tradeoffs involved in the design of the electronic circuits for wireless communications, and how those design parameters can affect ultimate system performance.

2.2 CURRENT AND NASCENT CONSUMER WIRELESS COMMUNICATIONS STANDARDS AND APPLICATIONS

One of the major forces that sustains the continual research and development in the wireless communication arena is the general user acceptance of wireless communications standards. The standards provide a coalescence within the industry to invest in the creation of suitable chip sets for the manufacturing of consumer subscriber units. The resulting high volume drives down the cost of the final product, and the industry is able to take advantage of the same trends that have lowered prices so dramatically in the PC industry. It also encourages investment in the development of communication infrastructure, such as base stations, switch offices and underground copper or optical fiber cables.

It is interesting to note that competing standards create a divisive effect in the market place. Unlike the market for consumer products, where competitive products have had a positive end effect of improved product quality and reduced cost, competing *standards* in the same market confuse the users. The result is a slow market development. In this respect, the rivalry between IS-54 [1] and IS-95 [2] serves as a good example.

The existing major analog cellular telephone standards in the world are listed in Tables 2.1 and 2.2 [3].

Table 2.1
Major Analog Worldwide Cellular Telephone Standards

Mobile Standards	*AMPS*	*TACS*	*ETACS*	*NMT 450*	*NMT 900*
Frequency plan (MHz)	824–849 869–894	890–915 935–960	872–905 917–950	453–457.5 463–467.5	890–915 935–960
Channel spacing (kHz)	30	25	25	25	12.5
Number of channels	832	1,000	1,240	180	1,999
Region	America Australia SE Asia Africa	Europe Africa SE Asia	United Kingdom	Europe Africa SE Asia	Europe Africa SE Asia

Table 2.2
Additional Analog Worldwide Cellular Telephone Standards

Mobile Standards	*C-450*	*RTMS*	*Radiocom 2000*	*NTT*	*JTACS* *NTACS*
Frequency plan (MHz)	450–455.74 460–465.74	450–455 460–465	192.5–199.5 200.5–207.5 215.5–233.5 207.5–215.5 165.2–168.4 169.8–173 414.8–418 424.8–428	925–940 870–885 915–918.5 860–863.5 922–925 867–870	915–925 860–870 898–901 843–846 918.5–922 863.5–867
Channel spacing (kHz)	10	25	12.5 12.5 12.5	25/6.25 6.25 6.25	25/12.5 25/12.5 12.5
Number of channels	573	200	560 640 256	600/2,400 560 480	400/800 120/240 280
Region	Germany Portugal	Italy	France	Japan	Japan

However, the current proposed or newly adopted wireless communications standards for personal communications services (PCS) are predominantly based on digital formats. This is because of the proven superiority digital techniques can provide. In contrast, much of the existing wireless RF circuit design techniques were founded upon analog voice systems, such as the AM and FM push-to-talk land or maritime mobile SMR telephones, the omnipresent cordless and cellular telephones, CB radios, and the SSB airline air-to-ground pay phone, to name but a few.

The modern digital standards for handheld and low-power wireless voice communications PCS are shown in Table 2.3 [4].

Table 2.4 shows the equivalent digital voice cellular telephone standards for various existing and proposed systems. These standards also support digital data transmission. Table 2.5 shows a summary of the wireless data communications standards. These standards are used for the transmission of digital data, rather than digitized voice. SMR services currently in operation also provide data communication services.

Apart from the data services indicated in Table 2.5, there are also wireless LAN standards that are being prepared, such as IEEE 802.11 [3–5] in the United States and ETSI/RES10 (otherwise known as HIPERLAN) in Europe. HIPERLAN has ERC allocation of 5.150 to 5.300 GHz and 17.000 to 17.300 GHz for its use. The IEEE 802.11 proposes to support 1 to 2 Mbps channel signaling rate, whereas HIPERLAN targets 20 Mbps.

The IEEE 802.11 is primarily intended to utilize the ISM bands of 902 to 928 MHz, 2.400 to 2.4835 and 5.725 to 5.850 GHz, which are facilitated by the FCC Part 15.247 rule [6] that allows peak transmit radiating power of up to 0 dBw using spread-spectrum techniques. Under this rule, the rigorous regulatory certification process of the intentional radiators is not in force as it is under other FCC regulatory classifications. There are a variety of wireless data products that are already in the U.S. market that take advantage of the FCC Part 15.247 rule. The most common consumer examples are the extended range cordless telephone and cordless stereo speakers using the 902- to 928-MHz band. Unfortunately, there

Table 2.3
Condensed Version of Wireless Communications Technology Standards for Low-Power Voice Systems

Mobile System	*DECT*	*PHS*	*CT-2*	*CT-3*	*UDPC*
Multiple access method	TDMA/FDMA	TDMA/FDMA	FDMA	TDMA/FDMA	TDM/TDMA/FDMA
Frequency bands (MHz)	1,800–1,900	1,895–1,907	864–868	862–866	Within 1.7–2.3 GHz
Duplex	TDD	TDD	TDD	TDD	FDD
Channel spacing	1,728 kHz	300 kHz	100 kHz	1,000 kHz	350 kHz
Modulation	FSK	$\pi/4$ QPSK	FSK	FSK	$\pi/4$ QPSK
TX max power	250 mW	80 mW	10 mW	80 mW	200 mW
Speech coding	ADPCM	ADPCM	ADPCM	ADPCM	ADPCM
Bit rate	32 Kbps	32 Kbps	32 Kbps	32 Kbps	32 Kbps
Speech per channel	12	4	1	8	10
Channel bit rate	96 Kbps	72 Kbps	640 Kbps	500 Kbps	
1152 Kbps					
Channel coding	CRC	CRC	None	CRC	CRC

Table 2.4
Summary of Digital Voice Cellular Telephone Standards

Mobile System	*IS-54*	*GSM*	*JDC*	*IS-95*
Multiple access method	TDMA/FDMA	TDMA/FDMA	TDMA/FDMA	FDMA/CDMA
Frequency bands	824–848 869–894	890–915 935–960	810–826 940–956 1,447–1,489 1,429–1,441 1,501–1,513 1,453–1,465	824–848 862–866
Duplex	FDD	FDD	FDD	FDD
Channel spacing	30 kHz	200 kHz	25 kHz	1,250 kHz
Modulation	$\pi/4$ QPSK	GMSK BT = 0.3	$\pi/4$ QPSK	QPSK
TX max power	600 mW	1W		200 mW
Speech coding	VSELP	RPE-LTP	VSELP	QCELP
Bit Rate	8 Kbps	13 Kbps	8 Kbps	1–8 Kbps
Speech per channel	3	8	3	12
Channel bit rate	48.6 Kbps 270.833 Kbps	42 Kbps	1,228.8 Kbps	
Channel Coding	Rate-1/2 convolution	Rate-1/2 convolution	Rate-1/2,1/3 convolution	CRC

Table 2.5
Summary of Mobile Wide Area Wireless Data Services

Mobile System	*ARDIS*	*MOBITEX*	*CDPD*	*IS-95*	*TETRA*
Multiple access method	FDMA/DSMA	FDMA/SLOTTED-ALOHA	FDMA/DSMA	FDMA/CDMA	FDMA/DSMA
Frequency bands (MHz)	800 band	935–940	824–849 869–894	824–849 869–894	400 band 900 band
Channel spacing	25 kHz	12.5 kHz	30 kHz	1,250 kHz	25 kHz
Modulation	4-FSK	GMSK BT = 0.3	GMSK BT = 0.5	QPSK/SS	$\pi/4$ QPSK

is also a plan for the U.S. regulatory agencies to reclaim portions of these frequency bands.

Another wireless communication family is known as specialized mobile radios (SMR). SMRs are used in various applications using often noncontiguous privately licensed frequency bands. As a result, the radio transceiver design and the performance requirements are often more stringent than those of standard consumer prod-

ucts. It is particularly true that the regulatory rules that govern SMR operations are based on traditional AM and FM push-to-talk land mobile radios with respect to the transmit spectrum mask, and thus impose a very tight constraint on the digital communication implementations. The general SMR performance and applications are governed under FCC Part 90 [7] and industrial recommendations can be found in TIA/EIA 603 [8].

There is also an expanded emphasis on wireless telephone usage for replacement of landline telephony local loop infrastructure. This application of wireless voice communications in these applications is commonly known as wireless local loop (WLL). Wireless local loop replaces the underground telephone cable drop to each user and connects the user directly to the public telephone switch network through a cellsite by means of a radio. The advantage of WLL is the deployment of a telephone network virtually overnight in areas where telephone services are not available, such as some cities in third-world countries or rural communities in the western world.

2.3 RADIO TRANSCEIVER TECHNOLOGY REQUIREMENTS

In radio transceiver design, it is often advantageous to have an *a priori* idea of the performance of the state-of-the-art radio transceiver component technology before it is integrated. Mature component data, though not always readily available, allows the systems designers as well as the circuit designers to assess how a required specification can be met in terms of its cost and technical risk. Systems tradeoffs are often dominated by the availability and performance of the concurrent new components in the wireless digital communication industry.

Figure 2.1 shows a typical IS-19 cellular telephone frequency plan and block diagram. Similarly, Figure 2.2 shows that of an IS-55 [9] digital dual-mode cellular telephone. It is interesting to see from Figures 2.1 and 2.2 that, although conceptually both radio transceivers have a similar structure, the actual implementation of these radios is significantly different in the frequency plan and circuit design topology. This is because of various design considerations, such as those related to the number of oscillators needed, the potential spurious frequency mixing products, the availability of mature components, and, most importantly, the unit costs. The resulting design is ultimately driven by the available radio transceiver components technology.

Table 2.6 shows an interesting example of practical product performance limits for an IS-55 dual-mode type cellular telephone using current radio transceiver component technology. The specification items selected are dictated by the limits of state-of-the-art component technology in particular. The limits depicted in Table 2.6 are derived from designs that used the most economical implementation and design techniques viable for commercial production. This table illustrates a baseline for practical achievable performance in radio transceiver design for the consumer market. It can also be used to estimate the performance requirements for other

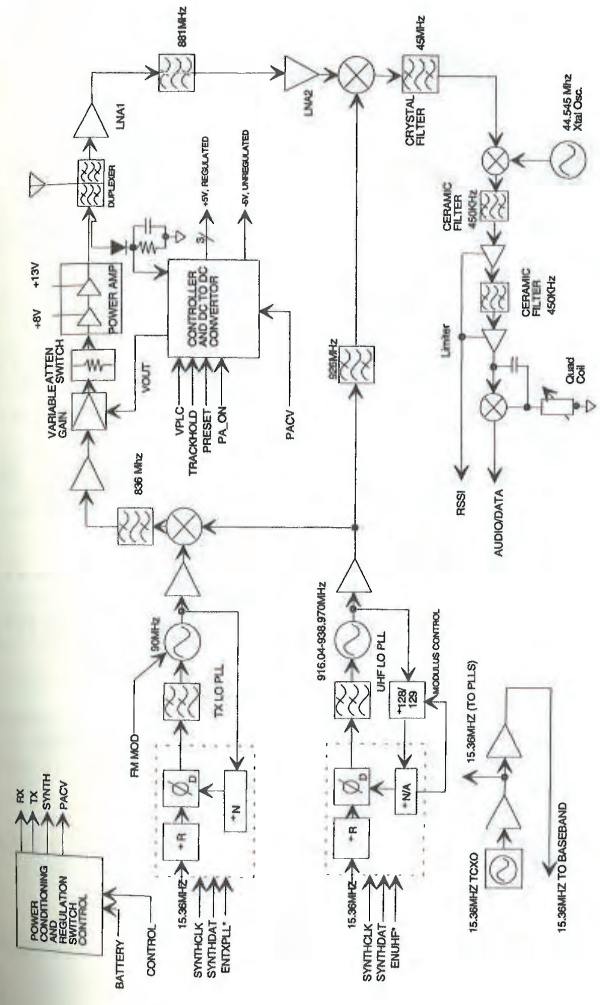


Figure 2.1 IS-19 cellular telephone RF section block diagram. A 45-MHz offset frequency oscillator generates the required receiver and transmitter local oscillator frequencies.

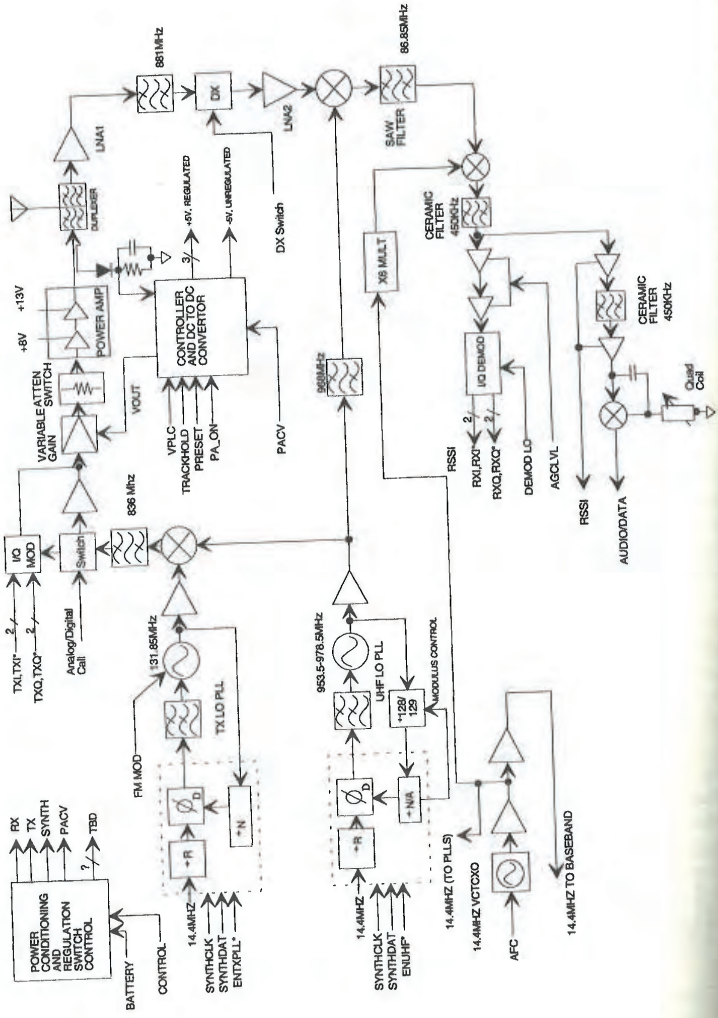


Figure 2.2 IS-55 block diagram. A narrowband IF filter is required for digital operation, as well as an analog-to-digital converter in the baseband.

frequency bands. One can probably assume that if the intended product has the component volume of the cellular market, then the performance and cost of the devices are proportional to the absolute frequency.

Table 2.6
Typical Performance of IS-55 Class of Product Compared with Specifications

IS-55A Specifications	*IS-55A Spec*	*Measured Values*
2. Receiver minimum standards:		
2.1.2.1 Sync. acquisition time (digital mode)		
–90 dBm, 8 & 100 km/h	< 130 msec	< 110 msec
–103 dBm, 8 & 100 km/h	< 250 msec	< 130 msec
2.2.2.2 Audio muting		
Audio level	> 40 dB	> 70 dB
Transient	< 25%; < 25 msec	~ 0%; 0 msec
2.2.2.3 Expander		
Above 0 dB	±1 dB	< ±2 dB
Below 0 dB	±2 dB	< ± 1 dB
Attack time	3 ± 0.6 msec	< 3 msec
Recovery time	13.5 ± 2.7 msec	< 4 msec
2.2.2.4 Hum and noise (dB)	> 32 dB	> 44 dB
2.2.2.5 Audio harmonic distortion (%)	5 %	1.30%
2.2.3.2 Supervisory audio tone decoding		
Detection time	< 250 msec	< 200 msec
Removal time	< 259 msec	< 200 msec
2.3 Receiver performance		
2.3.1 Analog		
2.3.1.1 RF Sensitivity (dBm)	–116 dBm	–121 dBm
2.3.1.2 Adjacent channel selectivity	> 16 dB	> 37 dB
Alternate channel selectivity	> 60 dB	> 71 dB
2.3.1.3 Intermod. spurious resp. atten., near	> 65 dB	> 72 dB
intermod, spurious resp atten., Far	> 73 dB	> 76 dB
2.3.1.4 Prot. against spurious resp. interference	> 60 dB	> 82 dB
2.3.2 Digital		
2.3.2.1 RF sensitivity		
Static (3% BER)	–110 dBm	–116 dBm
8 km/H (3% BER)	–103 dBm	< –103 dBm
100 km/H (3% BER)	–103 dBm	< – 103 dBm
2.3.2.2 Adj. and alt. channel desensitization @ –107 dBm		
30 kHz (+) adj. interferer @ –94 dBm	< 3 % BER	< 0.01 %
60 kHz (+) alt. interferer @ –65 dBm	< 3 % BER	< 0.03 %
30 kHz (–) adj. interferer @ –94 dBm	< 3 % BER	< 0.01 %
60 kHz (–) alt. interferer @ –65 dBm	< 3 % BER	< 0.01 %
2.3.2.3 Intermod spurious resp. atten. @ –107 dBm		
120/240 kHz (+) @ –45/–45 dBm	< 3 % BER	
120/240 kHz (–) @ –45/–45 dBm	< 3 % BER	0.3%

Table 2.6 (continued)

IS-55A Specifications	*IS-55A Spec*	*Measured Values*
2.3.2.6 Cochannel performance		
FACCH, 8 km/H 14 dB CIR	< 7.3 % WER	0.2 %
FACCH, 100 km/H 12 dB CIR	< 3.3 % WER	0.2 %
SACCH, 8 km/H 14 CIR	< 14 % WER	0.2 %
SACCH, 100 km/H 12 dB CIR	< 9.0 % WER	0.2 %
Class 1, 8 km/H, 17 dB CIR	< 7.1 % WER	0.2 %
Class 1, 100 km/H, 17 dB CIR	< 1.4 % WER	0.2 %
Data Field, 8 km/H, 17 dB CIR	< 3 % BER	0.9 %
Data Field, 50 km/H, 17 dB CIR	< 3 % BER	0.9 %
Data Field, 100 km/H, 17 dB CIR	< 3 % BER	0.9 %
2.3.2.7 Delay interval		
BER Under carrier to noise		
8 km/H, 10.3 usec delay, –103 dBm	< 3 % BER	0.7 %
8 km/H, 20.6 usec delay, –103 dBm	< 3 % BER	0.3 %
8 km/H, 41.2 usec delay, –103 dBm	< 3 % BER	0.07 %
50 km/H, 10.3 usec delay, –103 dBm	< 3 % BER	1 %
50 km/H, 20.6 usec delay, –103 dBm	< 3 % BER	0.5 %
50 km/H, 41.2 usec delay, –103 dBm	< 3 % BER	0.1 %
100 km/H, 10.3 usec delay, –103 dBm	< 3 % BER	1 %
100 km/H, 20.6 usec delay, –103 dBm	< 3 % BER	0.5 %
100 km/H, 41.2 Usec Delay, –103 dBm	< 3 % BER	0.1 %
BER under cochannel interference		
8 km/H, 17 dB Cir, –85 dBm	< 3 % BER	0.4 %
50 km/H, 19 dB Cir, –85 dBm	< 3 % BER	0.5 %
100 km/H, 22 dB Cir, –85 dBm	< 3 % BER	0.5 %
8 km/H, 16 dB Cir, –85 dBm	< 5 % BER	0.5 %
50 km/H, 17 dB Cir, –85 dBm	< 5 % BER	0.3 %
100 km/H, 17 dB Cir, –85 dBm	< 5 % BER	0.8 %
2.4 Conducted spurious emission (dBm)		
No spurious output at antenna terminal out of band	> –47 dBm	> –74 dBm
No spurious output at antenna terminal in RX band	> –80 dBm	> –80 dBm
No spurious output at antenna terminal in Tx band	> –60 dBm	> –60 dBm
2.5 Radiated spurious emission (dBm)		
470 MHz To 1,000 MHz	< –21 dBm	< –69 dBm
3 Transmitter minimum specifications		
3.1 Frequency requirements		
3.1.1 Analog		
3.1.1.2 Frequency stability	± 2.5 ppm	± 1.6 ppm
3.1.1.3 Carrier switching time (usec)		
Attack time	< 2 msec	6 usec
Release time	< 2 msec	4 usec
3.1.1.4 Channel switching time (msec)	< 40 msec	3.5 msec
3.1.2 Digital mode		
3.1.2.1 Reserved		
3.1.2.2 Frequency stability	± 200 Hz	± 10 Hz

Table 2.6 (continued)

IS-55A Specifications	*IS-55A Spec*	*Measured Values*
3.2 RF power output requirements		
3.2.1 RF power output		
3.2.1.1 Analog RF power output Pl0-Pl7	+2/–4 dBm	+1/–2 dBm
3.2.1.2 Digital RF power output Pl0-Pl7	+2/–4 dBm	+1/–2 dBm
3.2.2 RF Power transition time		
3.2.2.1 Analog (msec)		
Pl0-Pl7	< 20 msec	200 usec
Pl7-Pl0	< 20 msec	800 usec
3.2.2.2 Carrier-on state		
Inactive	< –60 dBm	< –65 dBm
3.3 Modulation requirements		
3.3.1 Analog		
3.3.1.1 Modulation type And stability (kHz)		
3.3.1.2 Voice modulation		
3.3.1.2.1 Compressor		
Above 0 dB	± 0.5 dB	± 0.1 dB
Below 0 dB	± 1 dB	± 0.2 dB
Attack time	3 ± 0.6 msec	3.58 msec
Recovery time	13.5 ± 2.7 msec	9.98 msec
3.3.1.2.3 Modulation deviation limiting	< 12 kHz	< 12 kHz
3.3.1.2.4 Audio voice path muting	> 40 dB	> 42 dB
3.3.1.3 Wideband data (kHz)	± 10%	± 1.4 %
3.3.1.4 Supervisory audio tone	± 10 %	± 1 %
3.3.1.5 Signaling tone	± 10 %	± 0.5 %
3.3.1.6 FM hum and noise	< –32 dB	–41 dB
3.3.1.7 Residual amplitude modulation (AM)	5 %	0.5 %
3.3.2 Digital modulation		
3.3.2.1 Modulation type and accuracy		
Error vector magnitude	12.5 %	8.5 %
Ten burst error	25 %	9 %
Origin offset	< –20 dBc	–30 dBc
3.4 Limitations on emissions		
3.4.1 Spectrum noise suppression-broadband		
3.4.1.1 Analog		
Voice plus sat, 20 kHz to 45 kHz from carrier	< 8.7 dBm	–4 dBm
Voice plus sat, > 45 kHz from carrier	< –33 dBm	–45 dBm
Voice plus sat, RX band	< –80 dBm	–90 dBm

2.4 RF COMPONENT REQUIREMENTS FOR TRANSCEIVERS

Radio transceiver circuit designers are often faced with the need to familiarize themselves with the new components that are continually being made available for a particular product family. In particular, the RF functions of a typical transceiver

are now being integrated at increasingly high levels onto a single semiconductor die, much the same way that digital IC technology is integrating more and more logic gates on a typical chip. This technological trend is reducing the cost of transceivers even further.

Examination of the performance results shown in Table 2.6 highlights some of the component areas limiting system performance. These include the need for:

1. Duplexers with low insertion loss;
2. Low-noise amplifiers (LNAs) with high intercept point and low current consumption;
3. Mixers with high intercept point and low noise;
4. Filters with low thermal center frequency drift and frequency response tolerance;
5. Low-noise PLLs with fast switching time.

The design considerations of each of these components will be considered more deeply in the following chapters of this book. In particular, the design of mixers is considered in Chapter 5, frequency synthesizers in Chapter 6, and amplifiers in Chapter 7.

2.4.1 Duplexer Requirements

Although the digital mode of the IS-55 specification is indeed half-duplex, and thus the need for a duplexer may be eliminated, the analog mode requirement of IS-55 (and virtually all digital cellular phones) prevents this removal from occurring. Furthermore, extensive GSM experience has indicated that the isolation provided by the duplexer, or alternatively a front-end receive filter, is necessary when a phone is operating in close proximity to one or more other phones. Usually, a duplexer adds an additional 3-dB noise figure in the receiver and 3-dB power loss in the transmit paths at 800 MHz.

In a half-duplex application, the front-end transmit filter requirements are greatly reduced. The > 120-dB transmit-to-receiver noise floor isolation is not required. It follows from this relaxed requirement that critical circuit board layout considerations are correspondingly less stringent. Commercial duplexers available for cellular telephone industry use have the largest physical volume of all other RF components in the transceiver. In handheld portable equipment, the duplexer dimensions and its losses can be a real problem in the enclosure design and the battery life considerations.

2.4.2 Amplifier Requirements

Perhaps one of the hardest specifications a designer of an RF transceiver faces is the balancing act between receiver sensitivity and the intermodulation requirements in conjunction with full transmit power. Current low-noise amplifier components, either in terms of discrete low-noise transistors or low-noise amplifier blocks, have very little margin for tradeoff between LNA gain, receiver sensitivity (or overall noise figure) at 3-dB CNR (carrier-to-noise ratio), and 73-dB intermodulation protection. These limitations are specified by IS-19/TIA-EIA-553 [10, 11] analog cellular telephone specifications and the 3% BER sensitivity at –110-dBm and 52-dB intermodulation limits of the IS-55 dual-mode cellular telephones. At the same time, mismanagement of the transmitter noise floor could certainly doom the overall design.

2.4.3 Mixer Requirements

It is commonly recognized by experienced designers that the selection of the first mixer design topology dominates the receiver chain performance. The mixer performance consideration is accentuated by the need to provide good local oscillator (LO) leakage isolation while maintaining a good balance between the front-end noise figure (NF) and intermodulation isolation. One common design approach replaces the front-end LNA with an active low-noise first mixer. In mobile RF transceiver design, where power drain is not a problem, there are more degrees of freedom in the mixer topology selection, given the fact that the LNA can be biased at higher current and voltage levels to obtain good front-end intercept point performance. In a handheld portable unit, such a luxury is out of the question due to the battery drain budget.

2.4.4 Filter Requirements

Filter designs are the most mature components used among the five items listed above. Filter designs in consumer wireless applications come in a few different forms, dependent on frequency. In the cellular industry, for instance, SAW, dielectric, ceramic, and crystal filters are the most common nonlumped element filters used. The technology for fabricating these elements is quite mature, and they tend to be very inexpensive. However, in modern digital communications, the accuracy of the filter frequency characteristics is frequently a prolific source of product failures due to the more stringent requirements placed on their performance.

The most common cause of severe performance degradation that can be attributable to filter performance in digital communication systems is the passband frequency drift with respect to center frequency of the filter. The passband frequency inaccuracy is usually due to the component production tolerance and thermal characteristics. Most of the digital modulation techniques are extremely sensitive to band-limiting phenomena. In addition, passband drifting also impairs the group delay budget for the modulation. In many cases, active filter equalizations are implemented

to counteract this problem. The real tradeoff is in the cost of the design. One can use a broader band filter and satisfy the inadequate adjacent channel rejection requirements by using a tight signal-matched filter with a high signal sample resolution.

In contrast, amplitude distortion seen in the mass-volume-produced commercial filters available today is generally not a serious limiting problem.

2.4.5 Frequency Synthesizer Requirements

The art of PLL design has improved tremendously in the last 10 years. The corresponding chip sets for various PLL structures are now commonplace. Even more esoteric PLL structures such as the fractional-n PLLs, which were not generally in use in the past, can now be seen in some designs. This is because the current challenge in PLL design is to achieve a fast switching time for a small-frequency step size without resorting to multiple PLL designs or using direct digital synthesis (DDS) techniques. Both multiple PLL and DDS options bear the penalty of cost, increased power drain, more printed wiring board (PWB) real estate, and increased EMC problems. Today, existing PLL chip sets are meeting the data communication's demand for switching time only by small margins.

2.5 SYSTEMATIC ANALYSIS OF TRANSCEIVER DESIGN

After mustering all the suitable components for an RF design, the overall transceiver system analysis can be estimated. The cascaded noise figure for an n-stage system is given by [12]

$$\mathrm{NF} = \mathrm{NF}_1 + \frac{\mathrm{NF}_2 - 1}{G_1} + \frac{\mathrm{NF}_3 - 1}{G_1 G_2} + \ldots + \frac{\mathrm{NF}_n - 1}{G_1 G_2 \ldots G_{n-1}} \tag{2.1}$$

where NF_1 is the linear noise figure of stage i and G_i is the linear gain of the stage i.

The second-order intercept point (IP2) and the third-order intercept point (IP3) as referenced from the output are given by

$$\mathrm{IP2} = -20\log\left[\sum_i \sqrt{\frac{1}{\mathrm{IP2}_i G_{i+1} G_{i+2} \ldots G_n}}\right] \qquad i = 1 \ldots n, \tag{2.2}$$

$$\mathrm{IP3} = -10\log\left[\sum_i \frac{1}{\mathrm{IP3}_i G_{i+1} G_{i+2} \ldots G_n}\right] \qquad i = 1 \ldots n, \tag{2.3}$$

Using (2.1–2.3), an analysis of the receiver design in Figures 2.2 and 2.3 was carried out, with the results shown in Tables 2.7 through 2.10.

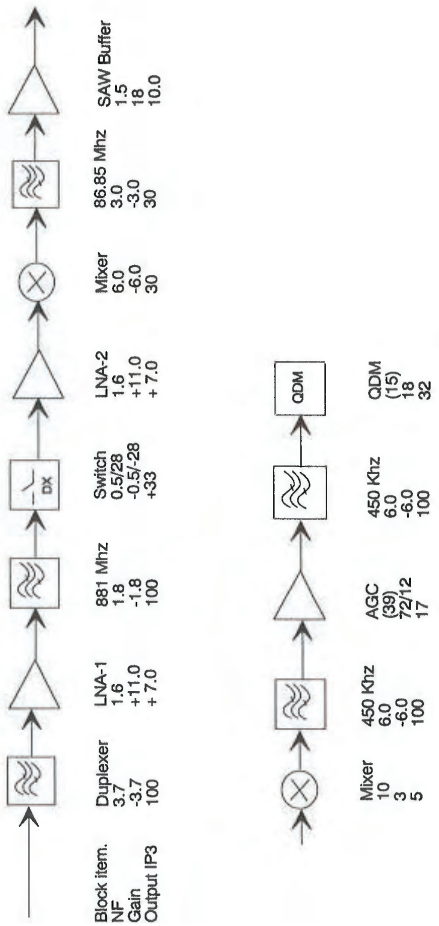


Figure 2.3 Receiver characteristics used for analysis.

Table 2.7
Small-Signal Analysis Results from Figures 2.2 and 2.3

Receiver Blocks	NF(dB)	Gain (dB)	Output IP3 (dBm)	Cascade NF (dB)	Cascade Gain (dB)	Cascaded IP3 (dBm)	Signal Strength (dBm)
Input							−110.0
Duplexer	3.7	−3.7	100.0	3.7	−3.7	98.46	−113.7
LNA-1	1.6	11.0	7.0	5.3	7.3	7.0	−102.7
881-MHz filter	1.8	−1.8	100.0	5.42	5.5	5.2	−104.5
DX switch	0.5	−0.5	33.0	5.46	5.0	4.69	−105.0
LNA-2	1.6	11.0	7.0	5.63	16.0	6.45	−94.0
UHF mixer	7.0	−6.0	30.0	5.75	10.0	0.44	−100.0
86-MHz filter	3.0	−3.0	30.0	5.87	7.0	−2.56	−103.0
SAW buffer	1.5	15	10.0	5.96	22.0		
VHF mixer	10.0	23.0	5.0	6.02	45.0		
450-kHz filter	6.0	−6.0	100	6.02	39.0	Narrowband Signal	
AGC amp.	(39)	68.0	17.0	6.99	111.0		
450-kHz filter	6.0	−6.0	100.0	6.99	105.0		
QDM	(20)	18.0	32.0	6.99	123.0		

Table 2.8
Summary of Small-Signal Results from Figures 2.2 and 2.3

Performance Parameter	Result
KTB (BW = 24.3 kHz)	−130.14 dBm
NF	6.99 dB
System noise floor	123.15 dBm
Minimum signal E_s/N_0 (For 4-phase differential detection @ BER = 10^{-3})	12 dB
Implementation loss	1 dB
Required CNR in AWGN	13 dB
Sensitivity @ BER = 1e − 3	−110.15 dBm
Intended minimum signal	−110 dBm
Input IP3 EISA IM method IS-55 -2.3.2.3	−9.56 dBm > 60 dB
IP3 /signal strength	> 20 dB minimum.

Table 2.9
Large-Signal Analysis Results from Figures 2.2 and 2.3

Receiver Blocks	NF(dB)	Gain (dB)	Output IP3 (dBm)	Cascade NF (dB)	Cascade Gain (dB)	Cascaded IP3 (dBm)	Signal Strength (dBm)
Input							−25
Duplexer	3.7	−3.7	100.0	3.7	−3.7	98.46	−28.7
LNA-1	1.6	11.0	7.0	5.3	7.3	7.0	−20.7
860-MHz filter	1.8	−1.8	100.0	5.42	5.5	5.2	−22.5
DX switch	28	−28	33.0	22.58	−22.5	−22.8	−50.5
LNA-2	1.6	11.0	7.0	24.15	−11.5	−1186	−39.5
UHF mixer	7.0	−6.0	30.0	25.01	−17.5	−17.86	−45.5
86-MHz filter	3.0	−3.0	30.0	25.72	−20.5	−20.86	−48.5
SAW buffer	1.5	15	10.0	26.22	−5.5		
VHF Mixer	10.0	3.0	5.0	26.54	−2.5		
450-kHz filter	6.0	−6.0	100	26.59	−8.5	Narrowband Signal	
AGC amp.	(39)	12.0	17.0	47.53	3.5		
450-kHz filter	6.0	−6.0	100.0	47.53	−2.5		
QDM	(20)	18.0	32.0	47.55			

Table 2.10
Summary of Large-Signal Results from Figures 2.2 and 2.3

Performance Parameter	Results
KTB (BW = 24.3 kHz)	−130.14 dBm
NF	47.55 dB
System noise floor	−82.59 dBm
Required signal CNR in AWGN	13 dB
Sensitivity @ BER=1e − 3	−69.59 dBm
Intended maximum signal	−25 dBm
IP3/signal strength	> 20 dB minimum

2.6 DIGITAL COMMUNICATIONS REQUIREMENTS

The rapid acceptance of digital communication techniques has created an interesting predicament for practicing RF circuit design engineers. The systems parameters suited to digital communication analyses are generally different from conventional RF circuit design parameters or their measurement parameters. Depending on a particular digital communication application, the digital communication system requirements impose different RF transceiver specifications for the same parameter value. Here, two of the fundamental system parameters that are often faced by RF transceivers at the beginning of a design are briefly discussed. They are the bit energy require-

ment of the system and phase noise specifications. Experience has demonstrated that these two parameters are the most common sources of confusion for RF transceiver designers.

For example, the aggregate receiver phase noise requirement for oscillators is commonly specified as a fixed rms double-side-band noise variance in radians over a signal bandwidth. The signal bandwidth for the phase noise calculation is a function bounded by the product of all the transfer functions of all the filters in the RF transceiver, the signal matched filter, and the highpass transfer function of the automatic frequency control (AFC). This integrated phase noise specification is itself dependent on the required bit error rate (BER) of the system.

The estimation of the required receiver sensitivity is usually based on the minimum E_b/N_0 of the digital communication specification for a set value of BER. Thus,

$$E_b = PT_b \tag{2.4}$$

where, E_b is the energy per bit, P is the power required and T_b is the bit time.

It follows that the theoretical probability of bit error in an additive white Gaussian noise (AWGN) environment with respect to the E_b/N_0 for M'ary PSK can be expressed as

$$\text{Pb} = \frac{1}{2}(1 - \text{erf}(\sqrt{E_b/N_0})) \tag{2.5}$$

In order to estimate the required receiver sensitivity, one needs to know the M'ary nature of the modulation scheme; that is, how many bits per hertz in the signal. For BPSK, the required E_b/N_0 for a 10^{-3} BER in AWGN environment from (2.5) is 6.8 dB. For QPSK, where there are two bits per symbol, the same energy E_b/N_0 value holds for the same BER, but the required symbol power spectral density within the signal bandwidth can then calculated from (2.4) with $T_b = 1/2T$, so the power density required per symbol is in fact twice that of BPSK.

Thus, for example, to achieve BER of 10^{-3} for QPSK symbol rate of 24.3 Ksps, the achievable sensitivity of the receiver using coherent detection can be calculated as

$$\begin{aligned}\text{Estimated Receiver Sensitivity} = {} & \text{Receiver Noise Floor} \\ & + \text{Required SNR} + \text{Implementation Loss}\end{aligned} \tag{2.6}$$

where

$$\text{Receiver Noise Floor} = \text{KT} + 10\log(\text{Bandwidth Hz}) + \text{NF(dB)} \tag{2.7}$$

and

$$\begin{aligned}\text{Required SNR} = {} & E_s/N_0(\text{dB}) \\ & + 10 \log (\text{Symbols per second}) - 10 \log (\text{Bandwidth Hz})\end{aligned} \tag{2.8}$$

On combining these, the receiver bandwidth cancels, leaving

$$\begin{aligned}\text{Estimated Receiver Sensitivity} = {} & \text{kT} + \text{NF(dB)} + E_s/N_0 \text{ (dB)} \\ & + 10 \log (\text{Symbols per second}) \\ & + \text{Implementation Loss} \\ = {} & -174 + 7 + E_s/N_0(\text{dB}) + 10\log(24.3 \times 10^3) + \text{Implementation Loss} \\ = {} & -123.14 \text{ dBm} + E_s/N_0(\text{dB}) + \text{Implementation Loss}\end{aligned} \tag{2.9}$$

assuming KT = −174 dBm/Hz (at room temperature) and system noise figure = 7 dB.

Now, the E_b/N_0 required = 6.8 dB (from (2.5)), and so the QPSK symbol power required is Es/N_0 = 6.8 +3 = 9.8 dB. If the estimated implementation loss = 1 dB, then the actual estimated receiver sensitivity at a BER of 10^{-3} is = −123.14 + 9.8 + 1 = −112.34 dBm.

An interesting comparison with this calculation can be made with an IS-55 digital mode receiver sensitivity requirement of BER = 3%, or $3 \cdot 10^{-2}$, in an AWGN environment, specified at −110 dBm using differential detection. Given that a four-phase differential detection performance curve is typically very close to 2.3 dB degradation equidistant from the AWGN BER versus E_b/N_0 curve for most of the E_b/N_0 range, excluding the nonreducible error floor; this is an interesting exercise for the RF transceiver designer new to digital communications to work out.

The next parameter often encountered by the RF transceiver designer is the phase noise requirement specifications. In a conventional phase noise measurement, as shown in Figure 2.4, phase noise is typically specified as a spectral density in dBc/Hz at a certain frequency offset. A crucial question to answer is: how much noise variance is introduced into the digital system due to this phase noise?

Ignoring the filter skirt residuals, the noise variance (σ^2) is the total phase noise existing within the single sideband of the phase noise profile shown in Figure 2.4 from the highpass 3-dB knee of AFC bandwidth to the lowpass 3-dB knee of the signal matched filter. A simple graphical integration method of Figure 2.4 using a straight-line approximation has been derived below [13,14]:

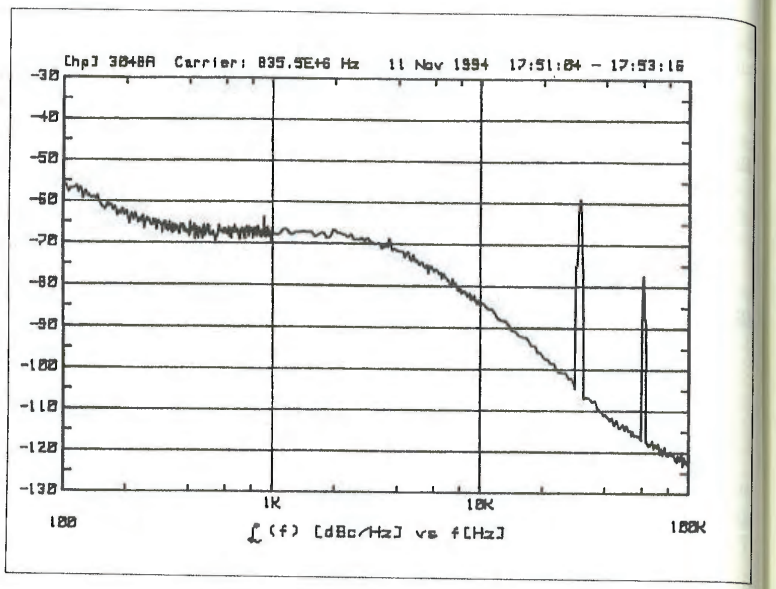


Figure 2.4. Measured phase noise plot of IS-55 transmit carrier. Integration bandwidth is between 200 Hz and 30 kHz.

$$\sigma^2 = \sum_i \left\{ \begin{array}{l} 10^{\frac{dBc_i}{10}} f_i \log\frac{f_{i+1}}{f_i} \rightarrow x = -1, \\ 10^{\frac{dbC_i}{10}} \dfrac{f_i}{x+1}\left[\left[\dfrac{f_{i+1}}{f_i}\right]^{x+1} - 1\right] \rightarrow x \neq -1 \end{array} \right\} \tag{2.10}$$

where

$$x = \frac{dBc_{i+1} - dBc_i}{10(\log f_{i+1} - \log f_i)}$$

With the value of σ^2 obtained above, one can now estimate the degradation in digital performance by calculating the probability of the bit errors. For BPSK, let $\rho = \sqrt{2\sigma}$. Then, the probability of error can be written as [13–15]

$$P_e = \frac{1}{\sqrt{2\pi\rho}} \int_{-\infty}^{\infty} \exp\left(-\frac{\phi^2}{2\rho^2}\right) \operatorname{erfc}\left(\sqrt{\frac{2E_b}{N_0}}\cos\phi\right) d\phi \tag{2.11}$$

Other modulation schemes require similar derivations to establish the impact of phase noise on their performance [16].

Using (2.10), the integrated signal sideband phase noise can be directly computed by a graphical integration method. Figure 2.5 shows the accumulation of the sectional integration in terms of σ in degrees rms of the phase noise profile depicted in Figure 2.4.

Using Equation (2.11), it can be shown that the phase noise contribution to the BER performance, as indicated by $\rho = \sqrt{2\sigma}$ for $\pi/4$ QPSK modulation, is not severe even if the phase noise performance is relatively poor. For example, the profile of Figure 2.5 is better than most digital equipment specifications—it shows a double sideband phase noise of 1.98 deg rms. In general, digital communication equipment specifications are overly restrictive in terms of phase noise. The conventional approach has been to use the empirical measurement of a typical system combination of oscillators—in terms of dBc/Hz related to an offset frequency from the carrier—as a profile for the required specifications. Often, the attempt to satisfy the phase

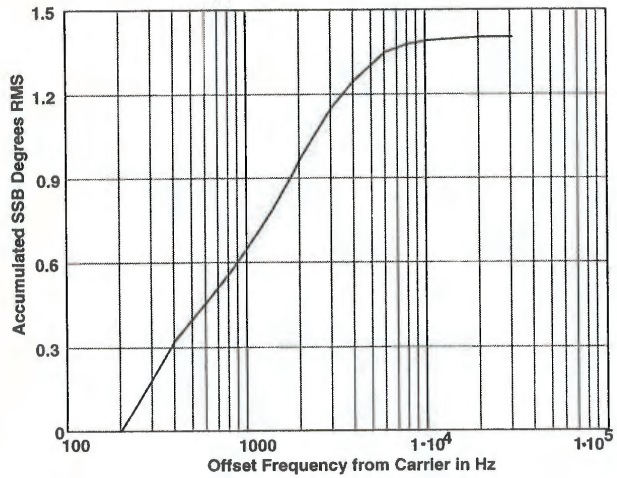


Figure 2.5 Accumulated single sideband phase noise in degrees rms of the phase noise profile shown in Figure 2.4, based on the graphical integration technique described in (2.10).

noise profile resulted in compromises at the expense of channel-switching time, frequency-switching transient response, and, more importantly, per unit cost. These requirements can be relaxed somewhat with an improved understanding of the effect of phase noise on digital system performance.

The introduction of the modulation accuracy [9] concept leads to a more realistic performance specification technique, which encompasses the phase noise contribution as one of many performance elements for the transceiver design. With this technique, the modulated signal is measured against a set percentage of rms signal vector error for its signal constellation. The imperfection in the modulated signal attributable to phase noise, I and Q imbalance, filter imperfections, and circuit noise floor can be determined by a single modulation error vector specification. This approach allows the circuit designer flexibility in circuit design tradeoffs and the communication system designers know explicitly the system performance that the transceiver can deliver.

2.7 ELECTROMAGNETIC COMPATIBILITY IN WIRELESS DEVICES

In RF design processes, one of the most common mistakes committed by RF designers is the omission of up-front electromagnetic compatibility (EMC) considerations. As a result, typical transceiver designs will validate the 20-80 rule; that is, it is common that the last 20% of the RF design takes 80% of the schedule.

Furthermore, consumer and commercial wireless products are not only governed by regulatory rulings, but they are also governed by user expectations. Indeed, the EMC performance requirements in this aspect are generally more stringent than those of a regulatory nature. For example, a consumer would not be satisfied with a mobile cellular telephone installation in his or her car if it means that the FM and AM radio will no longer work properly, even if the mobile cellular telephone passes the regulatory agency requirements. A wireless data modem that causes television interference is equally unacceptable.

To ascertain that the wireless data product design satisfies the EMC requirements in the early phase of the design cycle requires very careful project design and planning. This is because most of the baseband ASICs and noisy digital components may not be available when the tooling commitment for the plastic enclosures, as well as the constraints for other mechanical components, have to be finalized. As a result, the transceiver design is usually debugged without the benefit of the noise contribution information from the baseband circuitry until the final phase of system integration. By then, the RF designer is confronted with a tight project schedule and budget, as well as limited manpower and system workaround flexibility.

One good strategy to mitigate these problems is to use equivalent noisy baseband devices as placeholders for the yet-to-be available baseband ASICs in the initial circuit layouts to present a realistic noisy background for the transceiver circuits.

For example, fast DSPs can be placed to execute endless loop routines, and oscillators with fast-switching interfaces can be placed as dummy placeholders at circuit card locations where the real devices will be eventually go. As a result, EMC problems such as receiver desensitization, broadband frequency emissions, PLL signal contamination as a function of circuit board partition, and EMC shielding requirements can be minimized when options for workaround or design alternatives are easily available.

All wireless communication devices, irrespective of whether they have transceivers or receivers only, are required to pass FCC Part 15.109 Subpart (a) or (b) [17, 18] emission limits dependent on their classifications for use in the United States. This is a requirement for all classifications, and is common to all wireless communication products. It applies to all wireless communications products irrespective of whether a particular product has a transmitter or not, and also irrespective of whether the product can be suitably classified under other relevant emission regulations. FCC Part 15.109 Class A devices are for commercial users and are relaxed by approximately 10 dB compared with FCC Part 15.109 Class B devices, which are for consumer usage. Tables 2.11 and 2.12 show their legal limits.

As can be seen from the FCC Part 15.109 requirements, the emission interference potential to other established domestic appliances, such as broadcast radios

Table 2.11
FCC Part 15.109 Limits (Radiated Emission Limits for Class B Devices Measured at a Distance of 3 Meters)

Frequency of Emission (MHz)	*Field Strength (μV/meter)*
30–88	100
88–216	150
216–960	200
Above 960	500

Table 2.12
FCC Part 15.109 Limits (Radiated Emission Limits for Class A Devices Measured at a Distance of 10 Meters)

Frequency of Emission (MHz)	*Field Strength (μV/meter)*
30–88	90
88–216	150
216–960	210
Above 960	500

and garage door opener sets, are still considerable, even if the regulatory emission limits are satisfied. It is the duty of the RF transceiver systems designer to impose his or her own specific emission limits as a supplementary requirement to that of the regulatory agency. Outside the United States, one is required to satisfy foreign regulatory limits, and they are generally dominated by VDE and CISPR equivalents of the FCC regulations. These limits are a few decibels more stringent than that of the current equivalent FCC Part 15 rule in most cases. In the European Union countries, the Information Technology Equipment (ITE) family of emission standards are in a state of transition. The real impact of the new EMC requirements can only be assessed after 1996 when they are promulgated.

2.8 CONCLUSIONS

RF design engineers who engage in the design of wireless devices that utilize the maturity of digital communication technology are now facing a new knowledge gap. This gap is a consequence of the disparity of terminology and parlance of the digital communication technology with that of established RF design and measurement techniques. The consequence of this knowledge gap manifests itself in an unusually high design failure rate of the transceivers intended for digital communication use.

However, the explosive demand for wireless applications is just beginning to be felt by designers and the financial investors in these applications. A multitude of traditional analog voice and low-efficiency data communications devices have been updated or replaced by highly efficient digital standards. It is now necessary for practicing RF engineers and technologists to bridge this knowledge and language gap quickly and effectively to be ready for the challenge. In this chapter, some of the introductory facets of this effort have been briefly presented in a narrative manner. Comprehensive tutorials that enable a digital communications engineer to convey his or her requirements fully in the "traditional" RF parameters, or vice versa, are few and are thinly scattered in the open literature. In most instances today, successful designs are only possible when there is a group of close-knit engineers of both disciplines. This should not be necessary when the knowledge horizon of the RF engineers is extended into the digital domain.

References

[1] TIA/EIA Interim Standard, "Cellular System Dual Mode Mobile Station-Base Station Compatibility Standard," IS-54B, April 1992.
[2] TIA/EIA Interim Standard, "Mobile Station-Base Station Compatibility Standard for Dual Mode Wideband Spread Spectrum Cellular System," TIA/EIA/IS-95, July 1993.
[3] J. E. Padgett, C. G. Gunther, and T. Hattori, "Overview of Wireless Personal Communications," *IEEE Communications Mag.*, Vol. 33, No. 1, Jan. 1995, pp. 28–41.

[4] K. Pahlavan, and A. H. Levesque, "Wireless Data Communications," *Proc. of IEEE*, Vol. 82, No. 9, Sept. 1994, pp. 1398–1430.
[5] A. Flatman, "Wireless LANs: Developments in Technology and Standards," *IEEE Computing & Control Engineering Journal*, Vol. 5, No. 5, Oct. 1994, pp. 219–224.
[6] FCC 47 CFR, 10-1-93 Edition, pp. 581–582.
[7] FCC 47 CFR, 10-1-88 Edition, pp. 212–443
[8] TIA/EIA Standard, "Land Mobile FM or PM Communications Equipment Measurement and Performance Standards," TIA/EIA-603, Feb. 1993.
[9] EIA/TIA Interim Standard, "Recommended Minimum Performance Standard for 800 MHz Dual Mode Mobile Stations," EIA/TIA/IS-55, Dec. 1991.
[10] EIA Interim Standard, "Recommended Minimum Standards for 800 MHz Cellular Subscriber Units," EIA/IS-19-B, May 1988.
[11] EIA/TIA Standard, "Mobile Station-Land Station Compatibility Specification," EIA/TIA-553, Sept. 1989.
[12] R. Sagers, "Intercept Point and Undesired Responses," *Trans. IEEE Veh. Technol.*, Vol. VT-32, No. 1, Feb. 1983, pp. 121–133.
[13] J. Y. C. Cheah, K. Karimullah, and M. E. Davis, "Wireless Data Modem , Theory and Implementation," Reading, MA: Addison-Wesley, to be published.
[14] J. Y. C. Cheah, "Analysis of Phase Noise in Oscillators," RF Design, Nov. 1991, pp. 99–104.
[15] J. Y. C. Cheah, "Frequency Reference in VSAT," *IEEE Trans. Communications*, Vol. 42, No. 2/3/4, Feb./Mar/April 1994, pp. 233–236.
[16] R. P. Gilmore, "Specifying Local Oscillator Phase Noise Performance: How good is good enough?" *IEEE International Symp. on Personal, Indoor and Mobile Radio Communications*, Sept. 1991, pp. 166–172.
[17] FCC Gen. Docket No. 89-118, June 29, 1989.
[18] FCC Gen. Docket No. 89-116, June 29, 1989.

CHAPTER 3

LOW-POWER RADIO-FREQUENCY ICS FOR PORTABLE COMMUNICATIONS

Portions of this chapter originally appeared in the April 1995 *Proceedings of the IEEE Special Issue on Low-Power Electronics* and appear here with permission.

Asad A. Abidi
Electrical Engineering Department
University of California
Los Angeles, CA

3.1 INTRODUCTION

The portable revolution is upon us today. It promises to empower individuals throughout the world by giving them low-cost access to information wherever they may be, thus allowing them to make informed decisions and to be more productive in business and at home without necessarily being tied down to a physical location. It is expected that in the near future, individuals will be equipped with capabilities of local computing and of communications enabling them to perform almost all the tasks that today require the equipment on the office desktop: the telephone, the computer, its connection to a ubiquitous wired network, the fax machine, and so on [1].

The portable revolution has been many years in the making. The personal broadcast radio receiver and cassette player, as pioneered by Sony, was a runaway

global success. The user could construct a private audio environment anytime and anywhere, using a device so small and light that its presence was easily forgotten. Portable computers brought about the next wave of change. The personal computer has had such a large impact on the broad working habits of individuals that without it they are lost. The luggable computer has evolved in a matter of one decade into the portable, the notebook, and the subnotebook. With the establishment of a wide area network (WAN) of radio paging transmitters, the personal radio pager has also became very popular since the mid-1980s. Today, the pager network spans the entire continental U.S. and many other parts of the world, enabling the user to receive alphanumeric electronic mail messages. Two-way paging is actively under development. The cellular telephone became widely available shortly thereafter. The user could hook into the international switched-telephone network through the nearest cellular base station with a portable transceiver, which too has scaled down remarkably in size until its weight and volume is the smallest practical [2].

There are many competing visions of how these various portable devices and services will evolve and integrate in the next few years. These are covered extensively in the popular press and in numerous keynote speeches in technical meetings. A common theme is that users will want a multimedia terminal, capable of wireless access to a global network that can transport communications, images, and databases to the user in an on-demand, interactive fashion [3, 4]. Such a terminal will have capabilities of computing, image acquisition and display, and, obviously, communications. It will likely derive as a hybrid of the various portable technologies available today.

What obstacles must be overcome to realize this vision? These include the very highly integrated electronics, effective displays, and a philosophy of design based on low power dissipation to prolong the battery life of the portable device. Sometimes single-battery operation will impose the additional constraint of operation at low voltage, as low as 1V, which will require entirely new ways of doing electronic circuit design.

In the past few years, most designers of mass-market digital integrated circuits (ICs) have been preoccupied with low-power operation [5]. Principles such as operating CMOS logic at the lowest possible supply voltage have become widely known, and power-down modes, gear-shifting of operating clock frequencies, pipelining and parallelism, subthreshold operation, and other such methods that were once the province of specialized areas such as electronic wristwatches and implantable biomedical devices are becoming commonplace [6]. Studies into the fundamental thermodynamic limits to the energy required for computation are being initiated or revived. There is good reason to believe that all this activity will lead to significant improvements in the conventional circuit and system design styles for digital signal processing and computation.

How will this activity affect the communications aspects of the portable device? What similar principles to low-power digital design are there for energy-efficient

wireless communications transceivers? These questions do not have simple answers. Low-power communication systems will result from use of the correct architectures, a sensible partition between analog and digital signal processing, low-power circuit techniques everywhere, and a judicious division between active and passive components. There is still not a widely known, integrated vision on this subject. Furthermore, there remains a gap between the IC design used in the portable applications described above and the new designs that will be required over the next few years for advanced portable communicators. Consumers are demanding a great deal more functionality and performance, which is stressing present-day technology to its limits, and wireless communicator design itself is in transition from the classic analog modulation techniques used over the past 50 to 70 years to more sophisticated methods using digital signaling formats and signal-processing methods in transceivers.

This chapter summarizes the key developments in the discipline so far, and from them forecasts wireless IC design trends in the near future.

3.2 KEY SIGNAL-PROCESSING ISSUES IN WIRELESS TRANSCEIVERS

Were it not for the advent of the portable communications revolution, radio technique would almost certainly have become a lost art. The first edition of the last definitive textbook on the subject dates to 1943 [7]. Whereas once radio engineering was synonymous with electronics [8], few university electronics curricula today offer a course on radio communications circuits. Only a few modern textbooks on radio design have been written in the past 25 years [9–13]. Radio communication methods, at least for nonmilitary applications, have remained relatively unchanged since World War II, and the evolutionary improvements in consumer equipment mainly owes to the use of high-frequency discrete transistors [14], smaller passive components, and building-block ICs that improve the long-term reliability and manufacturability of radio and TV receivers. The major impact of IC technology in these consumer items has probably been at baseband, in adding more user features. In contrast, the front-end radio architectures have evolved almost not at all in the past 40 or so years. For instance, ICs have contributed digital volume control, digital frequency tuning, and features to alleviate manual effort on the part of the user, but the RF and intermediate frequency (IF) sections still contain discrete and passive components in rather conventional architectures.

Why is this? It is partly because radio frequencies were too high for the low-cost IC technologies traditionally used in the consumer electronics industry. It is also because advances in component packaging alone have led to rapid downscaling in the size of consumer devices, often obviating the need to rethink the electronics. As a result, only a few individuals in a handful of institutions worldwide have concerned themselves with thinking about these problems. Today, as conventional

solutions no longer suffice for the future wireless communications devices, there is a rekindling of interest in this subject, which has led to much rediscovery and some invention. Baseband IC designers are now attempting to apply familiar techniques to wireless, while microwave IC designers are exploring what to them are low-frequency commercial opportunities for their technologies.

To set the stage for further discussion, some of the unique problems of radio receivers and transmitters are first described. Unlike familiar wireline communications, the wireless environment accommodates essentially an unlimited number of users sharing different parts of the spectrum, and very strong signals coexist next to the very weak. The radio receiver must be able to select the signal of interest while rejecting all others. It must do so using less than perfect active and passive components. There are two important problems in the receiver: *image rejection* and *dynamic range*. Image rejection relates to the receiver's ability to select the desired signal from the array of signals occupying the spectrum. Ideally, it might do so with a tunable bandpass filter, whose center frequency could be positioned at will in the RF and whose passband was one channel wide. A filter with this small a fractional passband does not exist. Instead, a practical RF bandpass filer, which may or may not be tunable, will preselect an array of radio channels including the one of interest (Figure 3.1). The other preselected channels are then removed at a lower IF by translating them in frequency with a downconversion mixer and centering the desired channel within a bandpass filter at IF. The other mixer input is a frequency-tunable local oscillator (LO), offset by IF from the desired channel. As the preselected band after downconversion will very likely occupy an interval greater than (0, IF) on the frequency axis, the IF bandpass filter will select both the desired channel and another *image* channel the mixer has translated to –IF. The subsequent detector circuit will be unable to distinguish between the desired and the image channels, and therefore

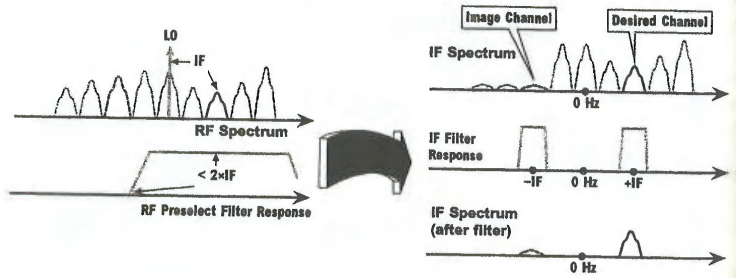


Figure 3.1 The image-rejection problem in radio receivers. The RF preselect filter passband must be determined with prior knowledge of the IF. The preselect filter is responsible for image channel suppression before downconversion.

its output will be the result of the superposition of both. However, if the stopband of the preselect filter lies less than 2 × IF away from the desired channel, it will attenuate the image, so only the desired channel will contribute energy at IF. The receiver designer first studies the available filter technologies and then chooses an appropriate IF, which yields an acceptable image suppression. A high IF relaxes the prefilter passband specification, but it also means that the downconverter signal requires high-frequency amplifiers, which are usually power-inefficient. Further, the IF filter requires a smaller fractional passband. In such cases, following image rejection at this high IF, the channel may be selected after downconversion to a second, lower IF. In such a double superheterodyne, or dual-conversion receiver, the first IF may actually lie at a higher frequency than the incoming RF to make image rejection easier.

The noise level and nonlinearity in the RF amplifier and first mixer usually set the receiver dynamic range. Consider reception of a weak channel surrounded by large undesired channels in the preselection band (Figure 3.2(a)). First, the input-referred noise of the receiver directly adds to the sought signal, corrupting its signal-to-noise ratio (SNR). Second, the large adjacent channels will experience the nonlinearities in the RF amplifier and mixer, and some of the products of the ensuing intermodulation distortion may overlap the desired channel. The IF filter cannot reject these unwanted products, which, like noise, will degrade the received SNR. As the receiver frequency response is normally bandpass, its nonlinearity is measured by applying two tones of equal amplitude closely spaced in frequency (f_1 and f_2) at its input, and measuring the rise in the third-order intermodulation products (at $2f_1 - f_2$ and $2f_2 - f_1$) with the input lever (Figure 3.2(b)). All other intermodulation tones usually lie outside the receiver passband. On a logarithmic plot, the third-order intermodulation level rises at a slope of three relative to the fundamental tone at the output. The two lines intersect at a point called the input-referred *third-order intercept* (IP3). The intercept point is usually extrapolated from measurements at low levels, because the receiver front end will saturate at large inputs. The 1-dB *compression point*, the input level that causes the receiver gain to drop by 1-dB relative to the small-signal gain, specifies the onset of saturation. The input-referred noise level may be included in this plot to define a spurious-free dynamic range (SFDR), although this is rarely used in radio specification. Usually the input noise figure (NF) and the input-referred interception point (IP3) are separately specified.

Similar specifications apply to the transmitter, which operates at much larger signals. Suppose, as is almost always the case, that the transmitter is required to emit a single-sideband suppressed-carrier output; that is, a signal where the LO and the image frequency have been suppressed. However, owing to circuit imperfections, it may also emit small amounts of the carrier and the unwanted sideband, which typically lie in the passband of the subsequent RF filter (Figure 3.3). These unwanted emissions may become interferers for adjacent channels. Nonlinearities in the power amplifier may also produce emissions of intermodulation products at other frequen-

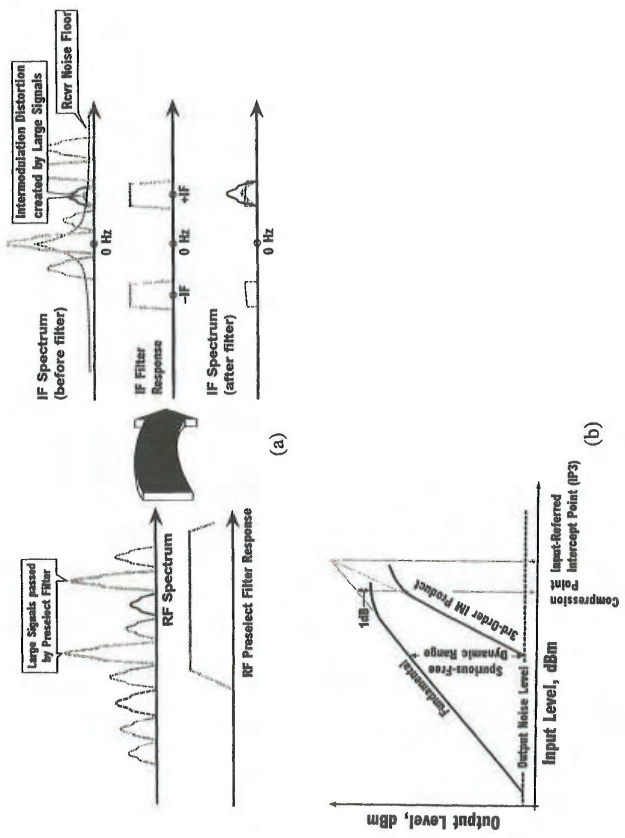


Figure 3.2 The dynamic range problem in the radio receiver. (a) Adjacent large signals may create intermodulation products superimposed on the desired channel. Receiver noise floor is a fundamental limit to sensitivity. (b) Receiver dynamic range is specified (box) in terms of extrapolated third-order intercept point and noise level.

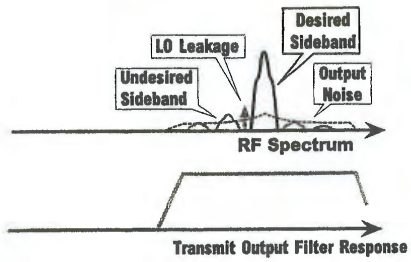


Figure 3.3 Transmitter imperfections (such as mismatches in a quadrature upconverter) result in the appearance of spurii in the emitted spectrum (referred to as "spectral regrowth"). The spurii may not be removed by the output filter, and may superimpose on adjacent channels. Emitted noise might overwhelm weak adjacent channels.

cies. Phase noise in the local oscillator responsible for upconverting to RF will convert to noise added to the signal, and the amplitude of this noise increases with the transmitted signal. This noise could possibly overwhelm nearby weak channels. Transmitter performance is usually specified in terms of the relative levels of unwanted signals to the desired signal and in terms of absolute spectral density of output noise at maximum output power.

To understand the rationale underlying receiver architecture, let us use as an example the familiar broadcast FM receiver. The architecture to be described is the same that Armstrong, the inventor of FM and the superheterodyne, had originally proposed for FM reception. The desired channel consists of a carrier in the 88- to 108-MHz band, modulated by up to ±75 kHz. Neighboring channels are spaced apart by 200 kHz. The receiver must select the desired channel while rejecting nearby channels, and it must be sensitive to a signal of a few tens of microvolts induced on the antenna. A simple FM antenna is wideband, and will pick up signals well outside the broadcast FM band. The low-noise bandpass amplifier in the front end may at best mildly attenuate the out-of-band signals—the actual channel selection must be done elsewhere. The RF amplifier uses an inductive load to resonate with the transistor and tuning capacitances, thereby transforming an inherently lowpass characteristic to bandpass centered on the frequencies of interest. The transistor f_{max} limits the highest frequency at which such a tuned amplifier can still provide a gain greater than unity. This figure of merit is familiar to microwave circuit designers and to device designers, whereas baseband IC designers deal more often with transistor f_T, the capacitance-limited unity current-gain frequency. In bipolar IC processes not optimized for small-signal high-frequency use, f_{max} is comparable to f_T, whereas in the best RF processes it may be twice f_T [15, 16].

It is impossible to select the desired channel at RF, because no tunable filters exist with the required fractional bandwidth of 0.15% (150 kHz/100 MHz). Therefore, following sufficient amplification at RF to overcome the noise level of the following circuits, the signal is mixed down by a variable-frequency local oscillator to a lower IF. Furthermore, a filter to select the desired channel at IF will have a fixed-center frequency and the fractional bandwidth in the passband will be larger. From this perspective, it is desirable to use as low an IF as possible. However, image rejection poses yet another constraint on the choice of IF. Conventional broadcast FM receivers use an IF of 10.7 MHz as a compromise. This IF guarantees that the image always lies outside the FM band (Figure 3.4). It is unlikely, however, that the preselect filter can entirely suppress this image, which will therefore either add noise or AM to the desired signal. However, the subsequent FM detector is inherently insensitive to both these forms of impairment. In this way, a medium-valued IF is made possible by exploiting properties of the detector, and thereby the fractional passband specification for the preselect filter is relaxed. The first mixer downconverts the entire FM band, with the desired channel centered at 10.7 MHz. A varactor-tuned Colpitts oscillator may be used as the first local oscillator. Prior to detection, a cascade of identical fixed-frequency ceramic bandpass filters, each with a 200-kHz passband centered at 10.7 MHz, passes the desired channel while rejecting neighboring channels [17]. In a high-quality receiver, the RF amplifier may be a discrete GaAs MESFET with a tuned load ganged to the LO tuning element, and another ganged tuned circuit may couple the antenna signal into the receiver [17].

To transplant this style of discrete radio circuit design to ICs, one would have to implement LC tuned circuits and filters on silicon. One can indirectly surmise these concerns in the Motorola series of IC design textbooks from the 1960s, which discuss loss in spiral metal inductors fabricated on silicon substrates [18] as well as issues relating to simulated inductors for active filters at IF for radios made with gyrators and capacitors [19]. However, on-chip spiral inductors of useful values

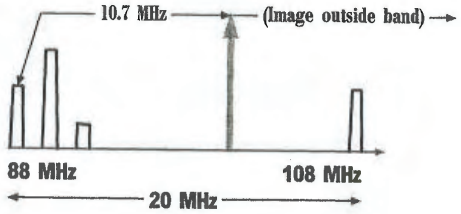


Figure 3.4 The 10.7-MHz intermediate frequency conventionally used in FM receivers is the lowest frequency that will guarantee that the image lies outside the FM broadcast band. The FM demodulator will reject a (non-FM) image channel entering the receiver. The 10.7-MHz passive bandpass filters in the IF strip select the desired channel.

were found to suffer excessive capacitance to the substrate, which lowered their self-resonant frequency to the point that they were not usable beyond the VHF band. It gradually became part of the collective consciousness of IC designers that on-chip tuned circuits are generally impractical.

When useful tuned amplifiers did appear on monolithic integrated circuits, it was not for the VHF-to-UHF range of relevance to consumer applications. Instead, it was military applications at much higher frequencies that drove the development of monolithic microwave integrated circuits (MMICs) during the 1980s. MMICs typically use MESFETs as the active device on semi-insulating GaAs substrates. This technology has enabled miniature radar, remote sensing, and communications at frequencies up to tens of gigahertz. The on-chip wavelengths are so small that monolithic distributed circuits may be built. MMICs take advantage of the semi-insulating substrate in two important ways. Transistors on these substrates have lower parasitic capacitance, which means that they amplify to higher frequencies. It is also possible to build low-capacitance interconnect with air bridge structures, and high-frequency passive components such as spiral inductors required for narrowband tuned circuits. Thus, on MMICs, the board-level design styles used hitherto by radio and microwave engineers could be miniaturized. However, over its many years of existence, GaAs MMIC technology has not had the major impact on consumer electronics that its adherents had hoped for. Makers of consumer electronics favor silicon IC technology wherever feasible because of its low cost, high yields, and the relative ease of mixing analog and digital circuits on a large scale.

It is anticipated that by the year 2000 about 300 million portable consumer wireless devices will be in use [20]. What IC technologies will enable the RF front end of these devices? Do miniaturization and long battery life call for architectural innovations in transceivers? What new circuit design styles will evolve in response? There is much curiosity and speculation on these matters, yet little is generally known about RF-IC design, or on the possible impact of large-scale integration and power-reduction strategies in the front end of wireless transceivers. The next sections present a brief survey of the use of IC technology in wireless receivers and transmitters since the 1970s to date, and from this projects some future trends. RF-ICs are roughly defined as integrated circuits operating in the band of frequencies from 400 MHz to 2500 MHz, which covers most consumer wireless communication devices. As opposed to MMICs, which were almost exclusively fabricated on III-V component semiconductor substrates at small scales of integration, mature silicon technologies will play a large, if not the dominant, role in RF-IC fabrication. It is the author's belief that in response to pressing demands of ubiquitous wireless access, both the underlying semiconductor technology and the design styles will rapidly evolve to realize the single-chip "VLSI radio" in the not too distant future.

3.3 ICs IN BROADCAST RADIO RECEIVERS

The two-way wrist radio has fascinated the popular imagination since its introduction in the popular American cartoon strip *Dick Tracy* in 1946 [21]. This is the ultimately

unobtrusive piece of consumer electronics. Let us now see how feasible it is to build an FM receiver of this size with microelectronics technology. For the average user to accept such a radio, its selectivity and sensitivity must be comparable to that of tabletop models. If, in the FM receiver described in the previous section, all the transistors in the receiver electronics were to be integrated onto one silicon chip, the radio would still need a considerable number of off-chip tunable inductors and ceramic filters, and in spite of state-of-the-art miniature packaging, the components could not plausibly all fit into a wristwatch. Neither would the power dissipation be commensurate with the life of a wristwatch battery.

The first generation of silicon bipolar ICs developed in the late 1970s for the IF and baseband portions of broadcast receivers more or less contained the transistors of conventional receivers assembled onto one or more ICs [22–24]. However, integration did afford freedom to use transistor-rich circuits for higher performance. Circuit techniques such as double-balanced mixers using Gilbert analog multiplier, PLLs as FM demodulators, and use of balanced on-chip signal paths to attain greater immunity to pickup and common-mode noise became widely used as a result. By eliminating many of the coupling coils and other noncritical discrete components found in older radio circuits, ICs contributed to lowering the cost of assembling and aligning the final product. In the RF section, though, the receivers still used the conventional 10.7-MHz IF superheterodyne architecture implemented with shielded discrete-component circuits.

In the early 1980s, Kasperkovitz (at Philips) [25] made the first significant explorations into alternative architectures for highly integrated radio receivers. He realized that to reduce receiver size and power dissipation, it was very important to eliminate the many off-chip passive components. If certain passive inductors, capacitors, and resistors could not be eliminated, they could at least be packaged in surface-mount outlines for very small size. However, neither the volume of the IF ceramic filters could be readily scaled down, nor could their characteristic impedance be scaled much above 50Ω. Each filter requires an on-chip analog driver of comparable impedance. Alternatively, the filter may be realized on-chip as an active bandpass circuit of sufficient selectivity and dynamic range. Although a gyrator-capacitor-based active filter is possible in principle, small phase shifts in the gyrator transistors at the 10.7-MHz IF can seriously upset the filter passband shape. Active resonators are also known to suffer from a larger internal noise level than their passive counterparts, and this discrepancy worsens with increasing pole-Q and pole frequency [26]. Kasperkovitz solved the problem with an *architectural innovation*, by dramatically lowering the IF from 10.7 MHz to 70 kHz. This makes it a great deal easier to implement an IF active channel-select filter, which now need only be lowpass. Further, at a given dynamic range, the power dissipation in an active filter also scales down with the IF [26]. The low IF eliminates the off-chip channel-select filter and reduces power dissipation, both very desirable properties. But what of the image frequency, the principal reason for the choice of 10.7 MHz?

At a 70 kHz IF, the image frequency lies halfway to the adjacent FM channel (Figure 3.5(a)). The image, therefore, is the interchannel noise in the FM band. As an RF preselect filter cannot possibly reject an image this close to the desired signal, it will pass unattenuated to worsen the received signal-to-noise ratio (SNR) by 3 dB. Another consequence of this choice of IF is that after the first downconversion, an instantaneous frequency deviation in the received FM signal of more than 70 kHz will alias around dc to produce distortion. This is avoided by compressing the frequency deviation to ±15 kHz with a negative feedback frequency-locked loop prior to downconversion (Figure 3.5(b)). The FM monostereo radio [27–29] requires, in addition to the single-chip receiver, only 15 small capacitors and 2 inductors, and this collection of parts readily fits inside a wristwatch. A miniature earphone is plugged into the watch and the earphone lead serves as the antenna. The radio when active drains 8 mA from a 4.5V supply.

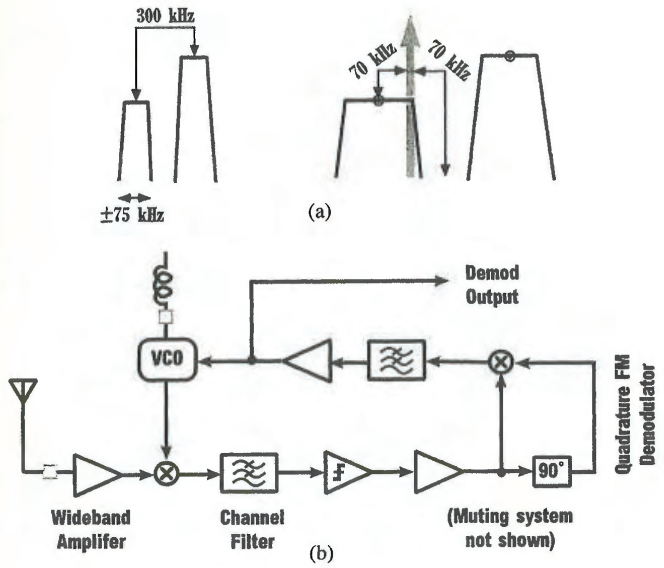


Figure 3.5 (a) An alternative choice of IF in the FM band, which places the image in the gap between adjacent channels. The IF strip, including the channel filters, now operates at a 70-kHz frequency. (b) An FM receiver using a 75-kHz IF. The channel filter is an active-RC implementation on-chip. A frequency feedback loop compresses the incoming frequency swing. Except for a tuning inductor, no high-frequency off-chip components are required.

Sony, one of the world's leading makers of miniature radios, has also recently modified its integrated FM radio-receiver architecture from the conventional 10.7-MHz IF [30] to low IF [31]. In the new architecture, a high first IF of 30 MHz is used, so any out-of-FM band image falls in the stopband of a fixed 80 to 110 bandpass preselect SAW filter after the antenna. A wideband IF amplifier boosts the received signal level of *all* FM channels *without* any filtering—amplification at 30 MHz is not a problem on this modern silicon bipolar IC process. Further, at a 30-MHz IF, the entire broadcast FM band falls to one side of the LO frequency, which means that any channel in the FM band may be selected after the first downconversion. Following this, another mixer converts to a low second IF of 150 kHz, and thereafter an on-chip ninth-order active-*RC* lowpass filter rejects adjacent channels. The low second IF, however, will pass an image FM channel as well as the desired channel, and as there is no filtering at all at the first IF, an *image-rejection mixer* is used for the second downconversion (Figure 3.6). The variable-frequency first LO tunes the desired channel, while the second LO, which must produce quadrature outputs for the image-rejection mixer, is at a fixed frequency.

The image-rejection mixer provides a trigonometric solution to a difficult filtering problem [32]. The desired channel and its image are frequency-converted into two paths by mixers driven by quadrature phases of an LO. The mixer outputs are then phase-shifted 90 degrees with respect to one another. The sum of these two signals will select the desired channel and suppress the image while, vice versa, the difference will select the image. The extent of image suppression depends on the gain matching of the two paths, and on the phase accuracy of the LO quadrature outputs. For these reasons, this concept has only become practical with IC technology, where the two paths are well-matched on-chip and track each other over temperature. Image suppression on the Sony chip is limited to about 40 to 45 dB by residual gain mismatch in the two paths. An allpass active RC-CR filter produces 90-deg phase-shifted versions of the downconverted input. This receiver drains about 15 mA from voltages as low as 0.9V in either FM or AM mode. All the necessary transistors are integrated on-chip—the RF amplifier portion, however, uses an off-chip load inductor, and the local oscillator needs an off-chip LC tuned circuit.

The circuit techniques that enable sub-1V operation are also interesting (Figure 3.7). The antenna signal drives the emitter of a common-base NPN, which forms the tail of a differential pair. The input resistance of the common-base stage matches the antenna impedance. The signal develops at one inductively loaded collector of the differential pair, while the other collector dumps a fraction of the signal current into the supply in response to an automatic gain control (AGC) differential control voltage. Following amplification by a resistively loaded differential pair in cascade, the balanced RF signal is level-shifted into the first mixer, a simplified double-balanced Gilbert-cell with resistors instead of current sources in the tails.

As inductors do not drop a dc voltage, they make convenient loads for stacked transistor circuits operating at a low supply voltage. The signal on the transistor

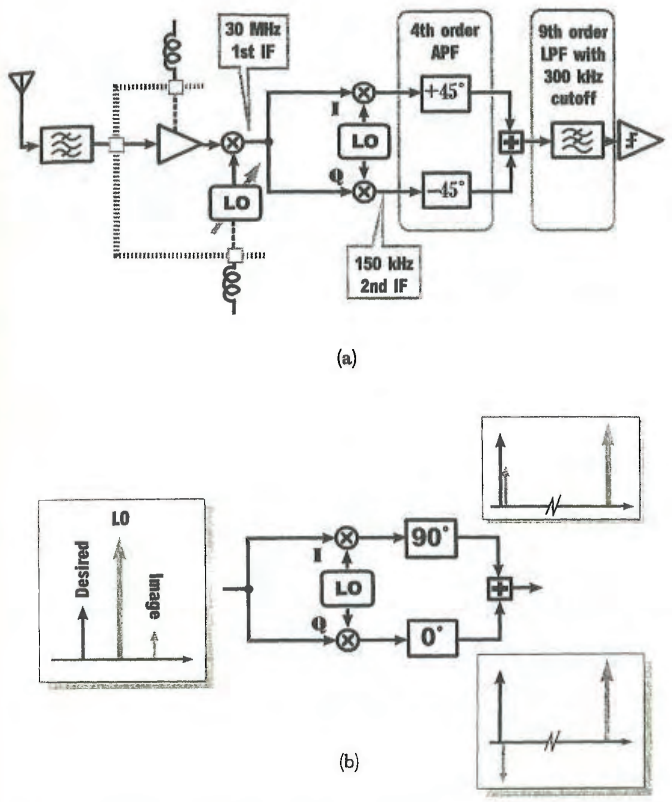


Figure 3.6 (a) An alternative architecture for a single-chip FM receiver, with a 30-MHz IF chosen for strong image rejection and to translate all the channels in the FM band to IF on to one side of dc on the frequency axis. The image-rejection mixer at 150 kHz selects the desired channel while rejecting the undesired one 300 kHz away. Aside from a noncritical RF preselect filter, the remaining filters are active on-chip, including the phase shifts in the two arms of the mixer. (b) The image-rejection downconversion mixer. The image and desired tones are at positive and negative frequency offsets from the local oscillator and are discriminated in the mixer by a relative inversion of polarities in the two arms.

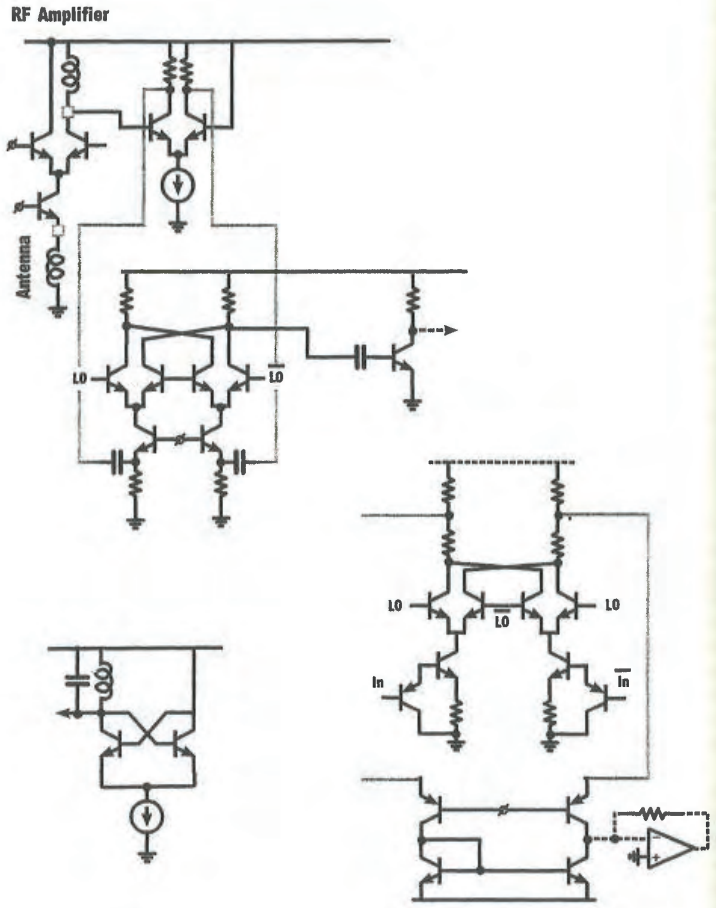


Figure 3.7 Low-voltage circuits capable of operation with a 0.9V supply. Note how the antenna carries the RF amplifier bias current, and the use of on-chip capacitors for level-shift in the RF signal path. As the signal is downconverted in frequency, active level-shift and op amp circuits appear. The cross-coupled differential pair negative resistance is popular for LC oscillators.

end of the inductor will swing above the supply. This is the case in the local oscillator, which implements a negative resistance of $-1/g_m$ with a cross-coupled differential pair, causing an LC tuned circuit across it to oscillate. The oscillation amplitude is limited by the differential pair nonlinearity to a small multiple of kT/q.

As the signal propagating through the receiver downconverts in frequency, low-frequency circuit techniques and devices with a lower f_T are used. For instance, the 30-MHz IF signal couples into the second mixer through PNP emitter followers. At the 150-kHz second IF, common-base PNPs and current mirrors level-shift down the signal, which is further amplified in an op amp.

This receiver is a significant example of how architectural rethinking combined with appropriate IC design styles has resulted in a very different solution to the well-established broadcast FM receiver.

3.4 ICs IN WIRELESS PAGING RECEIVERS

Miniature wireless communicators to page people on the move were first developed in the late 1950s. The Bell System's Bellboy™ paging receiver [33] anticipated many of the concepts underlying today's pagers. The system operated at a 150-MHz RF, addressing a receiver by frequency modulating the carrier with a unique set of three tones, which the intended receiver recognized at baseband by the simultaneous response of three passive reed resonators. The superheterodyne receiver operated at 4V using a total of only ten transistors [34], a notable early example of low-power and low-voltage circuit design. The low IF of 6 kHz meant that simple, capacitively coupled 10-kHz lowpass (rather than bandpass) filters could select the desired channel. Also, all low-frequency amplifiers after the RF section used transistors biased at small currents. The two stages in the cascode RF amplifier were inductively coupled to share, or reuse, the same bias current, a power-saving method found even in today's MMICs.

Paging receivers have been in continuous evolution since then. They are supported by a sophisticated nationwide wireless infrastructure. The modern pager uses digital signaling at rates anywhere from 500 to 1,200 bps, encoding binary data with a simple positive or negative offset of the carrier frequency—the binary frequency-shift keyed (FSK) modulation. In spite of market pressure to reduce the battery drain and miniaturize the unit, paging receivers until the early 1980s used conventional radio architectures without exploiting the powerful simplifications implied by this signaling scheme. Such a double-superheterodyne receiver [35] might consist of a *first upconversion* of the received signal to suppress the image channel with an RF crystal bandpass filter, and then a *downconversion*, followed by channel selection with a ceramic bandpass filter. An analog frequency discriminator demodulated the FSK. Much as in broadcast FM receivers, this architecture required tuned amplifiers and passive filters, which constrained further miniaturization of the paging receiver.

Vance, at ITT Standard Telecommunication Laboratories, first realized that by taking the idea of low IF to its limit with *zero IF*, an FM receiver could be scaled down to one chip with only one or two passive RF components. A quadrature downconversion mixer could discriminate positive and negative frequency modulation centered around dc [36]. There is now *no* image to be rejected, and a lowpass filter suppresses adjacent channels. If applied to analog FM as in broadcast signals, however, this scheme suffers from the dc offsets in the amplifiers and their flicker noise, which will seriously corrupt the SNR at midchannel. Furthermore, the limiting amplifier that analog FM receivers use in place of AGC does not respond to dc inputs, because there are no zero crossings. In paging receivers, however, the spectral energy clusters in two lobes on either side of dc owing to the relatively large modulation index, and zero IF (or direct conversion) is exactly the right solution [37, 38]. This remarkably simple receiver (Figure 3.8) consists only of a quadrature demodulator, lowpass filters in each arm, limiters, and a D-type flip-flop detector. When integrated on an early bipolar chip, it drains 2.5 mA from 1.8V when active, although in standby the current drain falls to a mere 50 μA. Large-value off-chip capacitors are used for ac coupling and for the lowpass filters. A second low-frequency digital CMOS IC performs all the user interface functions. Data is encoded by offsetting the carrier frequency by ±4.5 kHz, so capacitive coupling with a corner frequency below 4 kHz blocks out receiver dc offsets and lower frequency flicker noise from corrupting the SNR. A 10-kHz lowpass filter suppresses high-frequency out-of-band noise and adjacent channels. Vance also pointed out that data in a binary FSK signal downconverted in quadrature to dc may be recovered by a simple flip-flop when the limited output from one arm of the downconversion mixer is applied to the D-input, and from the other arm to the clock input. Although exceptionally simple, this flip-flop detector makes instantaneous decisions and therefore has poor immunity to a single noise spike. More sophisticated detectors must be used to get acceptably low bit error rates (BER) at typical received SNRs. For instance, the inputs of two flip-flops may be cross-coupled to the quadrature channel outputs, and the decision may be derived from the analog average of the two outputs [37]. The optimum binary FSK detector in the presence of Gaussian noise correlates the downconverted signal with the two possible offset frequencies, integrates the output, and declares a valid bit when one of the integrator outputs crosses a threshold [39]. Even the most complex detector will dissipate a small power, because it operates at the low baseband data rate.

Pager ICs from Philips originally used a frequency-offset receiver principle [40, 41] whereby the local oscillator frequency is adaptively offset from the received carrier by 2 kHz, thus converting the FSK tones to 2.5 kHz and (aliased to) 6.5 kHz. This avoids a quadrature downconversion to differentiate between positive and negative frequency, but it requires a fairly sophisticated automatic frequency control. A frequency discriminator detects data. In addition to the local oscillator crystal (operated here at its fifth overtone), this chip requires three off-chip tuned

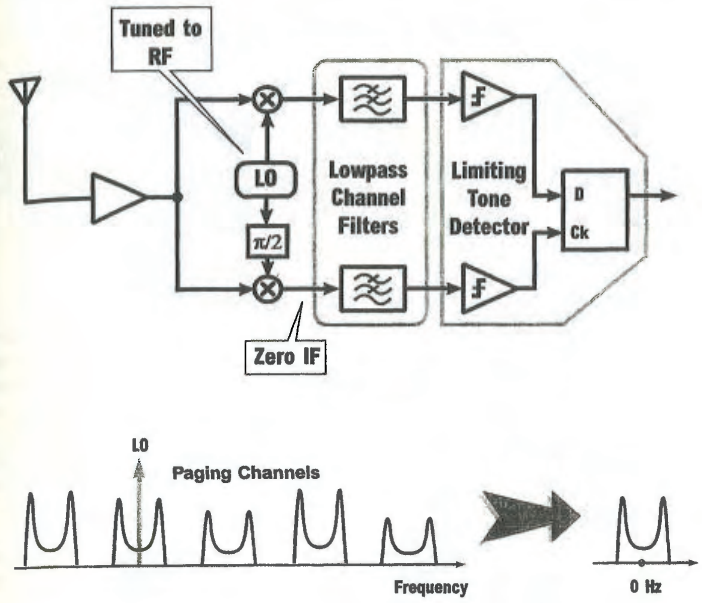


Figure 3.8 A direct-conversion single-chip receiver for FSK demodulation. There is no image channel here, and all channel selection filtering is on-chip at baseband. Note also that no AGC is required, only limiters, because all information is contained in the zero crossings. The signal path must be capacitively coupled to suppress the undesirable dc offsets in the receiver electronics.

circuits. Philips later recognized the simplicity of a zero-IF receiver for this application [42]. In this single-chip receiver, the RF amplifier requires one off-chip inductor, the quadrature phase-shift circuits in the mixer another; both inductors are combined into one off-chip signal path. The channel-select filters are entirely on-chip. They consist of a third-order active RC lowpass filter, followed by a seventh-order gyrator-based lowpass filter with a 15-kHz cutoff. Most of the signal amplification occurs at these low frequencies.

NEC's paging receivers evolved from the supeheterodyne [35, 43] to direct conversion [44] in an effort to reduce receiver volume and parts count. Others have developed similar zero-IF bipolar integrated front-ends [45, 46]. A notable feature of the chips is that much of the die area is taken up by the capacitors for ac coupling

the signal path and for the on-chip lowpass filters. Most pager ICs operate at supplies of 2V to as low as 1V and are implemented in silicon bipolar technology. In addition, full-featured pagers require 20,000 gate-equivalent digital ICs for the user interface [47], low-voltage EEPROMs for customization and software, and capability to drive a liquid crystal display. The basic paging receivers can fit within a wristwatch [48]. The direct-conversion FSK digital paging receiver concept has also been successfully used at very low carrier frequencies (hundreds of kilohertz) with much lower data rates in implanted devices for biomedical applications [49].

3.5 ICs IN CELLULAR TELEPHONE TRANSCEIVERS

Mobile and handheld cellular telephones are the first widespread two-way radios for consumer use. They were preceded by cordless telephones for local area use. These wireless telephones must meet stringent demands for low weight and volume, long battery life, low cost, and reliable network access to be successful with consumers. In contrast, walkie-talkie transceivers were always aimed at specialized markets and did not face these pressures for miniaturization. The average consumer, for instance, will not voluntarily accept a transceiver of the size and weight that policemen or soldiers carry as part of their outfit. Further, to support a large numbers of users in a crowded radio spectrum, wireless telephones use more internal signal processing than other common transceivers and must be capable of connecting to the public switched telephone network [50]. Features such as digitally selected channels, direct-sequence spread spectrum, and diversity selection are now becoming common. The transceivers perforce must use highly integrated, low-power electronics. Thus it may be said that with the advent of the modern cellular telephone, the conventions of wireless design are being re-examined, and sometimes rewritten.

The first generation of cellular telephones carried voice signals by analog frequency modulation of a carrier. In the U.S. AMPS system, for example, the handset receives at a carrier selected from the 869- to 894-MHz band, while it transmits on a carrier in the 824- to 849-MHz band. These 25-MHz widebands are separated by 45 MHz between the uplink and downlink, enabling the user to talk and listen at the same time much as on the wired telephone (users are not too fond of the "over, over-and-out" protocol). An antenna duplexer suppresses coupling from the transmitter into the sensitive receiver, acting as the equivalent of a two-to-four wire hybrid transformer in a telephone. This duplexer is a passive three-port designed to pass energy in the receive frequency band from the antenna port to the receive port, while attenuating energy in the transmit band from the transmitter into the receiver. It is either made with high-dielectric ceramic resonators [51] or with a SAW filter and coaxial resonator in parallel [52]. The receive portion of a conventional handset resembles a broadcast FM receiver (Figure 3.9). If the local oscillator at the first mixer lies at a higher frequency than 894 MHz, the image is guaranteed to lie outside

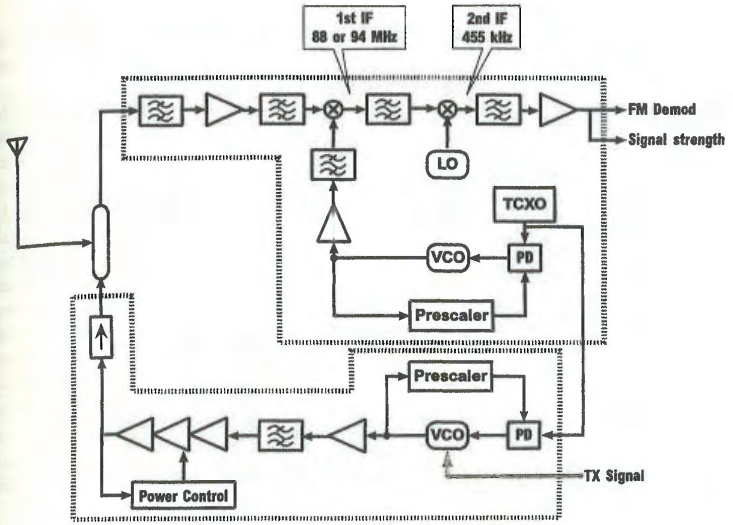


Figure 3.9 The transmit and receive front-end of an analog cellular telephone. Receiver uses conventional double-superheterodyne architecture, while transmitter is one-step upconversion.

the AMPS band. Furthermore, an LO offset of more than 45 MHz ensures that the image is attenuated by the receive band SAW filter, suffering at least 20-dB loss in each of the two filters in the handset [53]. A first IF of 90 MHz is therefore often used; this also avoids problems caused outside the handset by parasitic LO leakage through the antenna [54]. The desired channel is selected by locally synthesizing the first IF. This is followed by downconversion to a fixed second IF of 455 kHz, then demodulation by a frequency discriminator.

The first generation of small-scale ICs for portable communication devices offered building blocks for the intermediate-frequency chain, such as the mixer and LO conventional single or double-superheterodyne receivers [55], the IF amplifier chain and signal-strength indicator [55], or a stand-alone image-reject mixer [56]. Today, almost every major semiconductor company with an interest in the communications market offers building block ICs at this scale of integration.

Although ICs entered the IF portions, the RF front-end circuits continued to be made from discrete components. An RF amplifier and a first mixer may be mounted with the associated filters on a dense miniature board [57, 58]. Discrete bipolar transistors responded to needs for portable RF applications by offering, for

instance, a low noise figure and f_T exceeding 5 GHz at less than 1-mA bias currents [59, 60]. Manufacturers of passive components, too, have steadily scaled down their package sizes for high-density board mounting.

There is little argument, though, that the RF front-end components must also be integrated to reduce power dissipation. In competing with passive solutions, the RF ICs must cross some important thresholds of low price and high performance [61]. However, they offer the prospect of an order-of-magnitude reduction in physical volume of the front-end electronics, and power savings will accrue by routing RF signals at a high impedance on-chip, while eliminating the low characteristic impedance interconnects between discrete packages. Various reasons have been advanced for why GaAs MMIC technology is now the right choice for cellular telephones [62–64]. They are summarized as follows: first, that owing to the semi-insulating nature of the GaAs substrate, reasonable size inductors may be integrated with transistors to make high-frequency monolithic tuned circuits, which allows for lower current operation at a given frequency than would be possible with RC broadbanding techniques; and, second, that MESFETs afford lower noise figures at a given bias current than a bipolar transistor in a comparable silicon bipolar technology. Most GaAs MMICs integrate front-end components for cellular applications at a small scale. For instance, a chip may integrate a tuned RF low-noise amplifier in the 900-MHz band, or a mixer and a local oscillator [64–72].

The typical RF amplifier may consist of only one or two FETs, with LC matching circuits on the input and the output ports for stand-alone operation in a 50Ω environment (Figure 3.10). A powerful and popular method to match the capacitive FET input is to insert a series feedback inductor, L, in the FET source, which at high frequencies contributes a *resistance* $g_m L/C_{GS}$ at the input port [73]. This method is preferred to resistive feedback found in wideband amplifiers for impedance matching because, unlike feedback resistors, the inductor does not degrade the noise figure. MMICs from Matsushita favor the use of dual-gate MESFETs with RC matching instead of inductors [74], possibly because spiral inductors consume too large a chip area. GaAs IC designers must closely watch their chip area to remain competitive in price. The various low-noise amplifiers operate in the 0.9- to 2-GHz bands, with gains of 15 to 20 dB and noise figures of around 3 dB. These submicron MESFET ICs drain anywhere from 3 to 5 mA.

The received RF is typically downconverted by the local oscillator modulating the conductance of the mixer FET through a cascode FET (Figure 3.10). The mixers yield conversion gains greater than 10 dB, noise figures of 10 to 12 dB, and third-order input-referred intercept points of −5 to 0 dBm. Dual-gate MESFETs are naturally suited for mixer use, and offer a similar performance [74]. A recent 0.7-μm GaAs MMIC offers an LNA mixer pair draining 3 mA from 3V, with the LNA producing a 13-dB gain, a 3.6-dB noise figure, and an input-referred IP3 of −11 dBm [75].

Slicing across the system a little differently, another MMIC implements the downconversion and upconversion mixers for the receiver and transmitter, respectively, integrating their shared LO on the same substrate [76]. The four-FET switch mixer (Figure 3.11) is very linear, but it suffers from two disadvantages: it requires a large LO drive to turn the switches on and off [77]; and, unlike the bipolar transistor Gilbert-cell analog multiplier, the switch mixer is lossy, requiring additional signal amplification from the following stages.

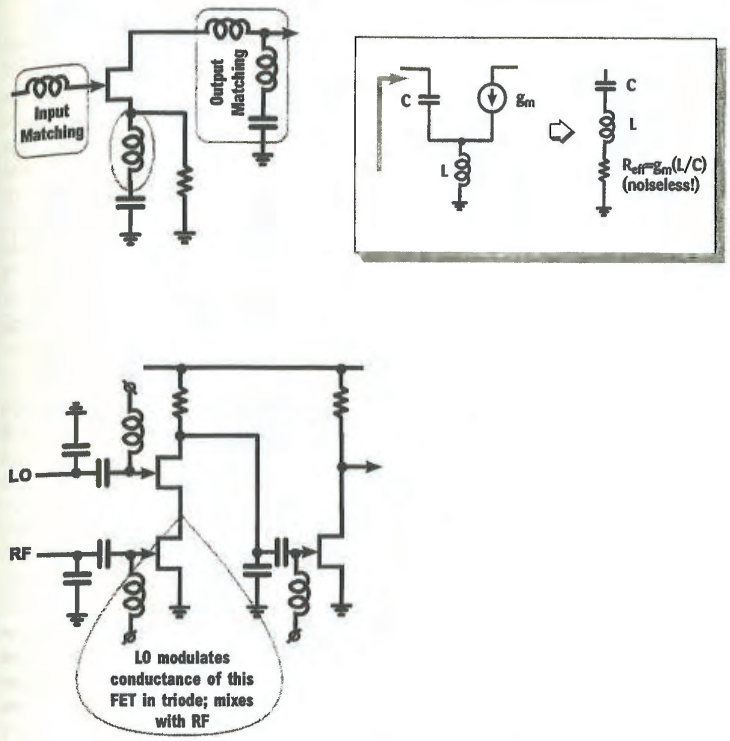


Figure 3.10 Typical stand-alone GaAs MMICs implementing low-noise amplifier (above) and downconversion mixer (below). Concept of series-feedback inductor for noise matching shown in box.

The current generation of analog cellular telephones transmits power levels of more than 1W (30 dBm). RF power amplifiers are usually packaged in a separate module with some form of integral heat sink. A preamplifier (or, in radio terminology, exciter) in the transmitter section boosts the modulated carrier level close to 0 dBm

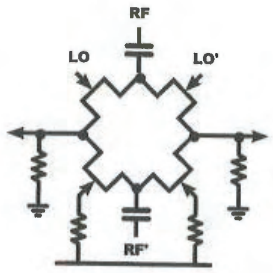


Figure 3.11 A four-FET commutating switch mixer. The balanced LO signal switches the FETs, which chop the RF signal to produce the sum and difference frequencies.

to drive the power amplifier input. The power amplifier module itself usually consists of a cascade of two or three FETs, tapering up to a single large-size FET that will deliver the required signal current into the antenna load. Furthermore, as the power amplifier is the largest single source of battery drain, it must have a high *conversion efficiency*. Much as in baseband power amplifiers, an efficient RF power amplifier is biased close to cutoff to reduce the dc standing current, and then driven by the input signal in Class A-B or Class B mode. Narrowband filters at the amplifier output remove harmonic distortion caused by nonlinear operation at RF [78]. These filters may be merged into the passive matching networks required for optimum power transfer from the amplifier to the load. As designers of RF power amplifiers have observed [79, 80], much of their work consists of synthesis and iteration of the interstage, input, and output *matching networks*. It is easier to design efficient power amplifiers for constant-envelope modulations, such as analog FM or digital FSK, where distortion may be tolerated because the useful information is all contained in the zero crossings. Further, it is argued that owing to the lower parasitic capacitance of a GaAs MESFET relative to silicon devices, a GaAs power amplifier at 1W power levels affords a higher efficiency (~ 60%) compared to silicon bipolars or FETs (~ 45%) [47]. A monolithic power amplifier consisting of a four-stage cascade of MESFETs with on-chip lumped LCR input and interstage matching networks, delivers 1W at 900 MHz at 63% efficiency from a 5.5V supply [80]. The distributed-element output matching network at 900 MHz would be exorbitantly large on an IC, and is therefore printed on an off-chip alumina substrate. The high efficiency is attributed to an improved method of suppressing the second harmonic at the amplifier output. This multicomponent module approach to mate ICs with matching networks is widely used in GaAs power amplifiers [81]. Another 900-MHz power module uses two discrete MESFETs, wirebonded to a hybrid IC containing a combination of distributed circuits and chip capacitors and resistors in the matching network to

attain 65% efficiency when delivering over 1.3W from 4.7V supply [82]. The output FET is 12-mm wide, with a 1-μm channel length. A separate negative supply is required in many of these MESFET power amplifiers to bias the gate. The circuit was recently modified [83] to produce the same output power equally efficiently, now with a single 3.5V supply and FETs of 0.5-μm channel length. The desired low-voltage operation, much more suitable for one battery, is the result of an improved device structure. The matching networks in the RF signal path are fabricated on separate passive-only GaAs ICs, which the authors claim are three times cheaper than GaAs ICs with FETs, while the bias networks reside on a miniature PC board. The various components are wirebonded to one another, and mounted on an AlN substrate with ten times higher thermal conductivity than alumina. Power amplifiers may also be built with silicon NMOSFETs: among the major semiconductor vendors, Hitachi has pursued this option. A MOSFET gracefully accepts large voltage swings on the gate without possibility of Schottky conduction, as in a MESFET gate. With an offset-gate FET structure and 0.8-μm channel length, a MOSFET power amp delivers 2W at 1.5 GHz with 55% efficiency from a single 6V supply [84]. The power amplifier module embeds these FETs into matching networks [85].

The foregoing discussion is not meant to imply that cellular telephones mainly use GaAs MMICs in the RF front-end. Silicon bipolar and BiCMOS technologies have made tremendous strides in improving f_T, and today silicon MMICs offer comparable performance to what is available in GaAs. A recent 1-GHz BiCMOS LNA-mixer combination [86] uses familiar circuits to baseband designers, such as a Gilbert-multiplier type mixer, and the chip yields a comparable performance to the MMICs described above. The RF signal path consists of bipolar circuits only, while the FETs are used as switches to select various power-down modes. There are no on-chip inductors to tune the low-noise amplifier, which is wideband; instead, the LNA output is routed off-chip into a passive bandpass filter, then returned to the chip for downconversion. This is part of a complete chipset from Philips for a digital cellular telephone handset [87].

The first LO in a cellular telephone receiver is programmed to the incoming RF channel with a PLL synthesizer, while the second LO at IF is fixed in frequency. The transmit LO oscillates in yet another frequency band. Both the receiver and the transmitter therefore require 900-MHz voltage-controlled oscillators. These are most often implemented with bipolar transistors, whose very small flicker noise means low phase-noise sidebands in the VCO. A stripline resonator sets the nominal frequency of a Colpitts oscillator, and this is voltage-controlled by a varactor diode in parallel [54, 57, 88]. Phase noise levels of –110 to –120 dBc/Hz are attained at a 50-kHz offset from the oscillation frequency. The VCO application has created a brisk demand for varactors with large voltage coefficient and low loss [60]. The VCO and its buffer may be connected in series to reuse the bias current between the two

stages [57, 88] (Figure 3.12). The nominal oscillation frequency may be slaved in a frequency multiplying PLL to a 12- or 15-MHz crystal oscillator.

The Motorola MicroTac, first introduced in 1989, set an industry standard for a miniature cellular handset. Small handsets [89] continue to use conventional architecture, but attain a small size with discrete filters in miniature packages [52, 53, 58, 90, 91] and with lower power electronics, which reduces battery weight. The electronics entail carefully designed standby modes, better software for the onboard microprocessors, and, among other items, lower power PLL prescalers—digital counters that must operate all the time for frequency synthesis and that often posed a significant power drain [92, 93]. In the transmit mode, greater efficiency is sought in the power amplifier to prolong battery life, as well as smart control circuits to maintain maximum efficiency across the range of output power levels [89]. It was generally agreed that by 1992, the handset volume should scale down to less than 150 cc, and that its weight should be less than 230 gm [20, 94]. NTT demonstrated the prototype of such a telephone in 1991 [95], while noting that the volume should be no smaller than this target, both for the sake of ergonomics and to ensure adequate heat removal from the package so that the internal temperature rises no more than 15°C. These small telephones might be powered by a 6V NiCd battery with 400-mA · h capacity.

With the emergence of the IS-54 standard, there is some convergence in the United States on a digital cellular technology. Upward compatibility is sought with the existing analog cellular bands, which is why handsets conforming to IS-54 are referred to as dual-mode cellular. A recent 12-GHz f_T silicon bipolar transceiver IC

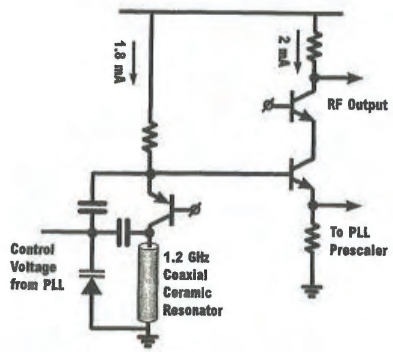


Figure 3.12 Colpitts oscillator with coaxial resonator. Varactor controls frequency with control voltage. Series-connected buffer distributes RF oscillation to mixer and to PLL prescaler.

from AT&T uses a conventional double-superheterodyne architecture with an 80-MHz first IF, and after IF amplification the signal path bifurcates into a conventional 455-kHz second IF FM demodulator or a digital I-Q PSK detector [96]. An interesting alternative is to downconvert the second IF, without preceding AGC, in a delta-sigma A/D converter for subsequent digital baseband signal processing [97].

3.6 DIGITAL CELLULAR AND CORDLESS TELEPHONES

Analog cellular telephones use the frequency spectrum inefficiently. The modulation schemes consume a large bandwidth and every cellular telephone transmits at constant power all the time it is in use, thereby appearing as an interferer to other users at nearby frequencies. As a result, the spectrum allotted to cellular phones in large metropolitan areas nears exhaustion.

Digital cellular telephones are one solution to better utilize the scarce spectrum. They use more efficient modulation schemes, such as minimum FSK or PSK, and multiple users may share time slots on the same part of the spectrum. The complex modulation formats used in these telephones and the greater capabilities required to withstand nearby blocking signals are prompting large-scale integration of the RF and IF electronics. Examples of digital telephony standards are the European GSM, North American Digital Cellular, and the emerging Japanese Personal Handy Phone (PHP) [98]. The salient characteristics of these various systems, as well as key digital cordless standards, DECT and CT-2, are summarized in Table 3.1, as well as Table 2.3 in Chapter 2.

The typical handset involves an RF/IF front-end, followed by a baseband digital signal processor (Figure 3.13). The first significant set of RF-ICs for GSM handsets appeared in 1990. Notable among them is a receiver and transmitter 7.5-GHz f_T silicon bipolar chipset from Siemens (Figure 3.14) [99, 100]. The double-superheterodyne receiver uses a selectable first IF of anywhere from 45 to 90 MHz, which places the image channel in the stopband of the 25-MHz-wide RF preselect bandpass filter. A 71-MHz IF, for instance, guarantees that the image lies outside the GSM band. The desired 200-kHz channel is selected by a SAW filter at the first IF. The receiver IC provides an on-chip buffer to drive this filter. Following an RF amplifier with an off-chip LC tuned load, the input signal is downconverted by a Gilbert-cell mixer. There is AGC with 70-dB range and a signal-strength indicator at the first IF, and the selected channel is then downconverted to baseband in a quadrature demodulator to acquire both amplitude and phase for GMSK vector demodulation. The on-chip gain may be as large as 80 dB at the highest VGA setting, but as it is distributed in different frequency bands there is little on-chip crosstalk, and the monolithic receiver operates stably. The receiver uses a fully balanced signal path and an input noise figure of 7 dB. It drains 27 mA from 5V in active mode and 10 mA in standby mode.

Table 3.1
Summary of Proposed Digital Cellular Telephone Standards

Specification	*GSM Europe*	*NADC North America*	*J-PHP Japan*	*CT-2 Europe, Asia*	*DECT Europe*
Downlink frequency band	935–960 MHz	869–894 MHz	1.9 GHz	864–868 MHz	1.88–1.9 GHz
Uplink frequency band	890–915 MHz	824–849 MHz	1.9 GHz	864–868 MHz	1.88–1.9 GHz
Multiple access method	TDMA	TDMA	TDMA/TDD	FDMA/TDD	TDMA/TDD
Modulation	GMSK	π/4-DQPSK	π/4-QPSK	B-FSK	GMSK
Speech data rate	13 Kbps	8 Kbps	32 Kbps	32 Kbps	32 Kbps
Handset output power	3.7 mW → 1W	2.2 mW → 6W	10 mW	1 mW → 10 mW	250 mW
Modulation rate	271 Kbps	49 Kbps	384 Kbps	72 Kbps	32 Kbps
Channel spacing	200 kHz	30 kHz	300 kHz	100 kHz	1.762 MHz
Burst length	156 bits	324 bits			424 bits

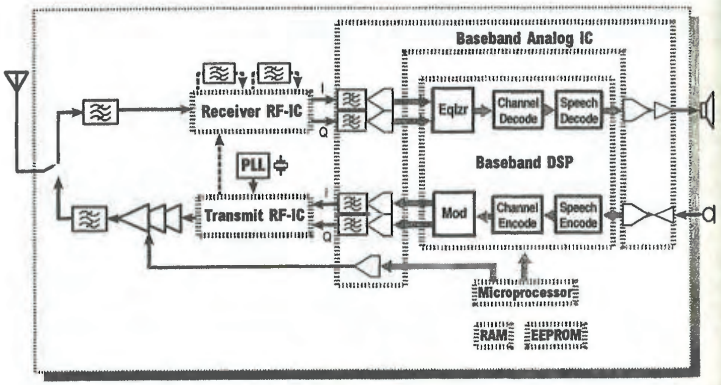
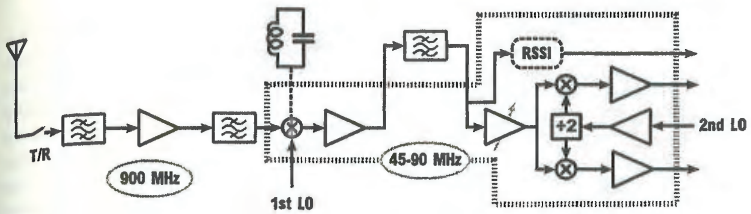
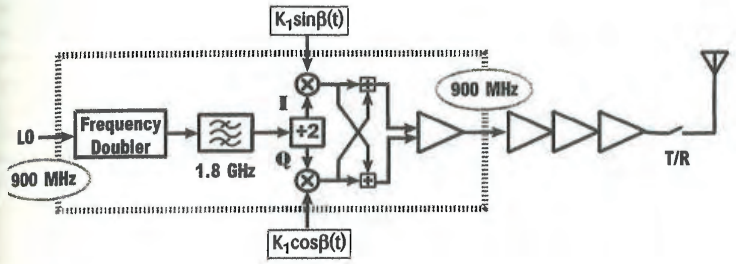


Figure 3.13 Block diagram of digital cellular telephone. Chip partitions reflect current state-of-the-art technology.



Receiver IC



Transmitter IC

Figure 3.14 Siemens receiver and transmitter ICs for GSM handsets. Receiver is double-superheterodyne architecture. Requires one off-chip tuned circuit, SAW resonator at first IF, and low-noise amplifier. Transmits doubles LO frequency and then divides by 2 to accurately obtain quadrature phases.

In the transmitter IC, a precision quadrature upconversion mixer produces a single-sideband, suppressed-carrier QPSK output from a baseband vector input. An RF differential input amplifier selects one sideband from the upconverted outputs in the two arms of the mixer. As the other sideband would occupy a nearby channel's spectrum, it must be adequately suppressed (relative to the level of the wanted sideband). Imperfect cancellation of the unwanted sideband arises from gain mismatch in the I and Q channels of the upconversion mixer, and in deviations from quadrature in the two LO outputs. The goal in this transmitter was quadrature phase errors of less than 2 deg, which with adequate gain matching implies unwanted sideband suppression of at least 40 dB relative to the wanted sideband. The stringent GSM requirements on low LO phase noise are only met with a 900-MHz off-chip

coaxial-resonator-based, varactor tuned Colpitts oscillator [101], which inherently provides a single-phase output. There are several methods to accurately derive quadrature phases. Siemens uses the double-frequency method. An on-chip nonlinear element doubles the oscillator frequency to 1.8 GHz, following which an on-chip bandpass filter removes harmonics. Then, the oscillation is divided by 2 by positive- and negative-edge triggered flip-flops, which will yield quadrature outputs with a phase accuracy set by how close the duty cycle of double-frequency oscillation is to 50%. Combined with an output stage to drive an off-chip power amplifier module, the IC drains 40 mA. Using a similar superheterodyne architecture with 71-MHz IF, AT&T Microelectronics has recently combined the GSM receiver and transmitter blocks, including the frequency synthesizer, onto one chip [102].

In an effort toward even greater miniaturization at Alcatel, the transmitter and receiver sections are both integrated onto one 9-GHz f_T silicon bipolar chip (Figure 3.15) [103, 104]. From the point of view of integration, the main advantage of the zero IF direct-conversion architecture is that it eliminates the IF passive high-frequency bandpass filters. The disadvantage, on the other hand, is that the downconverted signal has energy at dc, to which will add the receiver's offsets and low-frequency noise. An off-chip low-noise amplifier drives directly into quadrature (I-Q) mixers. The mixer dynamic range is wide enough so that a large blocking signal only 3-MHz away from the desired signal produces insignificant intermodulation distortion. The bipolar Gilbert-cell mixer is linearized with emitter degeneration

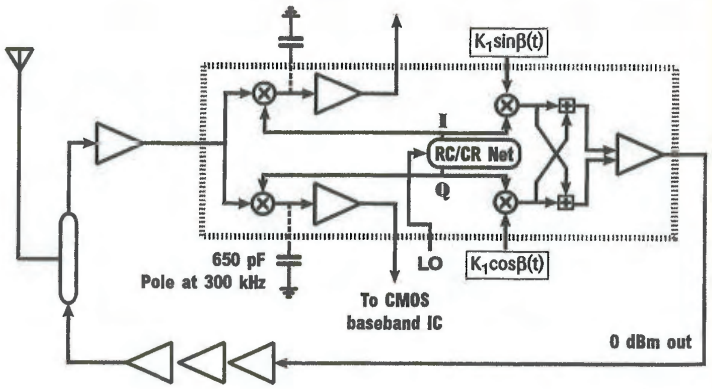


Figure 3.15 Alcatel direct-conversation single-chip transceiver. Channel selection done by switched-capacitor lowpass filters in companion CMOS mixed-signal baseband IC. On-chip, trimmed RC-CR network generates quadrature phases for mixers.

resistors, which also degrade its noise figure. The receiver section requires large (650 pF) off-chip capacitors at its output for antialias lowpass filters before the baseband signal is sent to a companion mixed-signal CMOS IC. This baseband CMOS IC contains a high-order, switched-capacitor lowpass filter for channel selection; a digital GMSK detector; and various DSP functions [105]. An RC and CR network with off-chip trimming shifts the phase of an external LO by precisely ±45 deg to produce the quadrature drive to the mixers. Upconversion mixers in the transmitter drive a power amplifier module. The chip drains 25 mA from 5V in the receive mode and 45 mA in the transmit mode, rather comparable figures to the previously described Siemens chipset. The transmitter suppresses the unwanted sideband and carrier by about 40 dB.

There is increasing interest in the 1.9-GHz band for digital cellular telephony. All the current approaches (Table 3.1) use time-division duplexed (TDD) receive and transmit frames, wherein users are assigned different timeslots. Two-way communications take place over the same frequency band. A host of ICs are appearing to serve the European DECT standard. The majority cater to a superheterodyne receive architecture with 110-MHz IF [106–109]; some of the literature [107, 109] also describes procedures for system design. The RF sections are usually integrated at small scales, embodying, say, the low-noise amplifier and first mixer on one chip, the quadrature modulator on another, and the exciter and power amplifier on a third and fourth chip. RF-ICs operating at 1.9 to 2.5 GHz exist in both GaAs [68, 110–112] and silicon [106, 113–115] technologies. Some of these stand-alone GaAs IC low-noise amplifiers achieve impressive gain and noise figure with 1- to 2-mA current drain from 3V [69, 116]. At the system level, though, functionality and overall dynamic range in these short-haul wireless links takes precedence over raw component performance. Thus, system-level input noise figures of 10 to 15 dB are acceptable [109], as are input-referred intercept points of -16 dBm and a 20-dB power amplifier control range [112]. The less than stringent system specifications lead to simplifications in the transmit path, such as direct VCO modulation by the baseband signal by opening the transmit PLL over the duration of the transmit-frame [107].

The power amplifier continues to be built almost exclusively as a separate GaAs IC, and at 1.9 GHz the on-chip wavelength is short enough that it is now possible to integrate distributed matching networks [68, 117, 118]. Efficiencies of 50% are attained when delivering almost 1W to the load from 3V. Either the transmitter or the receiver is active in TDD digital transceivers, so fast-switching power-down modes are designed into the various components. In particular, the switching trajectory must be shaped to suppress spurious emissions when the power amplifier is switched on and off [85]. Finally, the transmit and receive signals are directed to the antenna via a three-port passive circulator, or preferably through a low-loss, monolithic transmit/receive (T/R) RF switch [112, 119, 120]. A good FET switch must not appreciably distort the RF signal, or incur more than 1-dB insertion

loss when on, yet when off it should offer at least 30-dB isolation. Microcell applications, such as the Japanese Handy Phone, require an average output power of only 10 mW (20 dBm), although the $\pi/4$-QPSK modulation requires the power amplifier to handle larger peak powers. An exciter amplifier, power amplifier, and T/R switch have been integrated together [121]. The exciter and power amplifier attain a 44% efficiency at 23-dBm output from a 4.8V supply, and the on-chip matching network uses LC lumped elements. Although matching networks will eliminate harmonic distortion, they are ineffective in suppressing near-carrier spurii produced by intermodulation distortion. Feedback linearization techniques have been proposed that predistort the digitally synthesized exciter input waveform to anticipate the power amplifier nonlinearities [122]. Use of these techniques make a highly nonlinear but efficient power amplifier appear linear. These remain at the experimental stage today, and with the move to microcells and low emitted power levels, they may soon not be necessary in handsets.

The building-block ICs described so far are important advances in realizing small, relatively low-power transceivers, but the ultimate goal remains to integrate the entire transceiver on to a single chip. To this end, some early breadboard-level experiments show that direct-conversion receiver architectures seem well-suited to the DECT application [123, 124]. A 16-GHz f_T silicon bipolar IC from Alcatel operating at 1.9 GHz contains a complete direct-conversion DECT transceiver [125]. The architecture resembles the previously described Alcatel direct-conversion GSM transceiver [103], requiring a separate low-noise amplifier in the receive path, a power amplifier module, and a baseband CMOS mixed-signal signal processor IC. From a 5V supply, the transceiver drains 50 mA in receive mode and 80 mA in transmit mode. One of the challenges in realizing this system was a fast-switching PLL frequency synthesizer with low phase noise and low spurious output levels, which was built here with an improved charge pump and loop filter.

Siemens has extended its dual-conversion GSM transmitter/receiver chip set [99, 126] to a generalized front-end for digital cellular telephones operating in any part of the RF spectrum from 800 MHz to 2.1 GHz [127]. The 25-GHz f_T silicon bipolar chipset operates at a supply as low as 2.7V, the transmitter chip draining 60 mA and the receiver 33 mA. Both have power-down modes. An external power amplifier module is required. A notable low-voltage circuit on this chip is a modified Gilbert-cell mixer, with resistors instead of current sources to the negative supply [31]. To accommodate a wide input dynamic range, the gain of the RF low-noise amplifier in the receiver may be switched from −5 dBm to +15 dBm, a powerful technique that has also been used in paging receivers [46]. This assumes that the receiver does not instantaneously require a very wide dynamic range, but either receives mostly strong signals or mostly weak ones. The downconversion mixer, with a 13-dB noise figure and −3-dBm input IP3, is followed by an IF variable-gain amplifier with an 80-dB digitally programmable range [126]. The on-chip RF oscillator requires an off-chip resonator and varactor, and its phase noise at 2-kHz offset is −88 dBc/Hz. A balanced signal path is used throughout. One good measure of the quality of on-chip gain matching in the I and Q paths in the direct upconverter, as well as of the quadrature accuracy of the upconversion clocks, is given by the relative levels of spurious emissions from the transmitter IC. The unwanted sideband is suppressed by 48 dB, while the RF carrier tone, emitted due to dc offsets in the two paths, is 37-dB down. A third-order modulation tone at −46 dB appears due to mixer nonlinearity. The absolute output noise level at a 25-MHz offset is −141 dBm/Hz.

3.7 DIRECT-CONVERSION TRANSCEIVERS

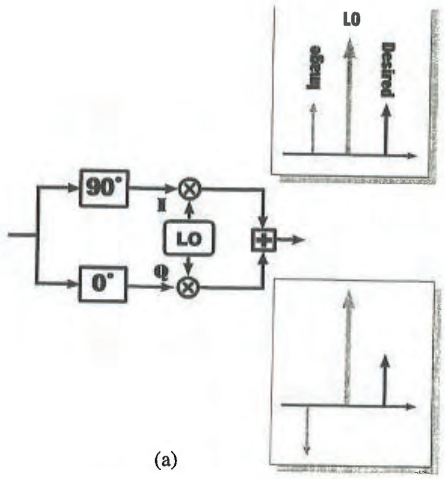
A receiver with zero IF is called a *direct-conversion* receiver. When the LO is synchronized in-phase with the incoming carrier frequency, this is also referred to as the *homodyne*. This architecture is sufficiently promising for single-chip transceivers to warrant a separate section to its study.

The desired channel is translated by the first mixer to a 0-Hz center frequency, and instead of adjacent channel rejection with a bandpass resonator, a more flexible and easier to implement lowpass filter is required—in effect, a bandpass filter centered at dc when the negative frequency axis is included. With zero IF, there is no image frequency. As early as 1924, radio pioneers had considered homodyne architectures for crude receivers requiring only a single vacuum tube, but it was in 1947 that a homodyne was first used to full effect, with a high-order, lowpass filter for channel selection in a measuring instrument for carrier-based telephony [32].

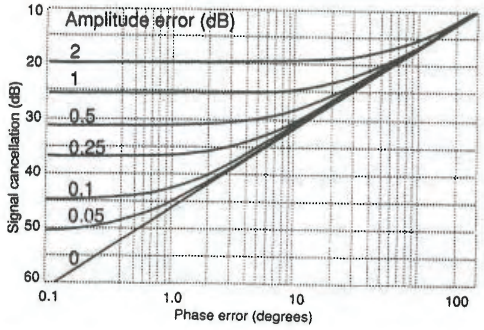
3.7.1 Direct-Conversion Single-Sideband Synthesizers

For reasons of spectral efficiency, the transmitted signal in digital communications is always single sideband with suppressed carrier. This is most often produced with the so-called *phasing method* [128]. The modulated signal is first synthesized in quadrature at baseband, *directly upconverted* into two paths by quadrature LO centered at the carrier frequency, and added or subtracted to select either the upper or lower sideband (Figure 3.16(a)).

The unwanted sideband is suppressed to an extent limited by the gain mismatch in the two upconversion paths, and by departures from quadrature in the two LO outputs (Figure 3.16(b)). Unequal dc offsets in the paths produce an output signal at the LO frequency. The unwanted sideband and LO leakage are spurious but unavoidable components of the transmitted spectrum. Although on the same IC the two upconversion paths are well-matched, a gain mismatch as small as 1% (0.1 dB) will limit suppression of the unwanted sideband to 45 dB. With this gain mismatch, a phase-error of up to 1 deg may be tolerated between the two LO outputs. These mismatches may be trimmed down at time of transceiver manufacture, or self-



(a)



(b)

Figure 3.16 (a) A single-sideband direct upconverter using the phasing method (also called the quadrature modulator). (b) Extent of suppression of unwanted (image) sideband vs. gain mismatch and phase errors from quadrature in two arms of modulator.

calibrated with loopback modes that are activated during idle times to sense and suppress the unwanted spurii.

An off-chip resonator, often connected to an on-chip unbalanced oscillator circuit, may also become a cause of spurious RF leakage if it couples energy into the power amplifier or the antenna. Frequency-offset upconversion schemes have been proposed [129] to combat this coupling problem. Other spurious output tones may arise from parasitic remixing of the modulated output with the baseband signal and by intermodulation distortion in the output stage [96]. Balanced circuit topologies, on-chip LOs requiring no external resonators [130], and lowered transmit power levels in microcells are all expected to lessen the magnitude of the spurious leakage problem.

Cellular wireless systems operate best with tightly regulated power control, the handsets transmitting only the minimum power required for reliable reception by the base station, and vice versa. Whereas in the receiver, the front-end is the main source of noise, in a transmitter, the upconversion LO phase-noise appears as added noise on the emitted signal. Direct upconversion has the advantage over two-step schemes in that only *one* LO contributes noise.

3.7.2 Direct-Conversion Receivers for GSM Digital Cellular Telephones

Among the various RF-IC suppliers for the European GSM digital cellular telephone handset, only Alcatel at present uses a direct-conversion receiver architecture [103, 125, 131]. This results in a relatively small silicon bipolar RF front-end chip, and the remainder of the signal processing, including lowpass channel-select filtering, is at baseband in mixed-mode CMOS. Why are the others reluctant to use direct conversion in their receivers? Given the many decades of superheterodyne experience, the most likely reason is conservatism. But direct conversion also suffers from some unique problems.

A well-known problem is that spurious LO leakage from the receiver into the antenna becomes an in-band interferer to other nearby receivers tuned to the same band. Superheterodynes, with their frequency-offset LOs, do not suffer from this problem. However, experimental studies suggest that with standard shielding in the receiver, this problem is not so severe as to handicap the use of direct conversion [132].

A more serious problem is *dc offset* in the receiver. Offset arises from three sources [131]: transistor mismatch in the signal path; the LO signal leaking to the antenna because of poor reverse isolation through the mixer and RF amplifier, then reflecting off the antenna and self-downconverting to dc through the mixer (Figure 3.17); and a large near-channel interferer leaking into the LO port of the mixer, then self-downconverting to dc. Good circuit design may reduce these effects to a certain extent, but they cannot be eliminated.

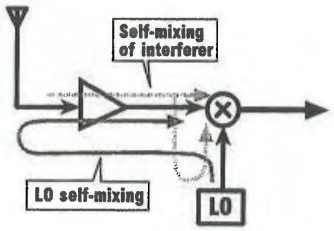


Figure 3.17 Sources of leakage in a direct-conversion receiver, and how through self-mixing they create dc offsets.

The spectrum of the GMSK modulation used in GSM has a peak at dc. Offsets will directly add to the spectral peak of the downconverted signal. These offsets are usually much larger than the rms front-end noise, and may therefore significantly degrade the SNR at the detector. To remove the offsets by ac coupling, the receiver will require impractically large capacitors if the signal-bearing spectrum around dc is not to be sacrificed. However, they may be compensated with DSP-based self-calibration [105, 133]. The baseband signal processor makes long-term measurements on the dc level in the receiver and subtracts off this level from the downconverted signal. The spectrum loss around dc is only a few hertz, and digital filtering does not distort the midchannel group delay in the receiver. Among the various RF-IC suppliers for the European GSM digital cellular system, Alcatel has demonstrated the feasibility of highly integrated direct-conversion receivers [103, 125, 131] where others use superheterodyne architectures. Alcatel's motivations are, predictably, that the channel filters may be lowpass and therefore active; a minimal silicon bipolar RF front-end is required; and that all baseband signal processing may be in CMOS.

3.7.3 Local Oscillators With Quadrature Outputs

A *carrier-frequency* LO with quadrature outputs is a key circuit component in direct-conversion transmitters and receivers. Usually, this oscillator is tuned by an LC circuit or an off-chip resonator and inherently produces a single-phase output. Quadrature phases are often derived by passing the oscillator output through a CR and an RC network, whose time constant is equal to the oscillation period. The two resultant outputs are then phase-shifted by +45 deg and −45 deg, respectively. Inaccuracies in the actual values of R and C will lead to errors in quadrature and are compensated by some form of on-chip trimming [103, 125].

An alternative method is to divide an oscillation at twice the carrier frequency with positive- and negative-edge triggered flip-flops [99]. The resulting two outputs at the desired frequency are in quadrature. Residual phase errors caused by unequal delays in the two flip-flops and by departures from 50% duty cycle in the double-frequency oscillation are found to be less than 1 deg. However, the double-frequency portions of the circuit may become a speed bottleneck.

Quadrature outputs may also be derived from a polyphase oscillator, such as a variable-frequency ring oscillator. In a four-delay stage ring oscillator, taps at diametrically opposite points will yield quadrature phases at all frequencies [134] (Figure 3.18). In oscillators with an odd number of unit delays, they may be synthesized from two taps by a voltage-controlled phase shifter in feedback around a quadrature-sensing circuit [135]. Mismatches in transistor characteristics limit the attainable phase accuracy. This type of resonatorless oscillator is practical when the specifications on phase noise are not too stringent (say up to −80 dBc/Hz at a 100-kHz offset from the carrier). The free-running oscillator must be slaved to a crystal reference in a frequency-locked loop, whose loop gain and bandwidth are specifically designed to suppress phase noise.

When circuit techniques cannot contribute further improvements to gain- and phase-matching in the two arms of a quadrature modulator or demodulator, digital



Figure 3.18 A four-phase voltage-controlled ring oscillator. Diametrically opposite taps give quadrature phase at any frequency. Unequal loadings and device mismatch set phase errors. VCO must be embedded in a PLL whose closed-loop bandwidth is tailored to suppress oscillator noise.

adaptive algorithms may be used at baseband to sense and compensate for these errors [136].

3.8 ICs FOR SPREAD-SPECTRUM WIRELESS TRANSCEIVERS

Spread-spectrum communications offer greater user capacity than narrowband techniques in a given piece of wireless spectrum. Much has been written on this subject [137]. Spread-spectrum techniques were developed during World War II as a form of secure communication with low probability of intercept and resilience to jamming [138]. Over the years, this technology has been further developed and refined for military communications. With the wireless revolution at hand and the IC technology now available to implement complex transceivers, spread spectrum has awakened commercial interest [139]. Spread-spectrum communications make it easy for users to access the wireless channel. Whereas conventional narrowband wireless communications require a careful discipline, which the FCC or other government agencies enforce by issuing licenses so as to prevent use of the same frequency by nearby operators, for spread-spectrum use, the FCC has allocated certain *unlicensed* bands in the United States, referred to as the instrumentation, scientific, and medical (ISM) bands, wherein the user is only required to spread spectrum by a minimum amount, and not to exceed an upper limit on transmitted power. The extent of spectrum spreading may be measured by the amount of *processing gain* in the receiver required to despread and detect the signal [140].

As analog cellular telephones begin to saturate the available RF allocations, spread-spectrum techniques are being deployed in cordless telephones and wireless modems [141]. These devices usually operate in the two lower ISM bands: 902 to 928 MHz, and 2.4 to 2.48 GHz. A ubiquitous wireless environment is envisioned, in which mobile users, wherever they may be, can access data and communications services through an intricate network of base stations [2]. The greatest hardware challenge in realizing this scenario is the development of a low-power miniature handset.

There are two different methods to spread the spectrum of a signal: by *direct-sequence* modulation or by *frequency hopping* [140]. Direct sequence is conceptually the simpler, as well as the more straightforward to implement. Each data bit (either +1 or −1) at the transmitter multiplies a prescribed sequence of bits, or *chips* as they are called. The chip sequence is selected for very low autocorrelation, and is often referred to as a pseudo-noise (PN) sequence. Each user is either assigned a unique PN sequence with very low cross-correlation with other user sequences, or a time-shifted version of some long PN sequence. The receiver, after an initial acquisition search to align itself with the start of the PN sequence, correlates the incoming sequence with the pattern it knows to be its own. On detection of a correlation peak, the sign of the peak signals the source data bit. The longer the PN sequence

for each bit, the greater the spreading factor, and the more reliable the detection. Here, the processing gain is the sequence length per bit. The spectrum of the transmitted data, composed from a concatenation of PN sequences, is noise-like; thus, the term *spread-spectrum*. Several baseband and IF modem ICs have already appeared to support BPSK and QPSK direct-sequence spread-spectrum transceivers [142–144].

Whereas direct-sequence modulation spreads the spectrum by randomizing the waveform, frequency-hopping spreads spectrum by *hopping* the carrier according to a prescribed sequence across the entire band. The effect in the frequency domain is similar to direct sequence, except for one fundamental difference: with direct sequence, spreading across a wide bandwidth requires each data bit to be mapped into a long PN sequence, and the resultant high "chip-rate" requires high-speed signal processing at the receiver front-end. On the other hand, if the carrier frequency is hopped with a wideband synthesizer, the output spectrum may span an arbitrarily wide frequency range, even at low bit rates. Therefore, if a fast and agile frequency synthesizer is readily available, the receiver front-end operates at the actual data rate frequency-hopped spread spectrum, rather than at the considerably higher chip-rate in a direct-sequence receiver. The former will very likely lead to a low-power solution.

As a notable recent example of a miniature spread-spectrum transceiver, Plessey has introduced a 700-Kbps frequency-hopped device operating in the 2.4-GHz ISM band. The transceiver is entirely contained on a 2-in × 3-in PCMCIA card for insertion into notebook computers [109, 145]. Data FSK modulates the carrier frequency, and a variable-modulus PLL synthesizer slowly hops the carrier to spread the spectrum across the 80-MHz band. All the active devices in the transceiver are on three ICs, consisting of a GaAs RF front-end, a silicon bipolar IF receiver, and a CMOS IC for the hopping-frequency synthesis. As in any spread-spectrum, two-way communication system, transmission and reception is time-division duplexed on the same frequency band. The receiver architecture is a conventional double superheterodyne. In addition, the transceiver requires 50 passive components, including 6 rather bulky filters, and when transmitting 100-mW RF power it dissipates more than 1W.

By the norms of GaAs MMICs, the single-chip IC front-end in this transceiver is highly integrated (Figure 3.19(a)) [146]. It includes the power amplifier and drivers for 2-GHz passive filters in the transmit and receive paths. Interesting features are the dc series connection of the single-ended two-stage low-noise amplifier, which shares the same bias current in both stages through a bypass inductor; the Clapp VCO; and the four-FET switched mixers in the transmit and receive paths (Figure 3.19(b)). In its first version, the IC includes more than 20 on-chip spiral inductors, perhaps the largest number on any MMIC at the time of this design. A second bipolar IC processes the 350-MHz IF signal. In receive mode, the GaAs front-end IC drains 30 mA from a 5V supply. The receiver selects one of two external antennas through an on-chip RF switch to attain spatial diversity. This work has spurred a

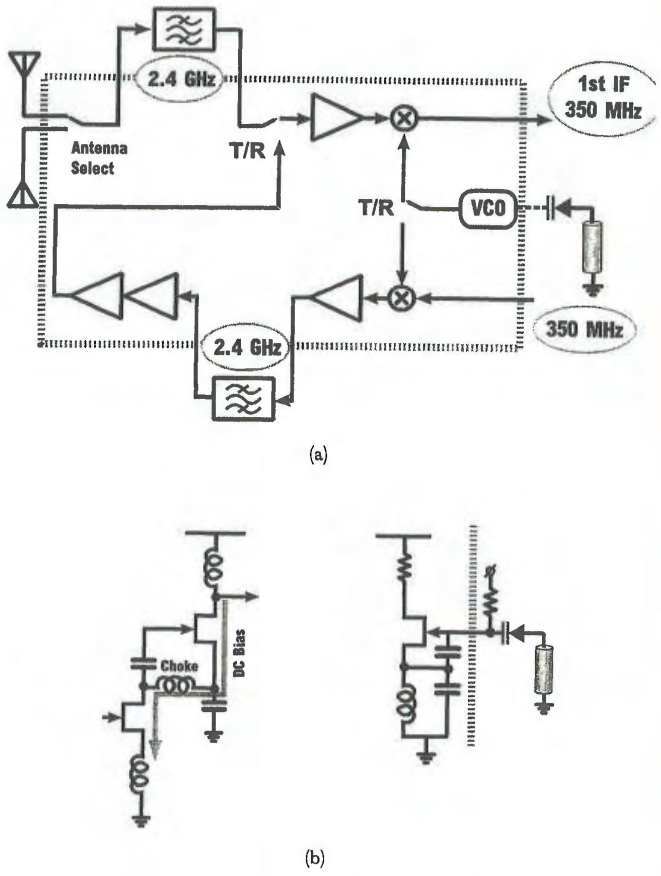


Figure 3.19 (a) Front-end GaAs MMIC transceiver for spread-spectrum communications in 2.4-GHz band. Receiver is double-superheterodyne. Note use of off-chip filters. (b) Low-noise amplifier in transceiver IF (left) reuses bias current in two stages. Clapp voltage-controlled local oscillator (right) combines on-chip LC network with off-chip resonators and varactors.

flurry of similar GaAs ICs at the same level of integration [147–149], all operating at ±5V supplies, with similar functionality and performance. One of them [149] uses a very high IF of 915 MHz, so that spurious products at the transmitter output lie well in the stopbands of the output filter, and where the IF signal is processed by a 915-MHz cellular-telephone-type IC.

3.9 ICs TO ENABLE FUTURE TRANSCEIVERS

What will portable wireless communicators of the future look like? We may draw some conclusions from the foregoing summary of RF-IC developments. The crowded spectrum means that future wireless communications will predominantly use spread-spectrum techniques. Coupled to this is a need for low-power dissipation, which will force *architectural innovations* and *higher levels of integration* in the electronics. As an example of such an advanced transceiver, the author, with his colleagues and their graduate students at UCLA, is investigating the architecture and circuit design of a frequency-hopped, binary frequency-shift keyed, zero-IF, all-CMOS two-chip transceiver capable of delivering up to 160 Kbps (the base ISDN rate) in the 900-MHz ISM band [39, 150]. The transceiver architecture is inspired by the modern paging receiver, a very low energy wireless device widely used today. The transceiver implementation (Figure 3.20) will freely mix analog and digital circuits, which makes CMOS the IC technology of choice. Further, to avoid the routing of high-frequency signals off- and on-chip and thereby save the power the buffers would use to drive stray capacitance and off-chip low-impedance lines, it is preferred to integrate all the blocks, *including* the RF front-end. Finally, it is most desirable to use an unmodified, standard production CMOS process.

Low-power operation requires the entire system to operate on a 3V supply. This supply voltage cannot be any lower because in many places the analog circuits contain stacked cascade transistors and the FET threshold voltages are about 0.8V. RF amplifiers are normally tuned with inductor loads, which on silicon ICs are almost always off-chip passive components [151]. However, this is itself wasteful of power, because in addition to the current supplied into the inductor, the circuit must also drive the various parasitic capacitances in off-chip routing the RF signal.

Spiral inductors on silicon substrates suffer from a large capacitance to the conducting silicon substrate. However, after many years of the belief that no useful inductors could be made on silicon, it was found that spiral inductors as large as 10 nH with self-resonance beyond 2 GHz could be fabricated with the standard interconnect metallization. These inductors were used to build passive filters, a tuned amplifier, and an LC voltage-controlled oscillator on a silicon bipolar IC [152–154]. This work has rekindled interest in the design and modeling of small-value spiral inductors on silicon (usually bipolar) ICs, with values in the range of 1 to 10 nH [155, 156]. In some instances, higher inductor Qs are obtained with thicker

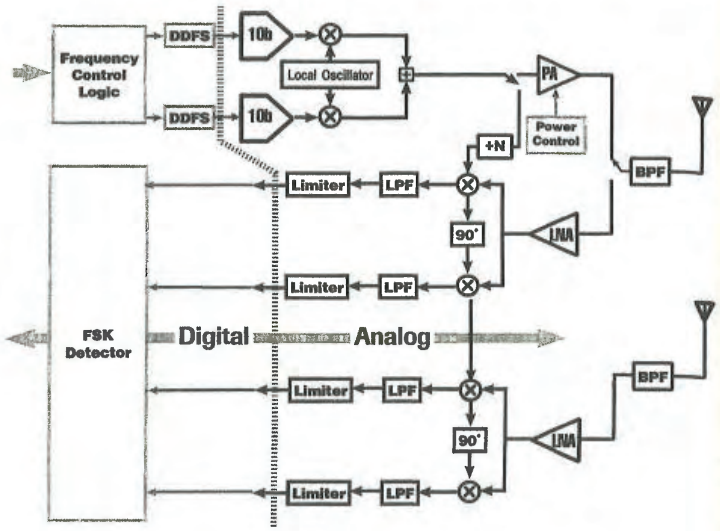


Figure 3.20 The UCLA frequency-hopped spread-spectrum transceiver. This uses direct conversion in both transmit and receive paths, and binary-FSK modulation. In the transmitter, the spread-spectrum waveform is synthesized at baseband and upconverted to RF. During receive mode, the DDFS is under control of a synchronization loop hops with the incoming data, and demodulates it to baseband. The entire mixed analog-digital, baseband/RF system is to be integrated on a 1-μm CMOS IC.

metallization, or higher resonant frequencies with thicker oxides. In search of larger value inductors for low-power, high-gain amplifiers, we have developed a fabrication method whereby the silicon substrate is selectively removed under the inductor, substantially reducing the capacitance to the substrate and thus extending self-resonance to a higher frequency (Figure 3.21(a)) [157]. The self-resonant frequency is now almost as high as it is on a semi-insulating GaAs substrate. Using two such inductors as loads in a balanced circuit, a 900-MHz RF amplifier has been built in 1-μm CMOS with a 30-dB gain draining only 3 mA (Figure 3.21(b)). The amplifier IP3 is about 0 dBm owing to the wide linear range of the MOS differential pair.

The front-end mixer in a direct-conversion receiver must be highly linear to suppress unwanted intermodulation produced by interferers. Also, as the local oscillator frequency is centered at RF in this direct-conversion receiver, inadequate reverse isolation or shielding may cause this frequency to leak through the antenna and to

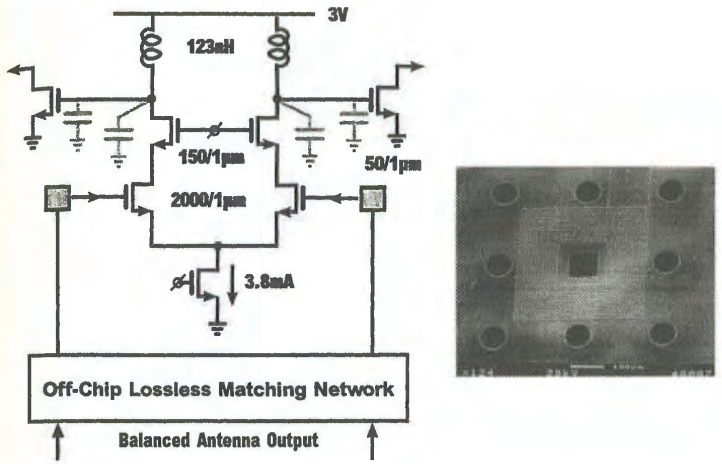


Figure 3.21 Balanced input CMOS RF amplifier. A cascode stage driving an on-chip inductor substantially enhances voltage gain. Outputs common-source stages are for measurement only. LNA will directly drive mixer. A microphotograph of a 100-nH spiral inductor suspended over a pit on substrate.

interfere with other nearby receivers tuned to the same RF. We use an unusual mixer—one that happens to be particularly well-suited to MOS implementation—to circumvent both these problems. The desired signal is downconverted to baseband by subsampling the incoming RF [158]. The mixer circuit is a track-and-hold, with such a wide track-mode bandwidth that it can follow the RF waveform, and a short enough aperture to acquire the *instantaneous* value of the RF waveform on receipt of the sampling clock edge (Figure 3.22). The received RF in our system is a 26-MHz wide spread-spectrum centered on a 915-MHz carrier, so by subsampling this waveform at a clock rate of 52 MHz or higher, the spectrum is translated to baseband without aliasing. However, the mixer also acquires wideband noise accompanying the signal from dc to up to the gigahertz track bandwidth, and aliases the rms noise into a bandwidth at half the sample rate, thereby raising the baseband noise spectral density by the ratio of the track bandwidth to the sample rate. Subsampling mixers therefore have a higher noise figure than conventional mixers. However, as such a mixer tends to be more linear than an analog multiplier, its dynamic range is also large, provided a high-gain, low-noise RF amplifier precedes it. The overall front-end spurious-free dynamic range is jointly set by the individual specifications of the RF amplifier and mixer. In this 1-μm CMOS prototype, the mixer circuit

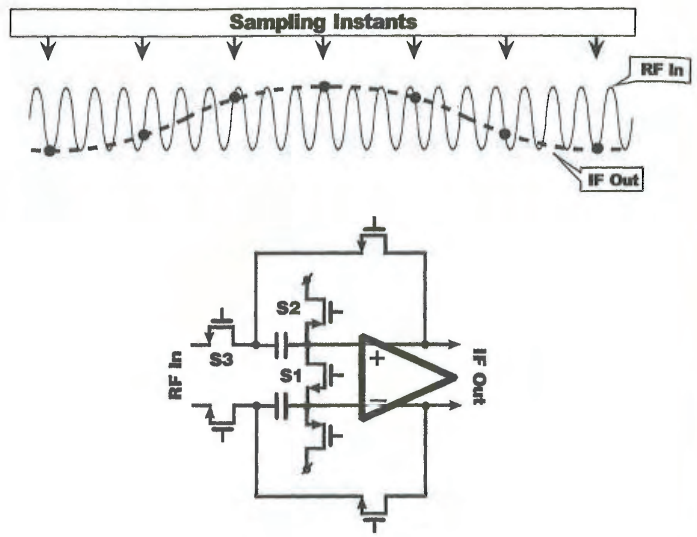


Figure 3.22 A downconversion mixer based on direct sampling of RF. During track mode, switches are closed and bandwidth of passive circuit is about 1 GHz. On receipt of sample clock, switches open within a very short aperture time, capturing instantaneous value of RF waveform. Op amp removes switch charge injection and buffers to subsequent circuits.

draws 4 mA from 3V to acquire samples of a 900-MHz modulated waveform at a 50-MHz rate, which it then translates to baseband. The linearity, as measured by a +26-dBm third-order intercept, exceeds that of most continuous-time 900-MHz monolithic mixers, and for the fundamental reasons given above, the noise figure is 18 dB.

A low-power, agile frequency synthesizer with high spectral purity is required in a frequency-hopped transceiver. Conventional synthesizers based on a PLL with a variable-modulus divider in feedback suffer from limited agility because the loop bandwidth may limit the speed of frequency switching, and it is difficult to design a VCO that both covers a wide range of frequency spreading and produces a pure spectrum. The DDFS-DAC combination is an alternative solution to this problem, which so far has only been used by the military in spread-spectrum communication systems. A direct-digital frequency synthesizer (DDFS) [159] consists of an accumulator and ROM, which produces a sequence of digital words representing samples of a sinewave at a frequency set by an input control word. A D/A converter converts

the digital samples into a discrete-time analog sinewave, and a subsequent analog filter may smooth this into a continuous-time sinewave. Round-off errors in the DDFS and nonlinearity in the DAC and filter will limit the attainable spectral purity of the synthesized sinewave [159]. In practice, DAC imperfections may contribute the largest nonlinearity in the system. A characteristic feature of digitally-synthesized sinewaves is that the major imperfections do not necessarily produce harmonic tones, but instead spurious tones at rational multiples of the fundamental. These spurious tones often cluster around the fundamental and cannot be filtered. The DAC can add spurious tones of its own. We use a highly efficient DDFS combined with a high-speed, low-power charge-redistribution DAC to produce an FSK-modulated, frequency-hopped spread spectrum at baseband, which a fixed-frequency local oscillator upconverts to RF. A CMOS implementation of a DDFS-DAC at 3V dissipates only 40 mW when clocked at 50 MHz (of which the quadrature ROM-accumulator accounts for 35 mW), with worst-case spurious levels of –57 dBc or less across the entire spreading range (Figure 3.23) [160].

The 915-MHz LO is a four-stage MOS ring oscillator locked in a PLL to a lower frequency crystal reference. The quadrature outputs at 915 MHz for the image-reject mixers are tapped at diametrically opposite points [130, 134] to a phase accuracy of better than 1 deg. As the LO operates at a fixed frequency, it is locked to a low-frequency crystal reference in a PLL solely optimized to reduce phase noise. At a 100-kHz offset, a phase noise level of –75 dBc/Hz is measured on a prototype.

The power amplifier is a binary-weighted array of FETs biased near threshold, which is digitally selected to deliver power levels from –15 dBm to +15 dBm through a matching network to the antenna with a 40% conversion efficiency (Figure 3.24) [161].

The transceiver achieves spatial diversity with two individual receive channels, which are powered up continuously and connected to two different types of miniature antennas. Data decisions are made on an equally weighted sum of their baseband outputs. A high-order switched-capacitor elliptic lowpass filter selects the desired channel, which the onboard DDFS-based synthesizer course has dehopped to dc in the course of downconversion. The capacitively-coupled output is then amplitude-limited to produce a one-bit stream fed to an optimal FSK digital detector [39]. The detector quantizes the zero crossings in time by oversampling the limited baseband data and seeks correlations between one-bit representations of the two possible data values. A logarithmic amplifier connected to taps on the limiting amplifier chain measures the received signal strength [162]. Operating from a 3V supply, the 1-μm CMOS transceiver is expected to drain 70 mA in receive mode, and 100 mA in transmit mode (the power amplifier is on-chip). Considering that the transceiver communicates on 26-MHz-wide spectrum, this is indeed a low-power dissipation.

This research project is now exploring issues related to the single-chip integration of this transceiver. Such a highly-integrated CMOS "VLSI radio" would repre-

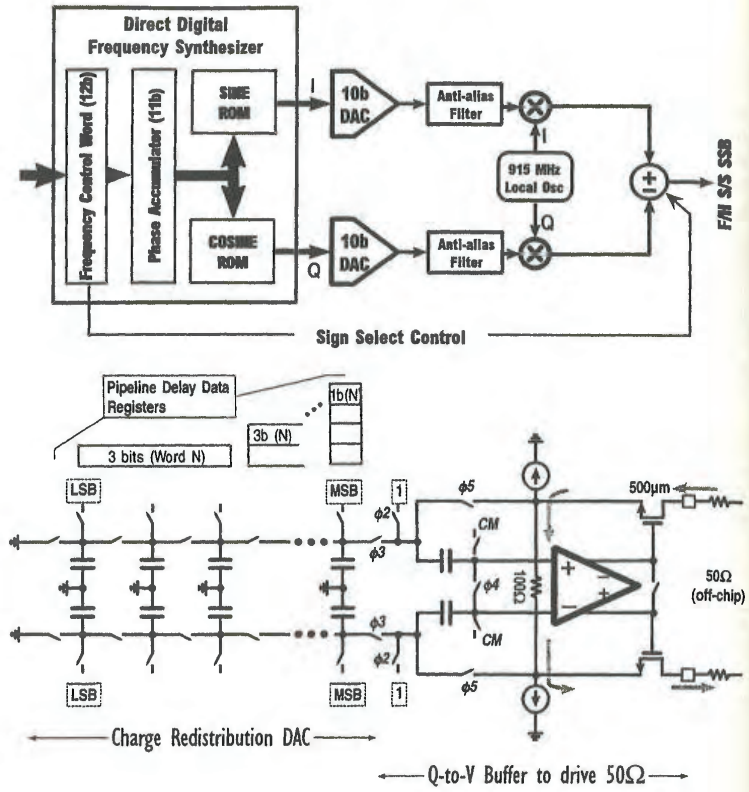


Figure 3.23 An agile frequency synthesizer. The DDFS produces digital samples of a sine and cosine at a frequency set by an input word. DACs convert these to analog domain, and after quadrature upconversion, either the upper or lower sideband is selected. Thus, 0- to 13-MHz baseband frequency synthesis results in output spanning 902 to 928 MHz. A first-order antialias filter suffices if DDFS/DAC operate at 80 MHz. DAC uses passive pipelined architecture.

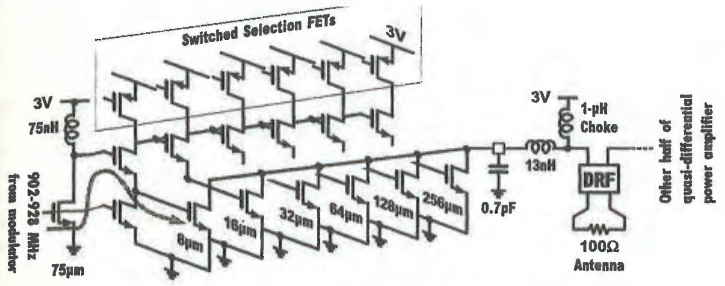


Figure 3.24 A digitally controlled power amplifier. The binary-weighted array of output NMOSFETs drive the antenna load differentially, to deliver up to +15 dBm from a 3V supply. A large on-chip inductor used as an exciter load gives swings above power supply. PMOSFETs selectively activate source follower buffers, and thus total delivered power.

sent a major step forward in the evolution of the radio integrated circuits described in this chapter.

3.10 CONCLUSIONS

This survey has covered some of the key existing and emerging communication applications that have prompted advances in RF-ICs. The emphasis was on widespread *portable consumer* applications. This has excluded coverage of integrated television tuners, for instance, which were historically the principal motivation for the development of UHF RF-ICs by the consumer electronics industry in the mid-1970s [163]. The fact that about 18 million tuners are produced every year means that there has been a steady stream of innovations in the building blocks and architectures for this application [164–178]. Then there are emerging areas, such as personal GPS receivers, where RF integration and low power will be the key for acceptance by consumers [179, 180]. Here, too, RF-IC designers are responding with integrated front-ends [181–184], although the greater challenges to realize a system with an acceptable precision may lie in the baseband processing.

The lucrative global wireless market is attracting a great deal of attention from circuit designers across the industry. The RF range of interest spans 400 MHz to 2.5 GHz. There is a multipronged industrywide assault to provide the right solutions. The solution must be low cost and low power, and it must give high functionality. MMIC designers bring to RF-ICs a microwave style derived from small-scale circuits that operated at much higher frequencies. On the other hand, as the frequency range lies within the capability of modern silicon IC technologies, there is a response from

this community. Bipolar and BiCMOS RF-ICs are often extensions of baseband-style circuits to high-frequency, often with more functionality than GaAs MMICs offer. A low-transistor f_T is no longer a handicap, and experience rapidly accumulates on how to solve some of the unique problems of silicon substrates, such as high-frequency losses in the substrate [185] and on-chip coupling problems between circuit blocks. In the not too distant future, as microprocessors and memories drive the linewidths of silicon IC processes to deep submicron, RF-ICs might be designed as untuned wideband circuits, in the style of ICs for video applications today.

Small-scale GaAs MMICs might sometimes perform better as stand-alone components, but embedded in a system such as a digital cellular telephone or a spread-spectrum transceiver, they require a large overhead in support circuits, and thereby lose an edge in certain specifications, such as power dissipation. Silicon RF-ICs offer higher levels of integration, but even within the silicon IC design community there are differing points of view. While integrated silicon bipolar transceivers are now to be found in the next generation of products, CMOS RF solutions, albeit at an exploratory stage, now look very promising. They combine elements from well-established techniques from voiceband and video ICs, and are unfettered in partitioning tasks between analog and digital signal processing, while using off-chip passive components when necessary. Digital control components and baseband signal processors are readily integrated with, indeed merged into, the RF and IF sections. This array of competing technologies offers communication system designers more creative opportunity, and the best wireless transceiver solutions may well emerge from system design evolving together with architecture, circuits, antennas, and power allocation plans.

What are the prospects for this multitude of approaches to low-power wireless communications? If we project on the experience gained from baseband and video telecommunications circuits, the future wireless transceiver may be only one BiCMOS or CMOS large-scale, mixed analog-digital IC, requiring one passive filter and a crystal resonator. Baseband digital signal processing may make up for imperfections in the front-end RF sections. As it will likely be used in a microcell environment at low transmit power levels, the transceiver may be encapsulated in a plastic package. There is no fundamental reason why it should need more than a single 1.5V battery as the power source. After decades of thinking of it only as science-fiction, the electronics world is prodded forth by the communications explosion into making the two-way wristwatch-size radio a reality. A large community of researchers, not trained in the classical radio art, is busy re-examining the conventions, and at times challenging the received wisdom. We are poised at a turning point in the evolution of radio-circuit engineering.

Acknowledgment

My own involvement in RF-IC design is considerably enriched by collaborations with the many talented colleagues and graduate students participating in the Low-Power Wireless Transceiver Program in the UCLA Electrical Engineering Department. Professors Henry Samueli, Greg Pottie, and Yahya Rahmat-Samii have made an exemplary teamwork possible on this project. The program is principally supported by the U.S. Defense Advanced Research Projects Agency, with key additional support, both intellectual and financial, from a consortium of semiconductor companies under the State of California MICRO Program.

The recent financial stringencies at UCLA have not so far eroded the marvelously rich engineering collection and online facilities at the UCLA Science and Engineering Library. It would be impossible to write this chapter without ready access to these resources. Ms. Karen Andrews and Ms. Aggi Raeder, two of the expert librarians there, enthusiastically helped me in using the latest tools to peruse the literature and amazed me with the ease with which they could track down obscure references from all corners of the world.

I thank Professor Bob Meyer for a very thorough reading of the draft manuscript and for his thoughtful comments on it, and Messrs Taiwa Okanobu and Mark McDonald who made useful suggestions. I am very grateful to Dr. Ran-Hong Yan, the Guest Editor of the *Proceedings of the IEEE* who solicited this manuscript for its original publication, for his extraordinary ability to obtain speedy reviews and patiently absorb delays on my part in bringing this work to completion.

References

[1] G. H. Heilmeier, "Personal Communications: Quo Vadis," *Int'l Solid State Circuits Conf*, San Francisco, CA, 1992., pp. 24–26.

[2] D. C. Cox, "Universal Digital Portable Radio Communications," *Proc. of IEEE*, Vol. 75, No. 4, 1987, pp. 436–476.

[3] S. Sheng, A. Chandrakasan, and R. W. Brodersen, "A Portable Multimedia Terminal," *IEEE Communications, Magazine*, Vol. 30, No. 12, Dec. 1992, pp. 64–75.

[4] B. A. Beatty, "The IBM Personal Communicator—Design Considerations," *Southeastcon*, Miami, FL, 1994, pp. 432–435.

[5] A. Matsuzawa, "Low-Voltage and Low-Power Circuit Design for Mixed Analog/Digital Systems in Portable Equipment," *IEEE J. of Solid-State Circuits*, Vol. 29, No. 4, 1994, pp. 470–480.

[6] E. A. Vittoz, "Low-Power Design: Ways to Approach the Limits," *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1994, pp. 14–18.

[7] F. E. Terman, *Radio Engineers' Handbook*, New York, NY: McGraw-Hill, 1943.

[8] W. L. Everitt, ed., *Fundamentals of Radio and Electronics*, 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 1958.

[9] K. K. Clarke, and D. T. Hess, *Communication Circuits: Analysis and Design*, Reading, MA: Addison-Wesley, 1971.

[10] H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, New York: John Wiley & Sons, 1980.

[11] W. H. Hayward, *Introduction to Radio Frequency Design*, Englewood Cliffs, NJ: Prentice-Hall, 1982.

[12] U. L. Rohde, and T. T. N. Bucher, *Communications Receivers: Principles & Design*, New York, NY: McGraw-Hill, 1988.

[13] D. O Pederson, and K. Mayaram, *Analog Integrated Circuits for Communications: Principles, Simulation, and Design*, Boston, MA: Kluwer Academic, 1991.
[14] M. Ibuka, "How Sony Developed Electronics for the World Market," *IEEE Trans. on Engineering Management*, Vol. EM-22, No. 1, Winter 1975, pp. 15–19.
[15] A. Laundrie, and K. Negus, "GPS Receiver Chip Targets Low-Power Commercial Markets," *Microwaves & RF*, Aug. 1992, pp. 135–138.
[16] N. Camilleri, D. Lovelace, J. Costa, and D. Ngo, "New Development Trends for Silicon RF Device Technologies," *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, San Diego, CA, 1994, pp. 5–8.
[17] W. Kanow, and I. Siewert, "Integrated Circuits for Hi-Fi Radios and Tuners," *Electronic Components and Applications*, Vol. 4, No. 1 Nov. 1981, pp. 11–27.
[18] R. M. Warner, and J. N. Fordemwalt, ed., *Integrated Circuits: Design Principles and Fabrication*. New York, NY: McGraw-Hill, 1965.
[19] C. S. Meyer, D. K. Lynn, and D. J. Hamilton, ed., *Analysis and Design of Integrated Circuits*. New York, NY: McGraw-Hill, 1968.
[20] Y. Hishida, and M. Kazama, "Mobile Communication System: Its Semiconductor Technology," *Hitachi Review*. Vol. 42, No. 3, June 1993, pp. 95–100.
[21] B. Crouch, ed., *Dick Tracy: America's Most Famous Detective*. Secaucus, NJ: Citadel Press, 1987.
[22] L. Blaser, and T. Taira, "An AM/FM Radio Subsystem IC," *IEEE Trans. on Consumer Electronics*, Vol. CE-23, No. 2, 1977, pp. 129–136.
[23] W. Peil, and R. J. McFayden, "A Single Chip AM/FM Integrated Circuit Radio," *IEEE Trans. on Consumer Electronics*, Vol. CE-23, No. 3, 1977, pp. 424–431.
[24] O. L. Richards, "A Complete AM/FM Signal Processing System," *IEEE Trans. on Consumer Electronics*, Vol. CE-24, No. 1, 1978, pp. 34–38.
[25] W. G. Kasperkovitz, "An Integrated FM Receiver," *Microelectronics Reliability*, Vol. 21, No. 2, 1981, pp. 183–189.
[26] A. A. Abidi, "Noise in Active Resonators and the Available Dynamic Range," *IEEE Trans. on Circuits and Systems*, Vol. 39, No. 4, 1992, pp. 296–299.
[27] W. G. Kasperkovitz, "FM Receivers for Mono and Stereo on a Single Chip," *Philips Technical Review*, Vol. 41, No. 6, 1983, pp. 169–182.
[28] W.H.A. van Dooremolen, and M. Hufschmidt, "A Complete FM Radio on a Chip," *Electronic Components and Applications*. Vol. 5, No. 3, 1983, pp. 159–170.
[29] J. Brilka, and G. Sieboerger, "A Single-Chip FM Radio," *Int'l Conf. on Consumer Electronics*, 1984, pp. 156–157.
[30] T. Okanobu, T. Tsuchiya, K. Abe, and Y. Ueki, "A Complete Single-Chip AM/FM Radio Integrated Circuit," *IEEE Trans. on Consumer Electronics*, Vol. CE-28, No. 3, 1982, pp. 393–408.
[31] T. Okanobu, H. Tomiyama, and H. Arimoto, "Advanced Low-Voltage Single Chip Radio IC," *IEEE Trans. on Consumer Electronics*, Vol. 38, No. 3, 1992, pp. 465–475.
[32] D. G. Tucker, "The History of the Homodyne and Synchrodyne," *J. of the British Inst. of Radio Engineers*, Vol. 14, No. 4, 1954, pp. 143–154.
[33] D. Mitchell, and K. G. van Wynen, "A 150-mc Personal Radio Signaling System," *Bell System Technical J.*, Vol. XI, No. 5, 1961, pp. 1239–1257.
[34] A. E. Kerwien, and L. H. Steiff, "Design of a 150-megacycle Pocket Receiver for the BELLBOY Personal Signaling System," *Bell System Technical J.*, Vol. XLII, No. 3, 1963, pp. 527–565.
[35] K. Nagata, M. Akahori, T. Mori, and S. Umetsu, "Digital Display Radio Paging System," *NEC Research & Development.*, No. 68, Jan. 1983, pp. 16–23.
[36] I.A.W. Vance, "An Integrated Circuit VHF Radio Receiver," *The Radio and Electronic Engineer*, Vol. 50, No. 4, 1980, pp. 158–164.
[37] I.A.W. Vance, "Fully Integrated Radio Paging Receiver," *IEEE Proc.*, Vol. 129, Part F, No. 1, 1982, pp. 2–6.

[38] R. C. French, "A High Technology VHF Radio Paging Receiver," *Int'l Conf. on Mobile Radio Systems & Techniques*, York, UK, 1984, pp. 11–15.
[39] J. Min, A. Rofougaran, H. Samueli, and A. A. Abidi, "An All-CMOS Architecture for a Low-Power Frequency-Hopped 900 MHz Spread-Spectrum Transceiver," *Custom IC Conf.*, San Diego, CA, 1994, pp. 379–382.
[40] K. Holtvoeth, "A Fully-Integrated Low-Power FSK Receiver for VHF and UHF Paging," *IEEE Trans. on Consumer Electronics*, Vol. 35, No. 3, 1989, pp. 707–713.
[41] S. Drude, and T. Rudolph, "New Chip-set Reduces Power Consumption of Radio Pagers," *Electronic Components and Applications*, Vol. 9, No. 2, 1989, pp. 112–124.
[42] J. F. Wilson, R. Youell, T. H. Richards, G. Luff, and R. Pilaski, "A Single-Chip VHF and UHF Receiver for Radio Paging," *IEEE J. of Solid State Circuits*, Vol. 26, No. 12, 1991, pp. 1944–1950.
[43] K. Nagata, D. Ishii, T. Mori, T. Oyagi, and T. Seo, "Slim Digital Pagers," *NEC Research & Development.*, No. 74, July 1984, pp. 55–64.
[44] K. Yamasaki, S. Yoshizawa, Y. Minami, T. Asai, Y. Nakano, and M. Kuroda, "Compact Size Numeric Display Pager with New Receiving System," *NEC Research & Development*, Vol. 33, No. 1, Jan. 1992, pp. 73–81.
[45] A. Burt, "Direct Conversion Receivers Come of Age in the Paging World," *GEC Review*. Vol. 7, No. 3, 1992, pp. 156–160.
[46] S. Tanaka, A. Nakajima, J. Nakagawa, and A. Nakagoshi, "High-Frequency, Low-Voltage Circuit Technology for VHF Paging Receiver," *IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E76-A, No. 2, 1993, pp. 156–163.
[47] B. F. Cambou, "MMIC Consumer Application and Production," *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, Albuquerque, NM, 1992, pp. 1–3.
[48] H. Komatsu, R. Norose, and S. J. Symonds, "The Wristwatch as a Personal Communications Device," in *Pacific Telecommunications Conf.*, Honolulu, 1993, pp. 242–244.
[49] M. Pardoen, R. Pache, and E. Dijkstra, "Direct-Conversion Receiver Provides CP-FSK Operation," *Microwaves & RF*. March 1994, pp. 151–155.
[50] G. J. Lomer, "Telephoning on the Move—Dick Tracy to Captain Kirk," *IEEE Proc.*, Vol. 134, Part F, No. 1, 1987, pp. 1–8.
[51] K. Wakino, "Recent Developments of Dielectric Resonator Materials and Filters in Japan," *Ferroelectrics*, Vol. 91, 1989, pp. 69–86.
[52] M. Hikita, T. Tabuchi, and A. Sumioka, "Miniaturized SAW Devices for Radio Communication Transceivers," *IEEE Trans. on Vehicular Technology*, Vol. 38, No. 1, 1989, pp. 2–8.
[53] K. Eda, "Ultra-Small SAW Filter Works in Mobile Communications," *J. of Electronic Engineering*. May 1992, pp. 54–57.
[54] Y. Tamura, T. Maru, N. Hirasawa, H. Okuno, M. Komoda, and M. Hotsumi, "Hand-Held Portable Equipment for Cellular Mobile Telephone," *NED Research and Development*, No. 87, Oct. 1987, pp. 34–43.
[55] D. Anderson, and R. J. Zavrel, "RF ICs for Portable Communications Equipment," *Electronic Components and Applications*, Vol. 7, No. 1, 1985, pp. 37–44.
[56] I. A. Koullias, S. L. Forgues, and P. C. Davis, "A 100 MHz IF Amplifier/Quadrature Demodulator for GSM Cellular Radio Mobile Terminals," *Bipolar Circuits and Technology Mtg.*, Minneapolis, MN, 1990, pp. 248–251.
[57] Y. Mori, H. Yabuki, M. Ohba, M. Sagawa, M. Makimoto, and I. Shibazaki, "Miniaturized RF-Circuit Modules for Land Mobile Communication Equipment," *IEEE Vehicular Technology Conf.*, Orlando, FL, 1990, pp. 65–70.
[58] M.. Hikita, A. Yuhara, and K. Oda, "New Low-Loss, Miniature SAW Filters," *Hitachi Review*, Vol. 42, No. 3, June 1993, pp. 119–124.
[59] L. Scharf, H. Frank, and T. Pollakowski, "High f_T and Low Noise," *Siemens Components*. Vol. XXV, No. 4, 1990, pp. 119–122.

[60] S. Yamada, and K. Oguri, "Low-Power Front-End Oriented Discrete Semiconductor Devices," *Hitachi Review*, Vol. 42, No. 3, June 1993, pp. 101–106.
[61] C. Huang, "Where Do Integrated Circuits Replace Discretes?," *Applied Microwaves & Wireless*, Summer 1994, pp. 5–6.
[62] S. P. MacCabe, "GaAs MMICs for Cellular Telephones, GPS and DBS Markets," *WESCON*, San Francisco, CA, 1989, pp. 279–281.
[63] M. Rocchi, "GaAs & Si MMIC Building Blocks: A Moot Point Revisited," *Microelectronic Engineering*, Vol. 15, 1991, pp. 685–692.
[64] Y. Aoki, K. Kobayashi, and W. Kennan, "Monolithic Microwave ICs (MMICs) for Telecommunication," *Fujitsu Scientific & Technical J.*, Vol. 30, No. 1, June 1994, pp. 32–39.
[65] Y. Imai, M. Tokumitsu, A. Minakawa, T. Sugeta, and M. Aikawa, "Very Low-Current and Small Size GaAs MMICs for L-Band Front-End Applications," *GaAs IC Symposium*, Monterey, CA, 1989, pp. 71–74.
[66] Y. Imai, M. Tokumitsu, and A. Minakawa, "Design and Performance of Low-Current GaAs MMICs for L-Band Front-End Applications," *IEEE Trans. On Microwave Theory & Techniques*, Vol. 39, No. 2, 1991, pp. 209–215.
[67] K. R. Cioffi, "Ultra-low dc Power Consumptions in Monolithic L-Band Components," *IEEE Trans. on Microwave Theory & Techniques*, Vol. 40, No. 12, 1992, pp. 2467–2472.
[68] C. Kermarrec, M. Takano, F. McGrath, P. O'Sullivan, G. Dawe, E. Heaney, and G. St. Onge, "High Performance, Low Cost GaAs MMICs for Personal Phone Applications," *Int'l Symp. on GaAs and Related Compounds*, Karuizawa, Japan, 1992.
[69] E. Heaney, F. McGrath, P. O'Sullivan, and C. Kermarrec, "Ultra Low Power Low Noise Amplifiers for Wireless Communications," *GaAs IC Symp.*, San Jose, CA, 1993, pp. 49–51.
[70] S. Hara, K. Osato, A. Yamada, T. Tsukao, and T. Yoshimasu, "Miniaturized Low Noise Variable MMIC Amplifiers with Low Power Consumption for L-Band Portable Communication Applications," *Microwave and Millimeter-Wave Monolithic Circuits Symp.*, Atlanta, GA, 1993, pp. 67–70.
[71] T. Ohgihara, S. Kusunoki, M. Wada, and Y. Murakami, "GaAs JFET Front-End MMICs for L-Band Personal Communications," *Microwave and Millimeter-Wave Monolithic Circuits Symp.*, Atlanta, GA, 1993, pp. 9–12.
[72] C. Kusano, E. Hase, and K. Sakamoto, "Low-Power, Low-Noise GaAs MMIC," *Hitachi Review*, Vol. 42, No. 3, June 1993, pp. 107–112.
[73] R. E. Lehmann, and D. D. Heston, "X-Band Monolithic Series Feedback LNA," *IEEE Trans. on Microwave Theory & Techniques*, Vol. MTT-33, No. 12, 1985, pp. 1560–1566.
[74] H. Sakai, A Tezuka, Y. Mori, M. Sagawa, T. Katoh, J. Itoh, and K. Fujimoto, "Low-Power GaAs ICs for Mobile Communication Equipment," *Int'l Symp. on GaAs and Related Compounds*, Karuizawa, Japan, 1992.
[75] V. Nair, R. Vaitkus, D. Scheitlin, J. Kline, and H. Swanson, "Low Current GaAs Integrated Down Converter for Portable Communication Applications," *GaAs IC Symp.*, San Jose, CA, 1993, pp. 41–44.
[76] C. Woo, A Podell, R. Benton, D. Fisher, and J. Wachsman, "A Fully Integrated Transceiver Chip for the 900 MHz Communication Bands," *GaAs IC Symp.*, Miami Beach, FL, 1992, pp. 143–146.
[77] P. E. Chadwick, "High Performance Integrated Circuit Mixers," *In Radio Receivers & Associated Systems*, London, UK, 1981, pp. 1–9.
[78] D. M. Snider, "A Theoretical Analysis and Experimental Confirmation of the Optimally Loaded and Overdriven RF Power Amplifier," *IEEE Trans. on Electron Devices*, Vol. ED-14, No. 12, 1967, pp. 851–857.
[79] J. J. Komiak, "Design and Performance of an Octave Band 11 Watt Power Amplifier MMIC," *IEEE Trans. on Microwave Theory & Techniques*, Vol. 38, No. 12, 1990, pp. 2001–2006.
[80] T. Takagi, Y. Ikeda, K. Seino, G. Toyoshima, and A. Inoue, "A UHF Band 1.3W Monolithic

Amplifier with Efficiency of 63%," *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, Albuquerque, NM, 1992, pp. 35–38.
[81] M. Gat, D. S. Day, S. Chan, C. Hua, and J. R. Basset, "A 3.0 Watt High Efficiency C-Band Power MMIC," *GaAs IC Symp.*, Monterey, CA, 1991, pp. 331–334.
[82] O. Ishikawa, M. Maeda, K. Nishii, M. Yanagihara, T. Yokoyama, Y. Ikeda, and Y. Ota, "Cellular Telecommunication GaAs Power Modules," *Applied Microwave & Wireless*, Vol. 4, No. 3, Fall 1992, pp. 83–88.
[83] M. Maeda, M. Nishijima, H. Takehara, C. Adachi, H. Fujimoto, and O. Ishikawa, "A 3.5V, 1.3W GaAs Power Multi-chip IC for Cellular Phones," *IEEE J. Of Solid-State Circuits*, Vol. 29, No. 10, 1994, pp. 1250–1256.
[84] I. Yoshida, M. Katsueda, S. Ohtaka, Y. Maruyama, and T. Okabe, "Highly Efficient 1.5 GHz Si Power MOSFET for Digital Cellular Front End," *Int'l Symp. on Power Semiconductor Devices*, Tokyo, 1992, pp. 156–157.
[85] T. Okabe, and K. Kobayashi, "Radio-Frequency MOS Power Module," *Hitachi Review*, Vol. 42, No. 3, June 1993, pp. 113–118.
[86] R. G. Meyer, and W. D. Mack, "A 1-GHz BiCMOS RF Front-End IC," *IEEE J. of Solid-State Circuits*, Vol. 29, No. 3, 1994, pp. 350–355.
[87] M. N. Sera, "Chip Set Addresses North American Digital Cellular Market," *RF Design*, Vol. 17, No. 3, March 1994, pp. 54–62.
[88] Y. Funada, "VCO Techniques Enable Compact, High-Frequency Mobilecom," *J. of Electronic Engineering*, Dec. 1993, pp. 40–43.
[89] Y. Tamura, A. Yonehata, S. Miyazaki, F. Kobayasi, Y. Fukuda, and J. Kamishiro, "Development of Advanced Mobile Telephone P3 (Personal Pocket Phone)," *NEC Research and Development*, No. 98, July 1990, pp. 60–70.
[90] H. Ohmori, "SAW Filters Hold Key to Creating Small, Portable Equipment," *J. of Electronic Engineering*, May 1992, pp. 63–67.
[91] G. Kainz, "Microwave Ceramics Boosts Miniaturization," *Siemens Components*, Vol. 27, No. 5, Sep-Oct 1992, pp. 11–14.
[92] Y. Kado, M. Suzuki, K. Koike, and Y. Omura, "A 1-GHz/0.9-mW CMOS/SIMOX Divide-by-128/129 Dual-Modulus Prescaler using a Divide-by-2/3 Synchronous Counter," *IEEE J. of Solid-State Circuits*, Vol. 28, No. 4, 1993, pp. 513–517.
[93] T. Seneff, L. McKay, K. Sakamoto, and N. Tracht, "A Sub-1 mA 1.5-GHz Silicon Bipolar Dual Modulus Prescaler," *IEEE J. of Solid-State Circuits*, Vol. 29, No. 10, 1994, pp. 1206–1211.
[94] M. Morishima, "Present and Future of Mobile Communications in Japan," *Microwave & Millimeter-Wave Monolithic Circuits Conf.*, Atlanta, GA, 1993, pp. 3–6.
[95] I. Shimizu, S. Urabe, K. Hirade, K. Nagata, and S. Yuki, "A New Pocket-size Cellular Telephone for NTT High-Capacity Land Mobile Communication System," *Vehicular Technology Conf.*, St. Louis, MO, 1991, pp. 114–119.
[96] I. A. Koullias, J. H. Havens, I. G. Post, and P. E. Bronner, "A 900 MHz Transceiver Chip Set for Dual-Mode Cellular Radio Terminals," *Int'l Solid State Circuits Conference*, San Francisco, CA, 1993, pp. 140–141.
[97] L. Longo, R. Halim, B.-R. Horng, K. Hsu, and D. Shamlou, "A Cellular Analog Front End with a 98 dB IF Receiver," *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1994, pp. 36–37.
[98] D. J. Goodman, "Second Generation Wireless Information Networks," *IEEE Trans. on Vehicular Technology*, Vol. 40, No. 2, 1991, pp. 366–374.
[99] J. Fenk, W. Birth, R. G. Irvine, P. Sehrig, and K. R. Schon, "An RF Front-End for Digital Mobile Radio," *Bipolar Circuits and Technology Mtg.*, Minneapolis, MN, 1990, pp. 244–247.
[100] E. Badura, "New Integrated Circuits for GSM Mobile Radio," *Siemens Components*, Vol. XXVI, No. 2, 1991, pp. 72–76.

[101] S. Beyer and G. Lipperer, Low-Current Oscillator Design for 900 MHz GSM Applications, in *Microwave Engineering Europe*. Oct. 1991, pp. 35–41.
[102] R. E. Jenkins, "Transceiver Chip Simplifies GSM Cellular Design," *RF Design*, June 1994, pp. 26–35.
[103] J. Sevenhans, A. Vanwelsenaers, J. Wenin, and J. Baro, "An Integrated Si Bipolar RF Transceiver for a Zero IF 900 MHz GSM Digital Radio Front-end of a Hand Portable Phone," *Custom IC Conf.*, San Diego, CA, 1991, pp. 7.7/1–4.
[104] J. Sevenhans, and A. Vanwelsenaers, "A Silicon Transceiver for a 900 MHz GSM Hand Set," *Microwave Engineering Europe*, Dec./Jan. 1992/93, pp. 59–63.
[105] D. Haspeslagh, J. Ceuterick, L. Kiss, and J. Wenin, "BBTRX: A Baseband Transceiver for a Zero IF GSM Hand Portable Station," *Custom IC Conf.*, San Diego, CA, 1992, pp. 10.7.1–10.7.4.
[106] M. McDonald, "A DECT Transceiver Chip Set," *Custom IC Conf.*, San Diego, CA, 1992, pp. 10.6.1–10.6.4.
[107] B. Madsen, and D. E. Fague, "Radios for the Future: Designing for DECT," *RF Design*, April 1993, pp. 48–53.
[108] M. Eccles, "Gigahertz Systems on a Chip," *Electronics World+Wireless World*, Aug. 1993, pp. 662–663.
[109] I. White, "UHF Technology for the Cordless Revolution," *Electronics World+Wireless World*, July/Aug. 1993, pp. 542–546/657–661.
[110] A. Boveda, G. L. Bonato, and O. Ripolles, "GaAs Monolithic 1.5-2.5 GHz Image Rejection Receiver," *Int'l Symp. on Circuits & Systems*, San Diego, CA, 1992, pp. 232–235.
[111] A. Boveda, F. Ortigoso, and J. I. Alonso, "A 0.7-3 GHz GaAs QPSK/QAM Direct Modulator," *Int'l Solid State Circuits Conference*, San Francisco, CA, 1993, pp. 142–143.
[112] M. Williams, F. Bonn, C. Gong, and T. Quach, "GaAs RF ICs Target 2.4-GHz Frequency Band," *Microwaves & RF*, July 1994, pp. 111–118.
[113] M. D. McDonald, "A 2.5 GHz BiCMOS Image Reject Front-End," *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1993, pp. 144–145.
[114] J. L. Wang, T. Tsuchiya, H. Takeuchi, K. Shirotori, S. Miyazaki, and T. Nakata, "A Miniature Ultra Low Power 2.5 GHz Downconverter IC for Wireless Communications," *NEC Research & Development*, Vol. 35, No.1, Jan. 1994, pp. 46–50.
[115] T. Tsukahara, M. Ishikawa, and M. Muraguchi, "A 2V 2 GHz Direct-Conversion Quadrature Modulator," *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1994, pp. 40–41.
[116] M. Nakatsugawa, Y. Yamaguchi, and M. Muraguchi, "An L-Band Ultra Low Power Consumption Monolithic Low Noise Amplifier," *GaAs IC Symp.*, San Jose, CA, 1993, pp. 45–48.
[117] D. Ngo, B. Beckwith, P. O 'Neil, and N. Camilleri, "Low Voltage GaAs Power Amplifiers for Personal Communications at 1.9 GHz," *MTT-S Int'l Microwave Symp.*, Atlanta, GA, 1993, pp. 1461–1464.
[118] T. Kunihisa, T. Yokoyama, H. Fujimoto, K. Ishida, H. Takehara, and O. Ishikawa, "High Efficiency, Low Adjacent Channel Leakage GaAs Power MMIC for Digital Cordless Telephone," *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, San Diego, CA, 1994, pp. 55–58.
[119] T. Tokumitsu, I. Toyoda, and M. Aikawa, "Low Voltage, High Power T/R Switch MMIC using LC Resonators," *Microwave and Millimeter-Wave Monolithic Circuits Symp.*, Atlanta, GA, 1993, pp. 27–30.
[120] K. Miyatsuji, S. Nagata, N. Yoshikawa, K. Miyanaga, Y. Ohishi, and D. Ueda, "A GaAs High-Power RF Single-Pole Double-Throw Switch IC for Digital Mobile Communication System," *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1994, pp. 34–35.
[121] P. O'Sullivan, G. St. Onge, E. Heaney, F. McGrath, and C. Kermarrec, "High Performance Integrated PA, T/R Switch for 1.9 GHz Personal Communications Handsets," *GaAS IC Symp.*, San Jose, CA, 1993, pp. 33–35.

[122] A. Bateman, D. M. Haines, and R. J. Wilkinson, "Linear Transceiver Architectures," *IEEE Vehicular Technology Conf.*, Philadelphia, PA, 1988, pp. 478–484.
[123] G. Schultes, A. L. Scholtz, E. Bonek, and P. Veith, "A New Incoherent Direct Conversion Receiver," *Vehicle Technology Conf.*, Orlando, FL, 1990, pp. 668–674.
[124] G. Schultes, E. Bonek, P. Weger, and W. Herzog, "Basic Performance of a Direct Conversion DECT Receiver," *Electronics Letters*, Vol. 26, No. 21, 1990, pp. 1746–1748.
[125] J. Sevenhans, D. Haspesalgh, A. Delarbre, L. Kiss, Z. Chang, and J. F. Kukielka, "An Analog Radio Front-end Chip Set for a 1.9 GHz Mobile Radio Telephone Application," *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1994, pp. 44–45.
[126] V. Thomas, J. Fenk, and S. Beyer, "A One-Chip 2 GHz Single Superhet Receiver for 2 Mb/s FSK Radio Communication," *Int'l solid State Circuits Conf.*, San Francisco, CA, 1994, pp. 42–43.
[127] W. Veit, J. Fenk, S. Ganser, K. Hadjizada, S. Heinen, H. Herrmann, and P. Sehrig, "A. 2.7V 800 MHz -2.1 GHz Transceiver Chipset for Mobile Radio Applications in 25 GHz fT Si-Bipolar," *Bipolar Circuits & Technology Mtg.*, Minneapolis, MN, 1994, pp. 175–178.
[128] D. K. Weaver, "A Third Method of Generation and Detection of Single-Sideband Signals," *Proc. of the IRE*, Vol. 44, No. 12, 1956, pp. 1703–1705.
[129] K. Negus, B. Koupal, J. Wholey, K. Carter, D. Millicker, C. Snapp, and N. Marion, "Highly Integrated Transmitter RFIC with Monolithic Narrowband Tuning for Digital Cellular Handsets," *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1994, pp. 38–39.
[130] A. A. Abidi, "Radio-Frequency Integrated Circuits for Portable Communications," *Custom IC Conf.*, San Diego, CA, 1994, p. 151–158.
[131] J. Wenin, "ICs for Digital Cellular Communication," *European Solid State Circuits Conf.*, Ulm, Germany, 1994, pp. 1–10.
[132] H. Tsurumi and T. Maeda, "Design Study on a Direct Conversion Receiver Front-End for 280 MHz, 900 MHz, and 2.6 GHz Band Radio Communication Systems," *IEEE Vehicular Technology Conf.*, St. Louis, MO, 1991, pp. 457–462.
[133] A. Bateman and D. M. Haines, "Direct Conversion Transceiver Design for Compact Low-Cost Portable Mobile Radio Terminals," *IEEE Vehicular Technology Conf.*, San Francisco, CA, 1989, pp. 57–62.
[134] A. W. Buchwald and K. W. Martin, "High-Speed Voltage-Controlled Oscillator with Quadrature Outputs," *Electronics Letters*, Vol. 27, No. 4, 1991, pp. 309–310.
[135] S. K. Enam and A. A. Abidi, "NMOS ICs for Clock and Data Regeneration in Gb/s Optical-Fiber Receivers," *IEEE J. of Solid State Circuits*, Vol. 27, No. 12, 1992, pp. 1763–1774.
[136] J. K. Cavers and M. W. Liao, "Adaptive Compensation for Imbalance and Offset Losses in Direct Conversion Transceiver," *IEEE Trans. on Vehicular Technology*, Vol. 42, No. 4, 1993, pp. 581–588.
[137] J. T. Taylor and J. K. Omura, "Spread Spectrum Technology: A Solution to the Personal Communications Services Frequency Allocation Dilemma," *IEEE Communications Magazine*, Vol. 29, No. 2, Feb. 1991, pp. 48–51.
[138] R. A. Scholtz, "The Origins of Spread-Spectrum Communications," *IEEE Trans. on Communications*, Vol. COM-30, No. 5, 1982, pp. 822–852.
[139] D. T. Magill, F. D. Natali, and G. P. Edwards, "Spread-Spectrum Technology for Commercial Applications," *Proc. of IEEE*, Vol. 82, No. 4, 1994, pp. 572–584.
[140] R. C. Dixon, *Spread Spectrum Systems*, New York, NY: John Wiley & Sons, 1976.
[141] M. Leonard, "Wireless Data Links Broaden LAN Options," *Electronic Design*. March 19, 1992, pp. 51–58.
[142] D. T. Magill, "A Fully-Integrated, Digital, Direct-Sequence, Spread Spectrum Modem ASIC," *Int'l Symp. on Personal, Indoor and Mobile Radio Communications*, Boston, MA, 1992, pp. 42–46.
[143] R. Jain, H. Samueli, P. T. Yang, and C. Chien, "Computer-Aided Design of a BPSK Spread-Spectrum Chip Set," *IEEE J. of Solid State Circuits*, Vol. 27, No. 1, 1992, pp. 44–58.

[144] C. Chien, P. Yang, E. Cohen, R. Jain, and H. Samueli, "A 12.7 Mchip/s All-Digital BPSK Direct Sequence Spread-Spectrum IF Transceiver in 1.2mm CMOS," *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1994, pp. 30–31.
[145] M. Leonard, "PCMCIA-Sized Radio Links Portable WLAN Terminals," *Electronic Design*. Aug. 5 1993, pp. 45–50.
[146] L. M. Devlin, B. J. Buck, J. C. Clifton, A. W. Dearn, and A. P. Long, "A 2.4 GHz Single Chip Transceiver," *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, Atlanta, GA, 1993, pp. 23–26.
[147] T. Apel, E. Creviston, S. Ludvik, L. Quist, and B. Tuch, "A GaAs MMIC Transceiver for 2.45 GHz Wireless Commercial Products," in *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, San Diego, CA, 1994, pp. 15–18.
[148] B. Khabbaz, A. Douglas, J. DeAngelis, L. Hongsmatip, V. Pellicia, W. Fahey, and G. Dawe, "A High Performance 2.4 GHz Transceiver Chip Set for High Volume Commercial Applications," in *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, San Diego, CA, 1994, pp. 11–14.
[149] M. S. Wang, M. Carriere, P. O'Sullivan, and B. Maoz, "A Single-Chip MMIC Transceiver for 2.4 GHz Spread Spectrum Communication," *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, San Diego, CA, 1994, pp. 19–22.
[150] J. Min, A. Rofougaran, V. Lin, M. Jensen, H. Samueli, A. A. Abidi, G. Pottie, and Y. Rahmat-Samii, "A Low-Power Handheld Frequency-Hopped Spread Spectrum Transceiver Hardware Architecture," *Virginia Tech's Third Symp. on Wireless Personal Communications*, Blacksburg, VA, 1993, pp. 10/1–8.
[151] M. Sakakura and S. Skiest, "Ultra-Miniature Chip Inductors Serve at High Frequency," *J. of Electronic Engineering*. Dec. 1993, pp. 48–51.
[152] N. M. Nguyen and R. G. Meyer, "Si IC-Compatible Inductors and LC Passive Filters," *IEEE J. of Solid State Circuits*, Vol. 25, No. 4, 1990, pp. 1028–1031.
[153] N. M. Nguyen and R. G. Meyer, "A Silicon Bipolar Monolithic RF Bandpass Amplifier," *IEEE J. of Solid State Circuits*, Vol. 27, No. 1, 1992, pp. 123–127.
[154] N. M. Nguyen and R. G. Meyer, "A 1.8-GHz Monolithic LC Voltage-Controlled Oscillator," *IEEE J. of Solid State Circuits*, Vol. 27, No. 3, 1992, pp. 444–450.
[155] D. Lovelace, N. Camilleri, and G. Kannell, "Silicon MMIC Inductor Modeling for High Volume, Low Cost Applications," *Microwave J.*, Vol. 37, No. 8, Aug. 1994, pp. 60–71.
[156] K. B. Ashby, W. C. Finley, J. J. Bastek, S. Moinian, and L. A. Koullias, "High Q Inductors for Wireless Applications in a Complementary Silicon Bipolar Process," *Bipolar Circuits & Technology Mtg.*, Minneapolis, MN, 1994, pp. 179–182.
[157] J. Y.-C. Chang, A. A. Abidi, and M. Gaitan, "Large Suspended Inductors on Silicon and their use in a 2-mm CMOS RF Amplifier," *IEEE Electron Device Letters*, Vol. 14, No. 5248, 1993, pp. 246.
[158] P. Y. Chan, A. Rofougaran, K. A. Ahmed, and A. A. Abidi, "A Highly Linear 1-GHz CMOS Downconversion Mixer," *European Solid-State Circuits Conf.*, Sevilla, Spain, 1993, pp. 210–213.
[159] H. T. Nicholas and H. Samueli, "A 150-MHz Direct Digital Frequency Synthesizer in 1.25 mm CMOS with -90 dBc Spurious Response," *IEEE J. of Solid State Circuits*, Vol. 26, 1991, pp. 1959–1969.
[160] G. Chang, A. Rofougaran, M. K. Ku, A. A. Abidi, and H. Samueli, "A Low-Power CMOS Digitally Synthesized 0-13 MHz Agile Sinewave Generator," *Int'l Solid State Circuits Conf.*, San Francisco, CA, 1994, pp. 32–33.
[161] M. Rofougaran, A. Rofougaran, C. Olgaard, and A. A. Abidi, "A 900 MHz CMOS RF Power Amplifier with Programmable Output," *Symp. on VLSI Circuits*, Honolulu, HI, 1994, pp. 133–134.
[162] K. Kimura, "A CMOS Logarithmic Amplifier with Unbalanced Source-Coupled Pairs," *IEEE J. of Solid State Circuits*, Vol. 28, No. 1, 1993, pp. 78–83.

[163] S. Watanabe, "Technology Transfer for High Frequency Devices for Consumer Electronics," *IEEE Trans. on Microwave Theory & Techniques*, Vol. 40, No. 12, 1992, pp. 2461–2466.
[164] U. Ablassmeier, W. Kellner, and H. Kniepkamp, "GaAs FET Upconverter for TV Tuner," *IEEE Trans. on Electron Devices*, Vol. ED-27, No. 6, 1980, pp. 1156–1159.
[165] K. Torii, S. Fujimori, S. Komatsu, S. Shimizu, K. Yoshihasa, J. Nagai, and Y. Yamagata, "Monolithic Integrated VHF TV Tuner," *IEEE Trans. on Consumer Electronics*, Vol. CE-26, 1980, pp. 180–187.
[166] J.-E. Muller, U. Ablassmeier, J. Schelle, W. Kellner, and H. Kniepkamp, "A Double-Conversion Broad Band TV-Turner with GaAs ICs," in *GaAs IC Symposium*, 1984, pp. 97–100.
[167] P. Dautriche, B. Y. Lao, C. Villalon, V. Pauker, M. Bostelmann, and M. Binet, "GaAs Monolithic Circuits for TV Tuners," in *GaAs IC Symp*, Monterey, CA, 1985, pp. 165–168.
[168] J. Fenk and R. Tauber, "TV VHF/Hyperband Tuner ICs," *IEEE Trans. on Consumer Electronics*, Vol. CE-32, No. 4, 1986, pp. 723–733.
[169] H. van Glabbeek, N. Baars, H. Schreurs, and W. Zwijsen, "VHF, Hyperband, and UHF Mixer/Oscillator Sections of a TV Tuner on One IC," *IEEE Trans. on Consumer Electronics*, Vol. CE-33, No. 4, 1987, p. 619–622.
[170] T. Nakatsuka, S. Sakashita, K. Goda, and S. Nambu, "A GaAs RF Amplifier IC for UHF TV Tuners," *IEEE Trans. on Consumer Electronics*, Vol. 34, No. 2, pp. 366–371, 1988.
[171] H. Mizukami, K. Sakuta, H. Hatashita, T. Nagashima, and K. Shinkawa, "A High Quality GaAs IC Tuner for TV/VCR Receivers," *IEEE Trans. on Consumer Electronics*, Vol. 34, No. 3, 1988, pp. 649–658.
[172] T. Ducourant, P. Phillippe, P. Dautriche, V. Pauker, C. Villalon, M. Pertus, and J.-P. Damour, "A3-Chip GaAs Double-Conversion TV Tuner with 70 dB Image Rejection," *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, pp. 87–90, Long Beach, CA, 1989.
[173] H. Yagita, A. Terao, M. Tsuneoka, A. Watanabe, T. Tambo, and S. Nambu, "Low Noise and Low Distortion GaAs Mixer-Oscillator IC for Broadcasting Satellite TV Tuner," *GaAs IC Symp.*, San Diego, CA, 1989, pp. 75–78.
[174] P. Philippe, and M. Pertus, "A 2 GHz Enhancement-Mode GaAs Down-Converter IC for Satellite TV Tuner," *Microwave Theory & Techniques Symp.*, Boston, MA, 1991, pp. 73–76.
[175] H. Mizukami, H. Ikedo, K. Ideno, T. Nagashima, and S. Yamada, "Low Supply Voltage GaAs ICs for a TV Tuner," *Int'l Symp. on GaAs and Related Compounds*, Karuizawa, Japan, 1992.
[176] K. Ideno, H. Mitzukami, T. Nagashima, and K. Sakamoto, "A Wide Band GaAs IC for a DBS/TV Receiver," *Int'l Symp. on GaAs and Related Compounds*, Karuizawa, Japan, 1992.
[177] K. Washio, T. Okabe, K. Norisue, and T. Nagashima, "An All-Band TV Tuner IC with 10-GHz 100-V Mixed Analog/Digital Si Bipolar Technology," *IEEE J. of Solid-State Circuits*, Vol. 27, No. 9, 1992, pp. 1264–1269.
[178] N. Scheinberg, R. Michels, V. Fedoroff, and D. Stoffman, "A GaAs Upconverter Integrated Circuit for a Double Conversion Cable TV Set-Top Tuner," *IEEE J. of Solid-State Circuits*, Vol. 29, No. 6, 1994, pp. 688–692.
[179] J. Hurn, *GPS: A Guide to the Next Utility*, Sunnyvale, CA: Trimble Navigation, 1989.
[180] W. O. Dussell and J. Medina, "Using Creative Surface Mount Manufacturing and Packaging as a Competitive Weapon," *Surface Mount Int'l Conf. & Exposition*, San Jose, CA, 1993, pp. 241–245.
[181] R. M. Herman, C. H. Mason, H. P. Warren, and R. A. Meier, "A GPS Receiver with Synthesized Local Oscillator," *Int'l Solid-State Circuits Conf.*, New York, NY, 1989, pp. 194–195.
[182] R. M. Herman, A. Chao, C. H. Mason, and J. R. Pulver, "An Integrated GPS Receiver with Synthesizer and Downconversion Functions," *Int'l Microwave Symp.*, Boston, 1991, pp. 883–886.
[183] R. Benton, M. Nijjar, C. Woo, A. Podell, G. Horvath, E. Wilson, and S. Mitchell, "GaAs MMICs for an Integrated GPS Front End," *GaAs IC Symp.*, Miami Beach, FL, 1992, pp. 123–126.
[184] K. J. Negus, R. A. Koupal, D. Millicker, and C. P. Snapp, "3.3V GPS Receiver MMIC Implemented

on a Mixed-Signal, Silicon Bipolar Array," *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, Albuquerque, NM, 1992, pp. 209–212.

[185] N. Camilleri, J. Kirschgessner, J. Costa, D. Ngo, and D. Lovelace, "Bonding Pad Models for Silicon VLSI Technologies and their effects on the noise figure of RF NPNs," *Microwave & Millimeter-Wave Monolithic Circuits Symp.*, San Diego, CA, 1994, pp. 225–228.

INTEL 1005 2 OF 3

CHAPTER 4

DIGITAL MODULATION AND DEMODULATION

Lawrence Burns
3Com Corporation
Santa Clara, CA

4.1 WHY DIGITAL MODULATION?

Most modern wireless communication systems use digital modulation/demodulation techniques. Digital modulation offers a number of advantages over analog modulation, such as increased channel capacity and the ability to transmit and receive messages with greater accuracy than an analog communication system in the presence of noise and distortion [1–9].

In digital communication systems, a finite number of electrical waveforms, or *symbols*, are transmitted. Each symbol can represent one or more bits. The job of the receiver is then to *estimate* which symbol was originally sent by the transmitter. The trick, of course, is estimating which symbol was sent after noise and distortion has been added. It is largely unimportant what size (amplitude) or shape (how distorted) the received signal is so long as the receiver can clearly distinguish one symbol from another.

In Figure 4.1(a), an undistorted binary bit stream is shown. As the figure shows, the original signal is free of distortion, with sharp, well-defined edges. It is easy at this point to determine the length of a bit period and whether each bit is a "1" or a "0." As the signal passes through the channel, distortion changes its shape.

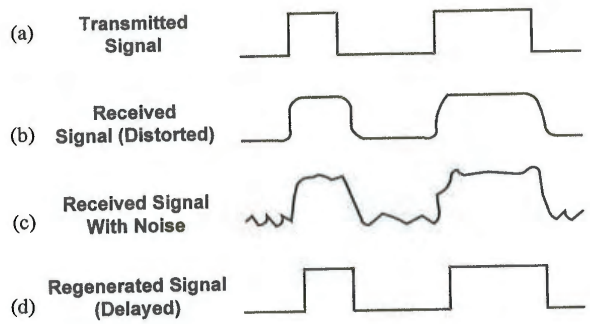


Figure 4.1 Waveforms in a typical digital communication system: (a) transmitted signal, (b) distorted receive signal, (c) distorted receive signal with noise, (d) regenerated signal (delayed).

Although it is somewhat less well-defined, it is still not hard to determine each bit period and its polarity. This is shown in Figure 4.1(b). Distortion in wireless communication systems can be caused by passing the signal through filters having insufficient bandwidth. This situation, where the effects of one bit affect succeeding bits, is called intersymbol interference (ISI). In addition to ISI, multiple versions of the original signal can be received at the same time. These delayed versions of the signal are produced when the transmitted signal reflects off of multiple objects on its way to the receiver. This effect is called *delay spread*. When the bit rate of the signal is high, delay spread can be on the order of one or more bit periods. Equalizers can be used to remove the ISI caused by filter effects and delay spread.

In addition to distortion, noise is added to the signal as it passes through the channel, as shown in Figure 4.1(c). Noise can come from the thermal noise within the receiver itself or can come from manmade or natural interference, such as motor noise or lightning. Other transmitters on adjacent channel frequencies can spill over into a receiver's band. This is called adjacent channel interference (ACI). In most cases, the amount of spectrum users are allowed to use is limited. This means users must reuse some or all of their channel frequencies. In general, transmitters operating at identical frequencies are placed some distance away from each other so that their transmissions do not interfere. When interference does occur, it is referred to as cochannel interference (CCI).

The addition of distortion and noise to the signal makes it harder to determine whether a 1 or a 0 was originally sent. If the amount of noise and distortion is small, then the receiver can easily decide what the original signal was. This is shown in Figure 4.1(d). In this case, the receiver was able to recreate the original bit stream without error or distortion. The ability of a digital communication systems to exactly

regenerate the original signal is an advantage digital communication systems have over analog communication systems. It is this very feature that makes compact disc players have such high-fidelity audio, and the reason why laser disks have such sharp picture quality.

If the level of distortion and noise is too large, the receiver's estimate of whether a 1 or a 0 was sent will occasionally be wrong. When this happens, a *bit error* has occurred. A useful parameter that can be used to gauge the quality of a digital communication system is the probability of bit error, P_E. Because our ears have the ability to decipher what another person is saying, even in the presence of noise, most wireless voice systems can tolerate one bit error in every 1,000 bits received ($P_E = 10^{-3}$), before their performance is deemed unacceptable. Other applications, such as data transfer, demand a much lower P_E. Typical values are 1 error in every 100,000,000 to 1 error in every 10 million bits sent ($P_E = 10^{-8}$ or 10^{-10}) for systems such as Ethernet and token ring networks.

In analog communication systems, the level of performance tends to fall off somewhat gradually with increasing noise. In contrast, when the level of noise in a digital communication system reaches the point where the receiver can no longer tell one symbol from another, overall quality deteriorates rapidly. For example, a 1-dB decrease in signal-to-noise ratio (SNR) can cause a tenfold increase in P_E.

The rest of this chapter will discuss various phase and frequency modulations as well as other digital modulations, which are variants of these. We will discuss how to filter the signal so as to reduce the transmitted bandwidth, yet have no ISI at the sampling instants (Nyquist filtering). Finally, we will give examples of wireless systems that use these various digital modulations.

4.2 PHASE MODULATION

Digital phase modulation generally changes the phase of the carrier to a number of different phase angles. In the simplest case—binary phase shift keying (BPSK)—the carrier phase is either 0 deg or 180 deg. We can even use different phase angles to represent groups of bits. For example, we can use one of four different phase angles (i.e., 45 deg, 135 deg, −45 deg, and −135 deg) to represent two bits at a time. This is called quadriphase or quadrature phase shift keying (QPSK). This idea can be extended further by using eight phase angles to represent groups of three bits at a time (8-PSK), or by using 16 different phase angles to represent four bits at a time (16-PSK).

4.2.1 Binary Phase Shift Keying (BPSK)

The simplest form of digital phase modulation is BPSK. Binary phase shift keying is often used in applications such as garage door openers, direct-sequence spread-

spectrum (DSSS) transceivers, and in pulse compression radars. In BPSK, the phase of the carrier is either 0 deg or 180 deg, depending on the input data. The phase domain of the carrier is shown in Figure 4.2. We can write the modulated carrier as

$$s_i(t) = A \cdot \cos[\omega_{RF} t + \varphi_i(t)] \tag{4.1}$$

where the instantaneous phase angle, $\varphi_i(t)$, is given by

$$\varphi_i(t) = \pi \cdot i;\ i = 0, 1 \tag{4.2}$$

and A is the amplitude of the carrier. In general, a logic "0" corresponds to a phase inversion of the carrier, while a logic 1 corresponds to no phase inversion, though this convention doesn't always hold. In general, a double-balanced mixer is used to modulate the carrier.

Binary phase shift keying is often referred to as an *antipodal* modulation, where each of the two binary waveforms is the negative of the other. In BPSK, the binary waveforms are the unshifted carrier or the inverted carrier. This can be more easily seen in time domain response, shown in Figure 4.3. As the figure shows, the binary bit stream inverts the phase of the carrier back and forth 0 deg or 180 deg. Each inversion of the carrier causes a sharp transition in the time domain response. As we'll see later, these transitions produce a very wide transmitted spectrum. In general, the carrier frequency is much higher than the modulation frequency, or data rate.

Figure 4.4 shows a block diagram for a typical BPSK modulator. An oscillator produces an unmodulated carrier that is fed to a double-balanced mixer. The non-return-to-zero (NRZ) data is mixed with the carrier to form the desired BPSK signal. As stated earlier, BPSK produces sharp transitions in the time domain, which result in a very wide transmitted spectrum. This is undesirable in most cases. While the transmitted signal could be directly filtered itself at its output, this is usually not done. A very narrow high-Q filter would be required and the center frequency of the filter would need to be changed for each channel frequency. A much better way

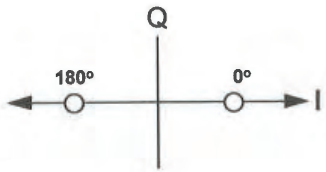


Figure 4.2 Phase domain of carrier for BPSK.

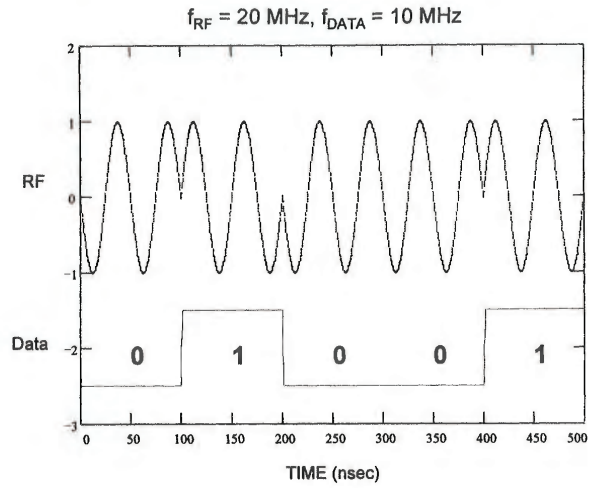


Figure 4.3 BPSK time domain waveforms.

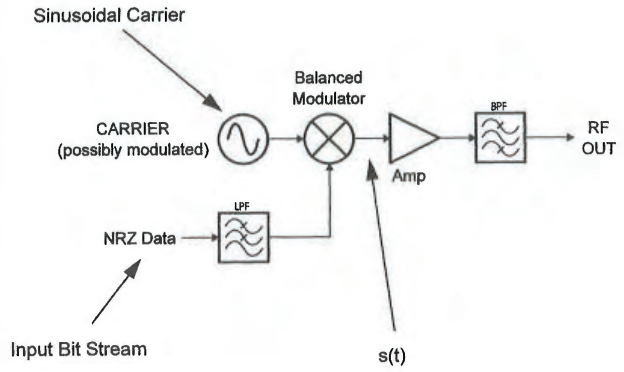


Figure 4.4 BPSK modulator.

to narrow the transmitted output spectrum is by filtering the baseband modulation signal itself. This can be done by inserting a lowpass filter between the NRZ data input and the double-balanced modulator. As we'll see later, some care must be taken when doing this in order to ISI and spectral regrowth.

In most cases, the signal coming out of the modulator is too weak to be sent to the antenna directly, so an amplifier is usually used to increase the transmitted signal level. Although quite clean near the fundamental frequency, most RF amplifiers produce harmonics of the signal at integer multiples of the fundamental transmitter frequency. For this reason, a final filter is used to eliminate these out-of-band signals. Often, this filter doubles as a "roofing filter" for the receiver, suppressing strong out-of-band received signals prior to downconversion.

As mentioned above, double-balanced modulators are often used in BPSK systems. A common form of such a modulator is the diode ring mixer, shown in Figure 4.5. In general, the LO waveform is at the RF carrier frequency and is overdriven, with a typical power level of +6 dBm to +7 dBm. As Figure 4.5 shows, this causes current to flow in and out to the mixer output at the same frequency as the LO. Diodes shown in solid lines are turned "on" and can be modeled as low-value resistors (almost short circuits). Diodes shown with dotted lines are "off" and can be modeled as small capacitors (almost open circuits). The modulation signal changes the sign of this output current in conjunction with the input data. The overall action of the mixer is equivalent to multiplying a square wave, $r(t)$, at the LO frequency with the modulation signal, $m(t)$.
This can be written as

$$r(t) = \frac{4}{\pi}\left[\cos(\omega_{RF}t) + \frac{1}{3}\cdot\cos(3\omega_{RF}t) + \frac{1}{5}\cdot\cos(5\omega_{RF}t) + \ldots\right] \tag{4.3}$$

$$m(t)\cdot r(t) = \frac{4}{\pi}\left[m(t)\cdot\cos(\omega_{RF}t) + \frac{1}{3}m(t)\cdot\cos(3\omega_{RF}t) + \frac{1}{5}m(t)\cdot\cos(5\omega_{RF}t) + \ldots\right] \tag{4.4}$$

Recall that an ideal square wave has energy only at the fundamental frequency and odd harmonics. To see where the frequency components of the modulated signal lie, we take the Fourier transform of both sides of the equation. Let $x(t)$ be a time domain signal. Its Fourier transform is given by

$$S(\omega) = \mathcal{F}[s(t)] = \int_{-\infty}^{\infty} s(t)\cdot e^{-j\omega t}dt \tag{4.5}$$

If we now apply the transform to the modulated waveform

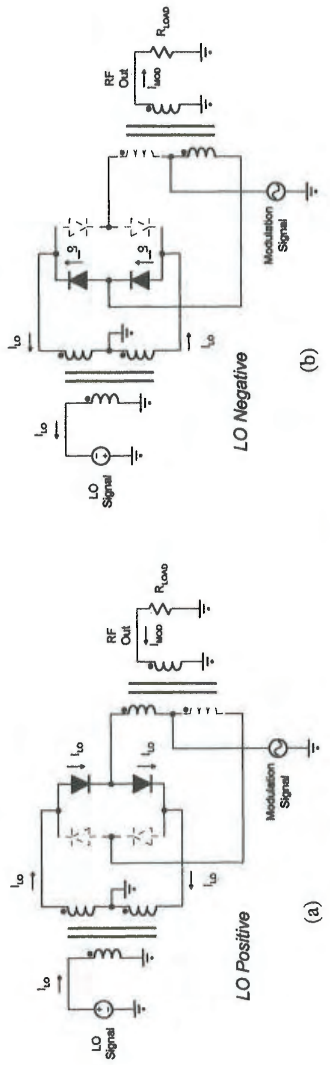


Figure 4.5 Diode-based BPSK modulator: (a) LO is positive, (b) LO is negative.

$$S_{MOD}(\omega) = \mathcal{F}[r(t) \cdot m(t)] = \int_{-\infty}^{\infty} r(t) \cdot m(t) \cdot e^{-j\omega t} dt \tag{4.6}$$

$$S_{MOD}(\omega) = \int_{-\infty}^{\infty} r(t) \cdot \left\{ \frac{4}{\pi} \left[\cos(\omega_{RF} t) + \frac{1}{3}\cos(3\omega_{RF} t) + \ldots \right] \right\} \cdot e^{-j\omega t} dt \tag{4.7}$$

Performing the integration, gives

$$S_{MOD}(\omega) = M(\omega - \omega_{RF}) + M(\omega + \omega_{RF}) + M(\omega - 3\omega_{RF}) + M(\omega + 3\omega_{RF}) + \ldots \tag{4.8}$$

where $M(\omega)$ is the Fourier transform of the baseband NRZ data signal. The action of the modulator causes the baseband data spectrum to be shifted to odd multiples of the LO frequency. Of course, the term we are interested in is the fundamental (and its negative frequency component), so after bandpass filtering we get

$$S'_{MOD}(\omega) = M(\omega - \omega_0) + M(\omega + \omega_0) \tag{4.9}$$

where $S'_{MOD}(\omega)$ is the filtered version of the signal.

Other circuit configurations can be used to perform the BPSK mixing function. Figure 4.6 shows the schematic of a Gilbert-cell based double-balanced modulator. This circuit is very convenient for monolithic IC implementation. The active devices can either be bipolar transistors (BJT or HBT) or field-effect transistors (MOSFET or MESFET). In this circuit, the LO signal is applied to the upper transistor pairs. These pairs act as current switches and steer the current to one side or the other of the output. In the example shown, a center-tapped transformer is used, though this could be replaced by a transmission-line type transformer. In fact, each output of the mixer (collectors of Q1 through Q4) could be singly terminated as well. This results in a simpler interface, but suffers from the disadvantages of lower output power level and a strong dc component in the modulated signal. In some monolithic implementations, an active push-pull, Class AB output stage is used. More details on the design of this circuit are included in Chapter 5.

The Gilbert-cell-based modulator has a number of advantages. All of the active devices of the Gilbert-cell can be fabricated on a monolithic IC. In most cases, the transformer is off-chip, though it is not uncommon to put the transformer on-chip when GaAs or other semi-insulating substrate IC processes are used. The LO input to the upper pair need only be large enough to switch the transistors on and off. In general, 100 mV (about -10 dBm) is sufficient when bipolar devices are used.

The operation of the Gilbert-cell-based modulator is as follows. The modulation signal is applied to the lower transistor pair, formed by Q5 and Q6. In the case of unfiltered BPSK, this input can also be thought as a switching input. However, in almost all cases, the BPSK signal is filtered at baseband (via a lowpass filter) before

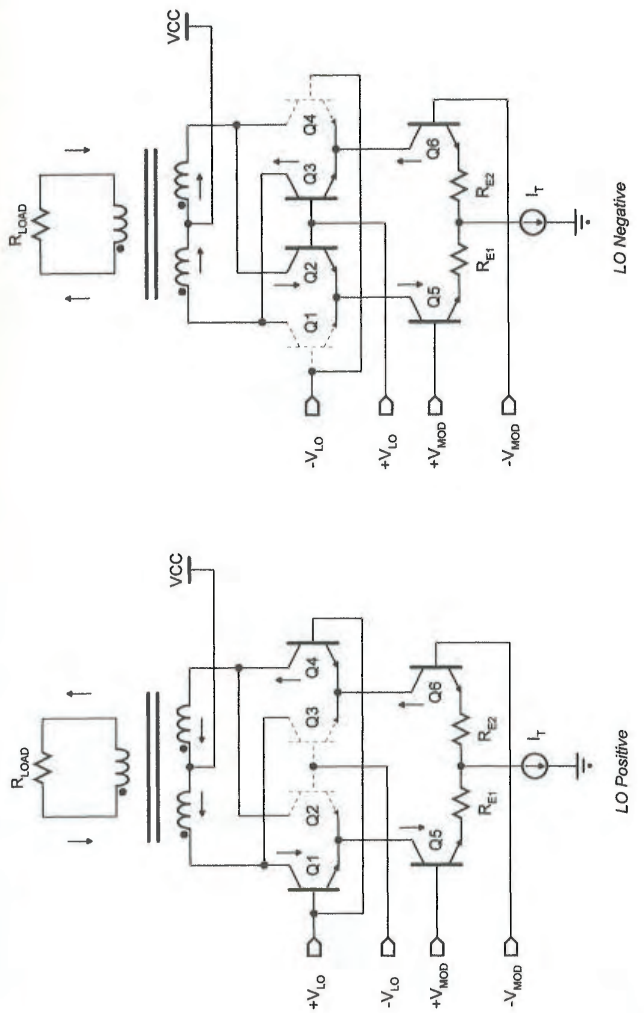


Figure 4.6 Gilbert-cell BPSK modulator (for monolithic IC applications).

being applied to the modulator. For this reason, the modulation input needs to be linear. This is accomplished by using emitter degeneration resistors in the lower pair. Unfortunately, the addition of degeneration resistors also increases the noise figure of the mixer to, typically, 10 to 15 dB. For this reason, the Gilbert-cell mixer is seldom used as the first mixer in a low-noise receiver. However, most modulators perform an upconverting function where noise figure is much less of a problem.

As stated earlier, the upper pair acts as current switches, configured in a cross-coupled fashion to change the polarity of the currents from the lower pair's collectors. Depending on which of the two LO inputs is greater, current is applied to the center-tapped output transformer either in-phase or out-of-phase.

A diagram of a typical coherent BPSK demodulator is shown in Figure 4.7. Basically, the action of the demodulator is the reverse of the modulator. The modulated signal is applied to a double-balanced mixer, which is mixed with an LO having the same frequency and phase as the modulated signal itself. Coherent demodulation, as opposed to noncoherent demodulation, implies that the exact phase and frequency of the modulated signal is known at the receiver. Circuits that produce this local carrier reference are called carrier recovery circuits. We can write the received, modulated signal as

$$r_i(t) = B \cdot \cos[\omega_{RF}t + \varphi_i(t)] \tag{4.10}$$

where B is the amplitude of the received signal and $\varphi_i(t)$ is the phase of the transmitted signal—either 0 deg or 180 deg. Multiplying the LO signal by the received signal gives

$$s_i(t) = B \cdot \cos[\omega_{RF}t + \varphi_i(t)] \cdot \cos(\omega_{RF}t) \tag{4.11}$$

$$s_i(t) = \frac{B}{2} \cdot \{\cos[2\omega_{RF}t + \varphi_i(t)] + \cos[\varphi_i(t)]\} \tag{4.12}$$

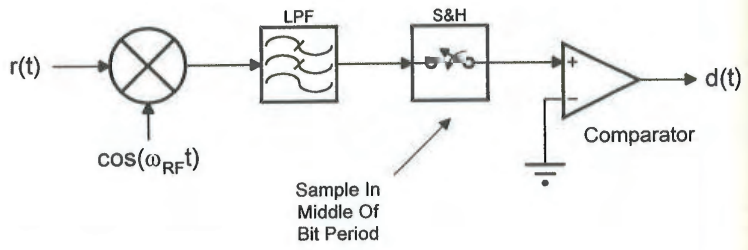


Figure 4.7 BPSK demodulator.

Lowpass filtering removes the component at twice the RF frequency, leaving

$$s_i(t) = \frac{B}{2} \cdot \cos[\varphi_i(t)] \tag{4.13}$$

If $\varphi_i(t)$ is 0 deg, then $s_i(t)$ is 1. If $\varphi_i(t)$ is 180 deg, then $s_i(t)$ is –1. Thus, the original binary NRZ signal is recovered by the demodulator. In Figure 4.7, the decision as to whether the received signal was 1 (or binary 1) or –1 (binary 0) can be performed by an analog comparator. In most cases, a special-purpose circuit, called a *bit (or data) slicer*, is used. The bit slicer tracks the dc level of the recovered data so as to sample the signal at precisely the middle of the data period, at the point of maximum SNR, providing for minimum P_E.

The NRZ data signal resembles a square wave. However, unlike a square wave, which alternates periodically between binary 1 and 0, the NRZ data signal alternates between binary 1 and 0 in a random fashion. This characteristic means that the NRZ signal's power spectral density is smooth (rather than consisting of discrete spectral lines), with a shape that resembles a sinx/x shape. The power spectral density of the NRZ baseband signal is

$$s(f) = 2(T_b A)^2 \frac{\sin^2(\pi f T_b)}{(\pi f T_b)^2} \tag{4.14}$$

where T_b is the bit period, A is the amplitude of the signal, and f is the frequency in hertz.

Recall that the modulated signal is produced by multiplying the baseband data signal by a square wave at the RF frequency. If we filter out all the terms except the fundamental, we are left with the original baseband data spectrum shifted to the RF frequency. This is given by

$$S_{TX}(f) = \frac{1}{2} \cdot [S(f - f_{RF}) + S(f + f_{RF})] \tag{4.15}$$

$$S_{TX}(f) = (AT_b)^2 \cdot \left\{ \frac{\sin^2[\pi(f - f_{RF})T_b]}{[\pi(f - f_{RF})T_b]^2} + \frac{\sin^2[\pi(f + f_{RF})T_b]}{[\pi(f + f_{RF})T_b]^2} \right\} \tag{4.16}$$

Figure 4.8 shows a plot of the power spectral density for an unfiltered BPSK signal. The data rate is 10 Mbps and the carrier frequency is 80 MHz. Only one side of the spectrum is shown for greater detail. Notice that the sharp transitions in the time domain lead to a relatively wide power spectral density that rolls off quite slowly. The first null occurs at a frequency equal to the (data rate)$^{-1}$ away from the carrier. The amplitude of the first lobe is only 13-dB down from its value

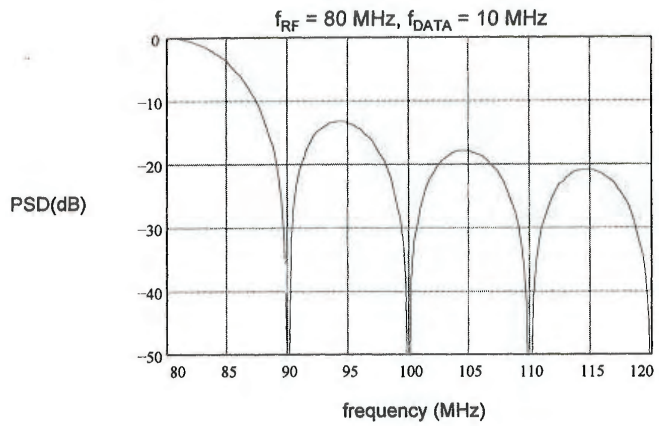


Figure 4.8 BPSK power spectral density.

at the carrier frequency. A good rule of thumb is that 99% of the transmitted power of an unfiltered BPSK signal lies within a bandwidth equal to the bit rate. As we will see later, filtering the signal can greatly reduce the transmitted signal bandwidth. However, care must be taken in order to avoid ISI and spectral regrowth. The theoretical spectral efficiency of BPSK is 1 bit/sec/Hz. In practice, a value of 1.4 bits/sec/Hz can be achieved using Nyquist filtering techniques (discussed in Section 4.2.5).

Another important parameter of any digital modulation is how the P_E varies with SNR. Figure 4.9 shows a BPSK received signal with and without noise. The demodulator decides whether the recovered signal was 1 or −1 by sampling the waveform at the center of the bit period. This is easy to do when there is no noise added to the signal, as shown in Figure 4.9(a). When noise has been added to the signal, as shown in Figure 4.9(b), it is easier to make an error. An error would occur if the noise level caused the signal to be above or below zero at the center of the bit period.

Before we can quantify how many bit errors occur for BPSK in the presence of noise, we must first have a means of quantifying the noise itself. In general, thermal noise is Gaussian in nature, having random amplitude values with a relatively flat frequency spectrum. Noise is added to the signal as it passes through the channel. We refer to such noise as additive white Gaussian noise (AWGN). A plot of AWGN is shown in Figure 4.10(a).

It is often convenient to assume that the noise power spectral density extends without limit in frequency in both directions. Of course, this can't really happen;

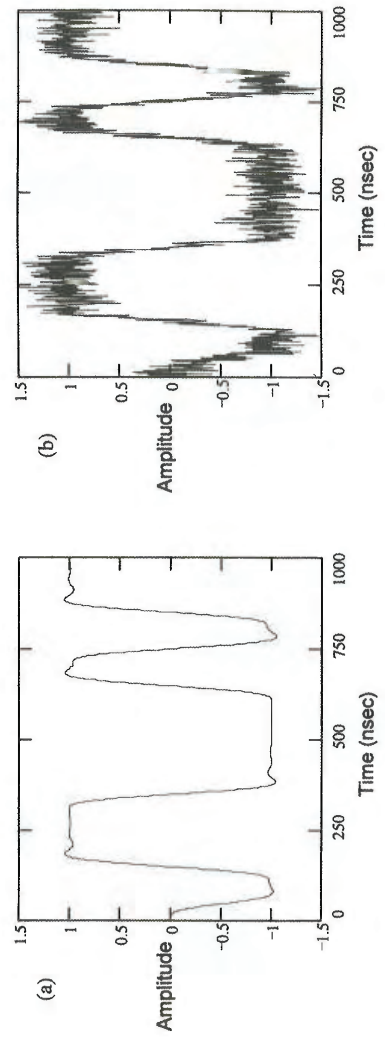


Figure 4.9 Demodulated BPSK waveforms: (a) without noise, (b) with noise.

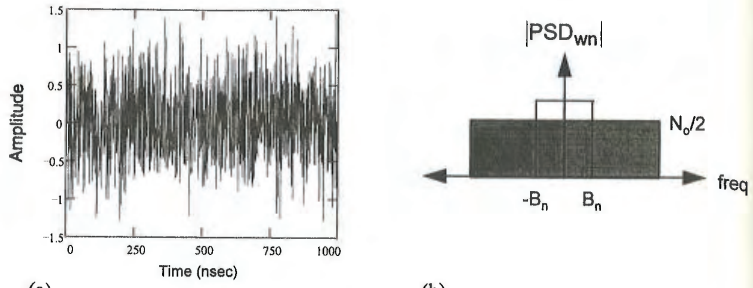


Figure 4.10 Additive white Gaussian noise: (a) noise time domain waveform, (b) noise power spectral density.

else, the noise would have infinite power. In reality, thermal noise does tend to fall off as the frequency approaches 10^{12} to 10^{14} Hz. For most of the frequencies we normally deal with in digital communications, assuming a flat power spectral density is a good approximation. A diagram of white noise versus frequency is shown in Figure 4.10(b). If we integrate the noise over a bandwidth ranging from $-B_n$ to $+B_n$ (where B_n is the one-sided noise bandwidth of the receiver), then we get the rms power of the noise. With an rms noise power equal to N, the noise power spectral density across the band must be equal $N_0/2$. This can be written mathematically by

$$N = \left(\frac{N_0}{2}\right) \cdot (2B_n) \tag{4.17}$$

Rearranging the equation, we can relate the noise power spectral density to the noise rms noise power by

$$N_0 = \frac{N}{B_n} \tag{4.18}$$

Figure 4.11 shows a plot of the probability of making a bit error versus a quantity called E_b/N_0 for BPSK. The quantity E_b/N_0 is the ratio of the *energy per bit* over the noise power spectral density. To see how E_b/N_0 is related to SNR, we can write

$$E_b = S \cdot T_b \tag{4.19}$$

where S is the average signal power of the carrier and T_b is the bit period. This gives

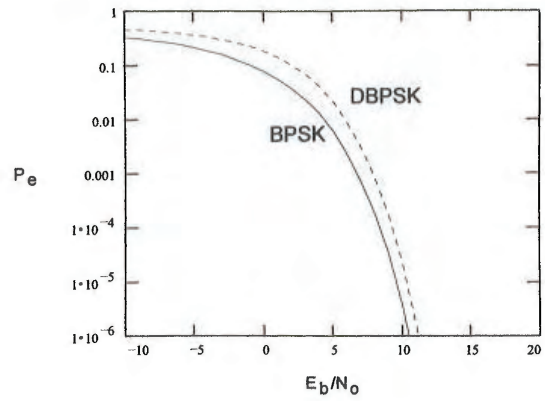


Figure 4.11 BPSK probability of bit error.

$$\frac{E_b}{N_0} = \frac{S \cdot T_b}{N/B_n} = \frac{S}{N} \cdot B_n T_b = \frac{S}{N} \cdot \frac{B_n}{f_b} \tag{4.20}$$

where f_b is the bit rate of the data. As (4.20) shows, E_b/N_0 is equal to S/N, or SNR, *when the noise bandwidth of the receiver is equal to the data rate.* This is a theoretical result, since the noise bandwidth of the receiver needs to be somewhat wider than the bit rate to avoid ISI in the received waveform. A practical rule of thumb is that a BPSK signal can fit into a bandwidth equal to 1.4 times the bit rate by using special filtering techniques. This is significantly narrower than the spectrum of the BPSK signal shown in Figure 4.8. These filtering techniques will be discussed in more detail in Section 4.2.5.

E_b/N_0 is often used to relate probability of bit error because it is a dimensionless quantity. By using (4.20), we can relate this quantity back to SNR. Two curves are shown in Figure 4.11. The curve labeled BPSK is the P_e for coherently demodulated BPSK. The curve labeled DBPSK is for differentially encoded/decoded BPSK. In DBPSK, we compare the received signal with a delayed version of itself. This is shown in Figure 4.12, where both the DBPSK modulator and DBPSK demodulator are shown. Notice that the data must be differentially encoded prior to modulation. If the original NRZ data stream is

$$0, 1, 1, 0, 1, 1, 1, 0, 0 \ldots \tag{4.21}$$

then the encoded NRZ data signal (assuming the initial bit was a 0) is

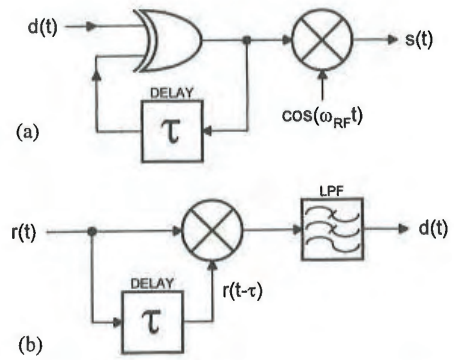


Figure 4.12 Differentially encoded/decoded BPSK (DBPSK): (a) DBPSK transmitter, (b) DBPSK receiver.

$$0, 1, 0, 0, 1, 0, 1, 1, 1, \ldots \tag{4.22}$$

In DBPSK, we can write the received signal as before

$$r_i(t) = B \cdot \cos[\omega_{RF}t + \varphi_i(t)] \tag{4.23}$$

The delayed version of the signal can be written

$$r_i(t - T_b) = B \cdot \cos[\omega_{RF}(t - T_b) + \varphi_i(t - T_b)] \tag{4.24}$$

Multiplying the delayed and undelayed signals together yields

$$r_i(t - T_b) \cdot r_i(t) = B \cdot \cos[\omega_{RF}(t - T_b) + \varphi_i(t - T_b)] \cdot B \cdot \cos[\omega_{RF}(t) + \varphi_i(t)] \tag{4.25}$$

After some algebra and lowpass filtering, we are left with

$$[r_i(t - T_b) \cdot r_i(t)]_{LPF} = \frac{B^2}{2} \cdot \cos[\varphi_i(t - T_b) - \varphi_i(t)] \tag{4.26}$$

This result also assumes that there are an integral number of RF carrier periods in one bit period.

If the last angle and the present angle are both the same, then their difference is 0 and the filtered output is 1. If the present angle and the last angle are different, then their difference is 180 deg, and the filtered output is −1. Since a delayed

version of the received signal, complete with added noise, is used as a reference for demodulation, we would expect that P_e performance for DBPSK would be worse than that of coherently demodulated BPSK. This is in fact the case, as Figure 4.11 shows.

Another way to view degradation due to noise is by picturing the transmitted signal as a vector. This is shown in Figure 4.13. Noise can be visualized as a random vector that is added to the tip of the received signal. In the case of BPSK, when the noise vector has greater amplitude than the signal vector—but in the opposite direction—a bit error has occurred. This way of viewing things becomes more useful when we talk about higher order modulations, such as QPSK.

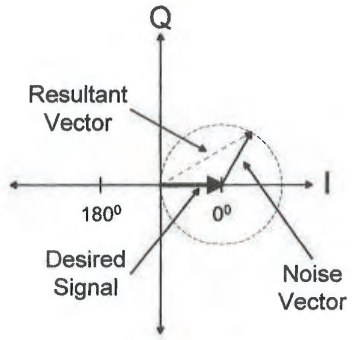
4.2.2 Quadriphase (or Quadrature) Phase-Shift Keying (QPSK)

In QPSK, we introduce the concept of *symbols*. A symbol is an electrical waveform that can represent one or more bits. To begin with, let each of the icons in Figure 4.14 represent a group of 2 bits. In Figure 4.14(a), the dog represents the bit pattern 00. The cat represents 01, the elephant represents 10 and, finally, the bear represents the pattern 11. Let the transmitted bit stream be

$$0, 1, 0, 0, 1, 1, 1, 0, \ldots \tag{4.27}$$

To send a stream of symbols corresponding to these bits, we first separate the bits into groups of two

$$0\ 1, 0\ 0, 1\ 1, 1\ 0 \tag{4.28}$$



$$|P_{Noise}| \leq |P_{Carrier}|$$

Figure 4.13 Resultant BPSK signal vector with noise.

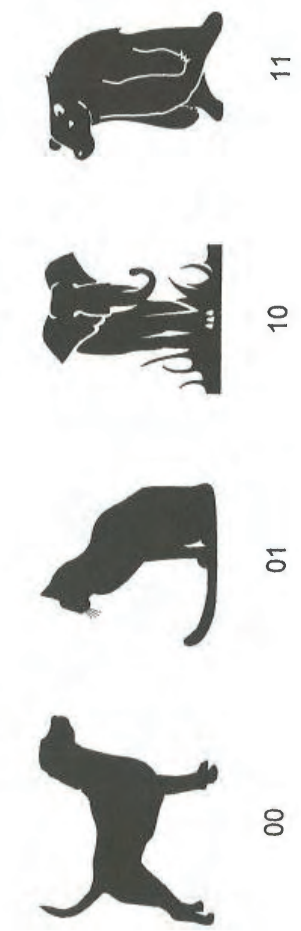




Figure 4.14 Using symbols to represent groups of two bits each: (a) symbol assignment for groups of bits, (b) symbol sequence for the bit stream 01001110.

This grouping results in the symbol stream of cat, dog, bear, and, finally, elephant, as shown at the bottom of the figure.

Notice that by using symbols to represent groups of one or more bits, we can reduce the rate at which the symbols are sent. In other words, if our bit rate is 1 Mbps, then by using symbols to represent groups of 2 bits each, the symbol rate is one-half the bit rate, or 500 Kbps.

In the example above, where animals were used, it is easy to tell which symbol was sent since each looked distinctly different from the others. For example, it is hard to confuse the picture of the bear with that of the dog. But, what if we decided to use pictures of different types of dogs. Can most people tell the difference between the silhouette of a dalmatian from that of a labrador retriever? It might be difficult, especially if the pictures were fuzzy or out of focus. This illustrates the point that each symbol should also be distinctly different from all the others, so that we can still decide what symbol was sent in the presence of noise.

So, what kind of symbols can we use in digital communications? If we are doing phase modulation, then we might choose different phase angles. This is exactly what we do in QPSK. Figure 4.15 shows the phase domain of a QPSK signal. Each group of two bits is represented by a different phase angle, 90-deg apart from the next. Notice that each angle lies 45-deg off the in-phase, or I, and quadrature, or Q, axes. Again, as in the case of BPSK, we assume that a logic 0 corresponds to a -1.

We can write the QPSK signal as

$$s_i(t) = A \cdot \cos[\omega_{RF}t + \varphi_i(t)] \tag{4.29}$$

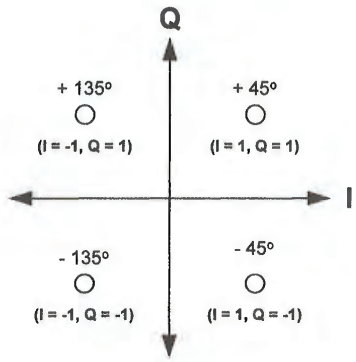


Figure 4.15 Phase domain of carrier for QPSK (gray-coded).

where A is the carrier amplitude, i is an integer, and $\varphi_i(t)$ is the instantaneous phase angle of the modulated signal with respect to the unmodulated carrier. The value $\varphi(t)$ can be written as

$$\varphi_i(t) = \frac{\pi}{4} \cdot (2i + 1) \tag{4.30}$$

where $i = 0, 1, 2, 3$.

As mentioned earlier, each symbol represents a group of two bits. Therefore, the symbol rate of QPSK is half the bit rate of BPSK. This means that the spectrum is only 1/2 as wide as that of BPSK for a given bit rate. QPSK is used in applications such as early telephone modems, satellite communications, the global positioning system (GPS) and in some forms of code division multiple access (CDMA).

One way to generate QPSK is by simply changing the phase angle of the carrier using a brute-force approach as shown in Figure 4.16. Techniques similar to this are used in active-array radars to steer the transmitted beam pattern. Unfortunately, this technique is not as useful in digital communication systems. Since the phase shifters operate at the carrier frequency, errors can be introduced into the transmitted phase via back-coupling from the powerful transmitted signal back into the phase shifting circuitry itself (phase pulling). Also, the phase shifters are open-loop devices,

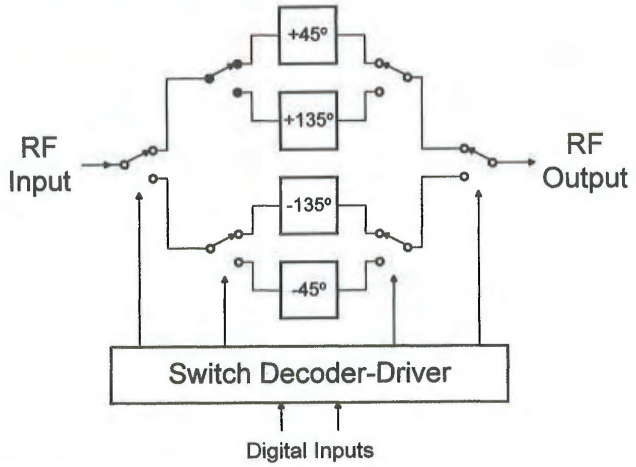


Figure 4.16 Brute-force QPSK modulator.

which tend to shift their phase with respect to one another over temperature. However, the most significant disadvantage of the brute-force phase shifter is that the phase transitions are abrupt. As we shall see, filtering the digital inputs to the QPSK modulator can greatly reduce the transmitted output spectrum.

A more common implementation used to perform QPSK modulation is shown in Figure 4.17. The serial data stream is demultiplexed into even and odd components via the serial-parallel converter (or demultiplexer). Each data component (now at one-half the bit rate) is applied to a set of double-balanced modulators. As will be discussed later, optional lowpass filters can be used to bandlimit the transmitted spectrum. The set of double-balanced mixers form what is known as a quadrature modulator. The quadrature modulator can vary both the phase and amplitude of the transmitted signal, though for QPSK, we only care about the phase. Finally, the outputs from each double-balanced mixer are added together and fed through an amplifier to a bandpass filter, which removes any harmonics of the modulated signal.

Figure 4.18 shows a simplified block diagram of the quadrature modulator. To understand the operation of the quadrature modulator, we can write the in-phase and quadrature components of the modulated signal in their exponential form

$$I \cdot \cos(\omega_{RF}t) = I \cdot \left[\frac{e^{j\omega_{RF}t} + e^{-j\omega_{RF}t}}{2}\right] \tag{4.31}$$

$$Q \cdot \sin(\omega_{RF}t) = Q \cdot \left[\frac{e^{j\omega_{RF}t} - e^{-j\omega_{RF}t}}{2j}\right] \tag{4.32}$$

The output of the quadrature modulator is then

$$s(t) = I \cdot \cos(\omega_{RF}t) + Q \cdot \sin(\omega_{RF}t) = I \cdot \left[\frac{e^{j\omega_{RF}t} + e^{-j\omega_{RF}t}}{2}\right] + Q \cdot \left[\frac{e^{j\omega_{RF}t} - e^{-j\omega_{RF}t}}{2j}\right] \tag{4.33}$$

Rearranging the terms in the equation gives

$$s(t) = \frac{1}{2} \cdot [(I - jQ) \cdot e^{j\omega_{RF}t} + (I + jQ) \cdot e^{-j\omega_{RF}t}] \tag{4.34}$$

If we let

$$I - jQ = \sqrt{I^2 + Q^2} \cdot e^{j\varphi} \tag{4.35}$$

$$I + jQ = \sqrt{I^2 + Q^2} \cdot e^{-j\varphi} \tag{4.36}$$

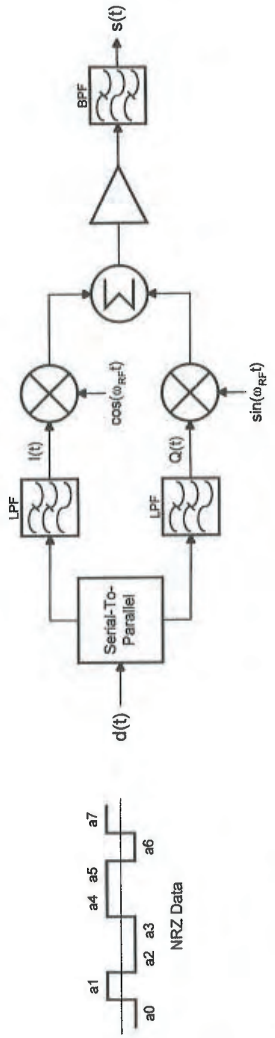


Figure 4.17 QPSK modulator.

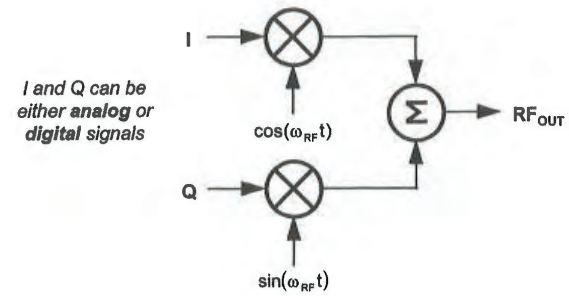


Figure 4.18 Quadrature modulator.

where

$$\varphi = \tan^{-1}(Q/I) \tag{4.37}$$

then we can write

$$s(t) = \sqrt{I^2 + Q^2} \cdot \left[\frac{e^{j\omega_{RF}t+\varphi} + e^{-j\omega_{RF}t-\varphi}}{2}\right] \tag{4.38}$$

$$s(t) = \sqrt{I^2 + Q^2} \cdot \cos(\omega_{RF}t + \varphi) \tag{4.39}$$

We see that the quadrature modulator can be used to modulate both the amplitude and the phase of the carrier. If I and Q are chosen to be ±1, then the amplitude of the modulated signal is a constant $\sqrt{2}$. Notice also that having I and Q equal to ±1, all phase angles lie along 45-deg lines off the axes. In general, we assume that the carrier amplitude is normalized to a value of 1, therefore we assume that both I and Q inputs can range between $\pm 1/(\sqrt{2})$, or ±0.707. We can also use the quadrature generator to perform modulations that require both amplitude and phase variation. An example of such a modulation is quadrature amplitude modulation (QAM), in which each symbol has a different phase and/or amplitude from the other symbols.

A number of circuits can be used to generate the in-phase and quadrature components of the carrier, and some examples are shown in Figure 4.19. In Figure 4.19(a), two op amp allpass filters are used to generate two phase-shifted signals having constant amplitude. The allpass filters are designed such that there is a region in frequency where the relative phase shift at the output of the two filters is 90 deg,

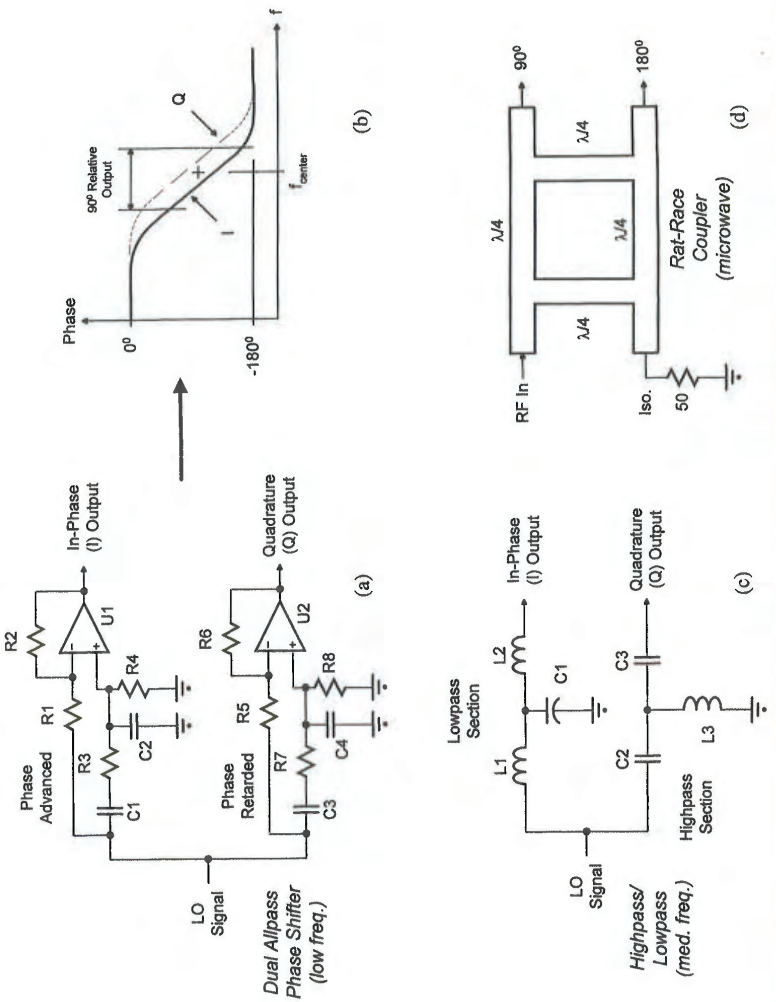


Figure 4.19 Methods of generating in-phase and quadrature LO references: (a) OPAMP allpass filter phase shifter I/Q reference, (b) output phase of OPAMP allpass filter, (c) lumped-element highpass/lowpass I/Q reference, (d) rat-race coupler I/Q reference.

as shown in Figure 4.19(b). With two second-order filters as shown, the phase difference can be maintained to an accuracy of a few degrees over a decade in frequency. This approach is limited to relatively low frequencies where the op amps work well.

Figure 4.19(c) shows a higher frequency version of the differential phase shifter. In this approach, a lowpass filter is used to provide a relative −45-deg phase shift and a highpass filter is used to provide a relative +45-deg phase shift at the center of the frequency band. The lowpass/highpass approach can be used to much higher frequencies than the op amp approach, but is limited to a smaller useful bandwidth due to the fact that the amplitude of each filter changes with frequency.

An approach that is used at microwave frequencies is shown in Figure 4.19(d). A rat-race coupler is used to provide a relative 90-deg phase shift at each output. In general, this approach works over about an octave in frequency. As the frequency is lowered, the size of the rat-race coupler gets physically large, limiting its usefulness to relatively high frequencies. The phase shifters of Figure 4.19(c, d) are discussed in more detail in Chapter 5.

The quadrature signals can also be generated by digital means. Figure 4.20 shows two methods by which this can be done. In Figure 4.20(a), an input at four times the desired LO frequency is applied to a divide-by-4 circuit constructed using two D-type flip-flops. The divider action ensures that each output of the circuit is a perfect square wave, shifted by one period of the input signal (or 90 deg). The digital circuit can be used from dc up to the highest frequency of operation of the divider. Nowadays, digital logic circuits can operate well above 1 GHz, even tens of gigahertz. In order to ensure the circuit tracks out all errors over temperature and process variations, it is essential that every node in the divider is equally loaded. It is not uncommon to use dummy gates to equally load each output.

The major disadvantage of the divide-by-4 circuit is that a signal at four times the desired LO frequency needs to be applied as an input. If it can be guaranteed that the input is perfectly symmetrical (i.e., has a 50% duty cycle), then the circuit can be connected as shown in Figure 4.20(b) where one of the flip-flops is driven out-of-phase from the other. This approach requires the input signal to be only twice the desired LO frequency.

We've discussed how to generate the required in-phase and quadrature versions of the LO signal, but how are the two double-balanced modulators and summing junction implemented? Figure 4.21 shows a schematic of a dual Gilbert-cell mixer. Recall from earlier discussions on BPSK modulation, that the Gilbert-cell mixer can be used as a double-balanced modulator. Since the outputs of the Gilbert-cell mixer are really currents, we can form the summing junction by simply "wire OR-ing" the outputs together. The dual Gilbert-cell implementation is often used in monolithic versions of the quadrature modulator.

As shown in Figure 4.17, the input to the quadrature modulator is split into two parallel bit streams, each at one-half the rate of the original serial data stream.

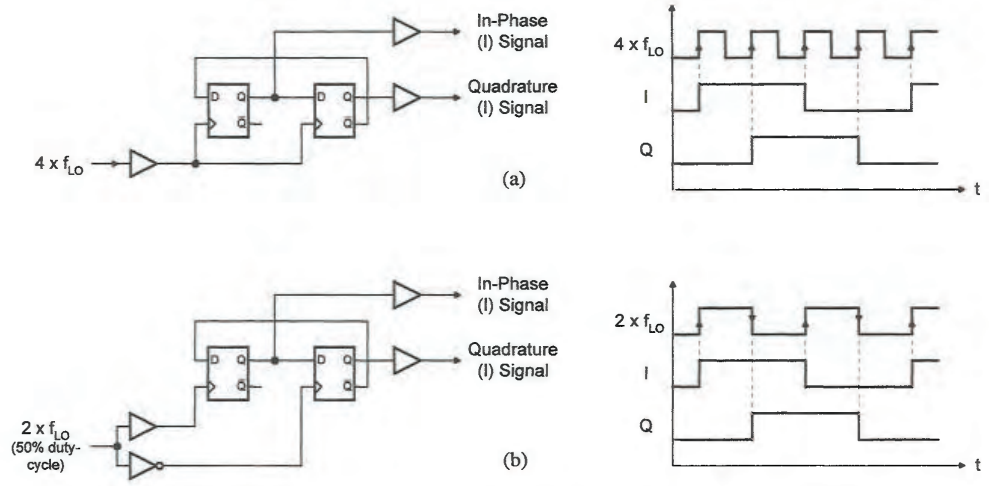


Figure 4.20 Digital *I/Q* references: (a) 4X LO reference (no duty cycle restrictions), (b) 2X LO reference (50% duty cycle required).

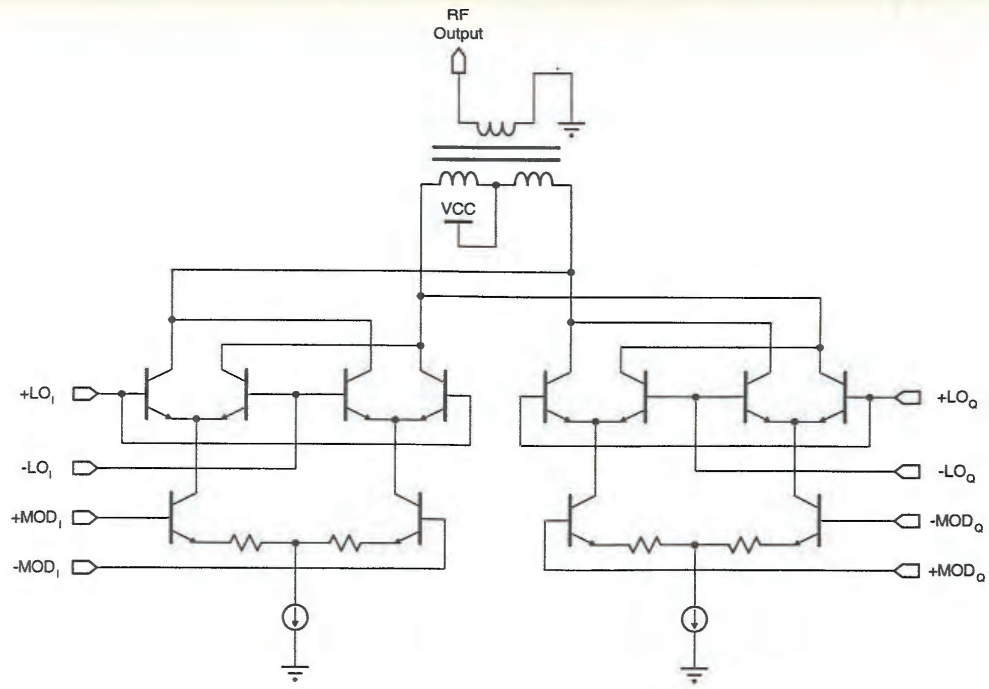


Figure 4.21 Dual Gilbert-cell QPSK modulator.

Figure 4.22 shows the block diagram for a digital demultiplexer that does this. The first two D-type flip-flops take snapshots of alternating bits of the serial data. The third flip-flop adds a 1-bit delay to the *I* data to realign the two bit streams.

Figure 4.23 shows the time domain response for QPSK. As the figure shows, the phase of the carrier can be changed 0 deg, ±90 deg, or 180 deg from its present

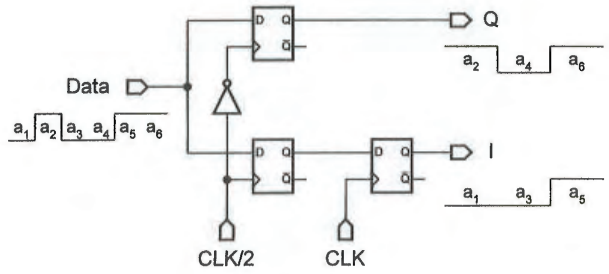


Figure 4.22 Data demultiplexer (serial-to-parallel) for QPSK.

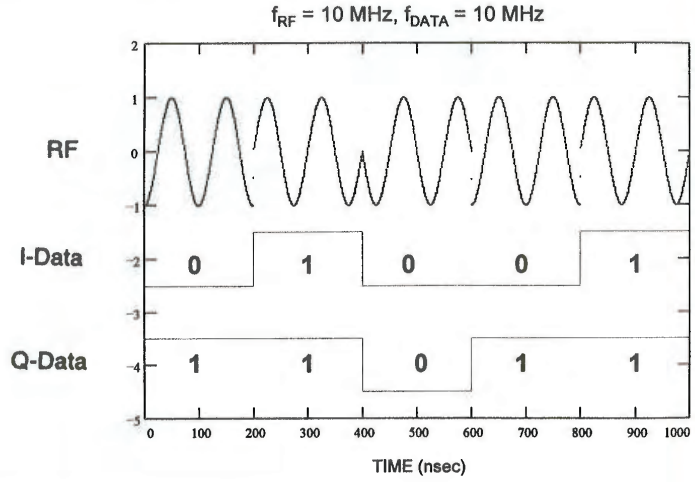


Figure 4.23 QPSK time domain waveforms.

value. The two bit streams in the bottom of the plot correspond to the demultiplexed *I* and *Q* data streams. In this example, the input data bits are unfiltered. This causes abrupt changes in phase in the modulated carrier, much in the same fashion as BPSK. Figure 4.24 shows the transmitted power spectral density of a typical QPSK signal at a carrier frequency of 80 MHz and a data rate of 10 Mbps, along with a BPSK spectrum having the same parameters. Not surprisingly, the QPSK spectrum is only half as wide as the BPSK spectrum. Notice however, that the spectrum still has sharp nulls and rolls off very slowly with relatively large sidelobes. In theory, the spectral efficiency of QPSK is 2 bits/sec/Hz, which is better than BPSK by a factor of two.

As we might expect, the QPSK demodulator is simply the reverse of the modulator. Figure 4.25 shows a block diagram of a typical QPSK demodulator, which is for all intents and purposes simply two BPSK demodulators. The received signal is multiplied by cosine and sine versions of the local carrier. The block labeled C/R contains the carrier recovery circuitry, which provides a local in-phase reference precisely aligned in phase to the transmitted carrier. Of course, the quadrature signal is generated by simply shifting the in-phase signal by 90 deg. The block labeled STR contains the symbol timing recovery circuitry, which extracts the data clock and samples the lowpass filtered outputs at precisely the optimum position within the bit period. Finally, a data multiplexer reinterleaves the data back into the serial bit stream.

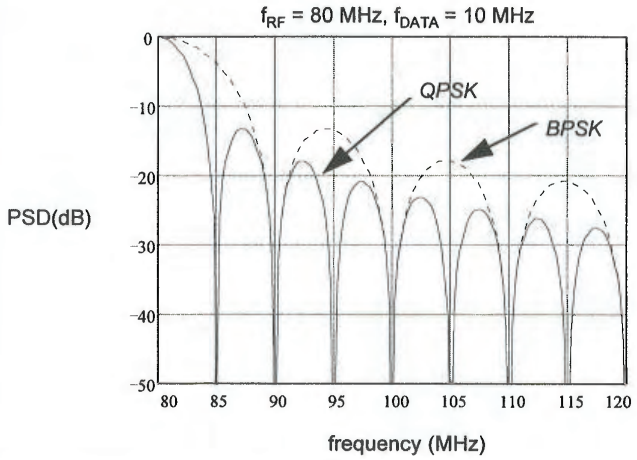


Figure 4.24 QPSK power spectral density.

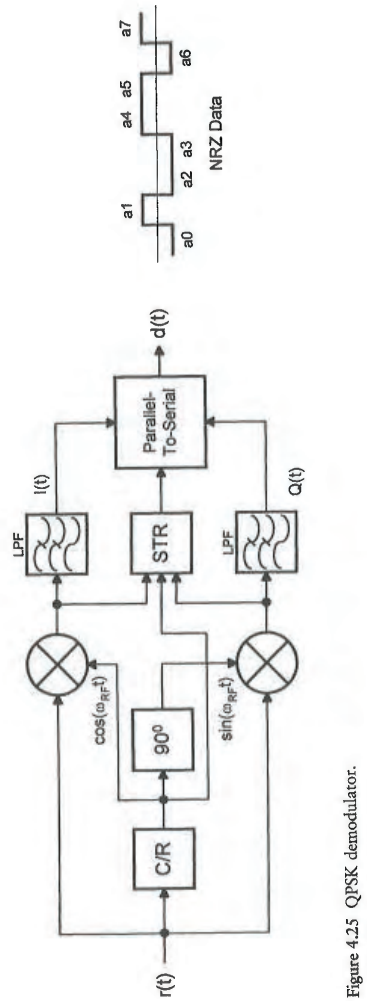


Figure 4.25 QPSK demodulator.

As in the case of BPSK, we can think of the recovered QPSK signal as a vector to which gets added a noise vector. This is shown in Figure 4.26. It is easy to see that any added noise need only be strong enough to knock the resultant vector into the next quadrant for a symbol error to result. In fact, the noise required to maintain a given symbol error rate for QPSK must be 3-dB lower than that required to maintain the same P_E in BPSK. However, when we demodulate the QPSK signal, we break it into two BPSK bit streams by multiplying by cosine and sine waveforms, as shown in the figure. This is further illustrated in Figure 4.27, where we have assumed, for now, that the carrier has no extra phase noise imparted onto it from the noisy received signal. Looking at the recovered signal in this form leads us to conclude that the bit error rate of QPSK is the same as for BPSK. This is in fact the case.

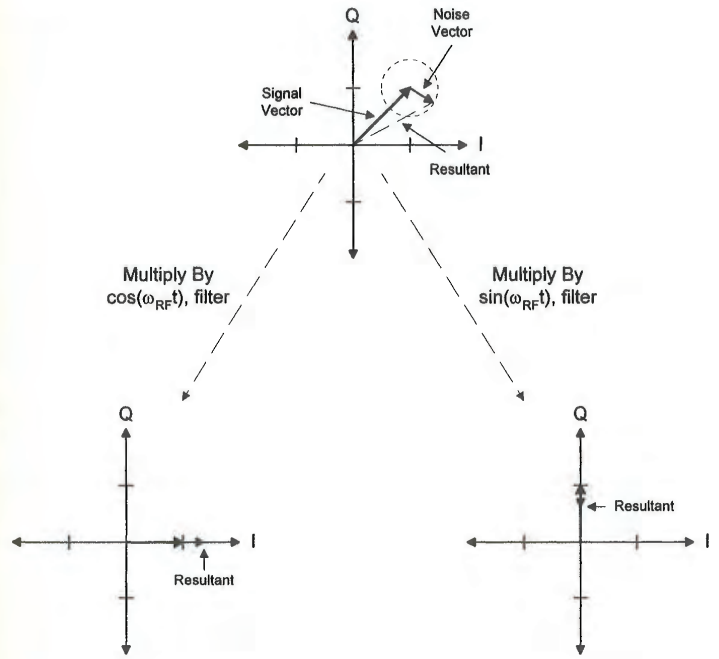


Figure 4.26 Resultant QPSK signal vector with noise (broken into I and Q components).

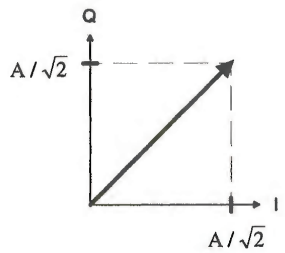


Figure 4.27 QPSK signal vector in the *I/Q* phase plane.

Recall, from our earlier discussions that we can write E_b/N_0 as

$$\frac{E_b}{N_0} = \frac{S}{N} \cdot \frac{B_n}{f_b} \tag{4.40}$$

where E_b is the energy per bit, N_0 is the noise power spectral density, B_n is the effective noise bandwidth of the receiver and f_b is the bit rate. In QPSK, we have created two data streams, each at one-half the rate of the original bit stream and each having maximum amplitude of $\pm(\sqrt{S})/\sqrt{2}$ (recall that S is the power of the received signal, therefore it is the square of its amplitude). Plugging this information into (4.40) gives

$$\frac{E_b}{N_0} = \frac{(\sqrt{S}/\sqrt{2})^2}{N} \cdot \frac{B_n}{(f_b/2)} = \frac{(S/2)}{N} \cdot \frac{B_n}{(f_b/2)} = \frac{S}{N} \cdot \frac{B_n}{f_b} \tag{4.41}$$

Therefore, we see that the BER of QPSK is the same as that of BPSK. Figure 4.28 shows a plot of the probability of bit error for QPSK versus E_b/N_0, which is identical to BPSK.

As we discussed above, a noise spike could very easily cause the resultant vector to fall into the wrong quadrant. If we code the data properly, a symbol error will result in a single bit error and not two bit errors. Figure 4.29 illustrates this point. In this figure, the symbols on the QPSK phase plane are shown for different assignments of bits. Because the noise is Gaussian by nature, it is much more likely that a noise spike will cause the resultant vector to fall in an adjacent quadrant, rather than the opposite quadrant. If the symbols are coded as shown in Figure 4.29(a), then falling into an adjacent quadrant could result in two bit errors. However, if we *gray code* the data, as shown in Figure 4.29(b), then each symbol error into an adjacent quadrant causes only one bit error. A gray code is a mapping of

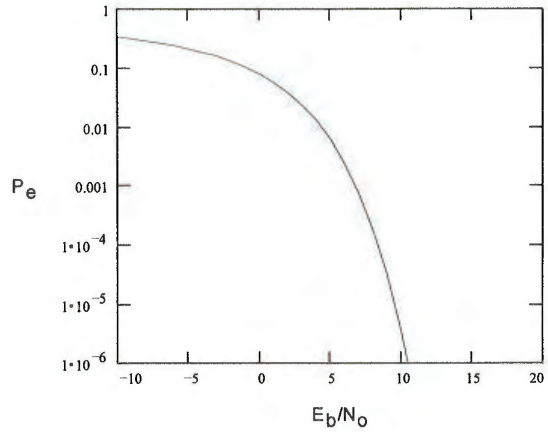


Figure 4.28 QPSK probability of bit error.

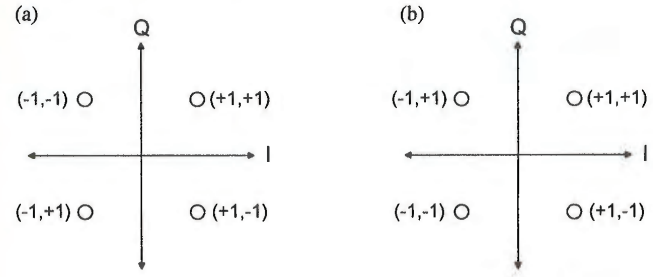


Figure 4.29 Assignment of symbols for QPSK: (a) nongray-coded QPSK, (b) gray-coded QPSK.

input bits to output bits, such that only one output bit at a time changes for sequential input bit changes.

Earlier, we saw that the power spectral density of QPSK was similar to that of BPSK, but only half as wide, due to the reduced symbol rate. In most wireless communication systems, such a spectrum would be too wide to meet FCC regulations and would also cause unwanted interference by spilling into adjacent channels. We could narrow the transmitted spectrum by placing a very narrow bandpass filter on

the output of the modulator. However, this is usually not practical, since the Q of the filter would have to be very high. This is due to the fact that the transmitted carrier frequency is usually much higher than the data rate. Such a high-Q filter would be difficult to build, expensive, and would cause significant distortion on the transmitted waveform due to rapid phase variations at the band edges. In addition, if we have many channels, we would also need to design such a filter so as to be tunable.

A better way to limit the transmitted output spectrum is by filtering the baseband I and Q modulation signals before applying these to the double-balanced modulators, as shown in Figure 4.30. Figure 4.31 shows the resulting transmitted spectrum of filtered and unfiltered QPSK. As the figure shows, the action of the filter greatly reduces the bandwidth of the transmitted signal. Figure 4.32 shows the corresponding time domain signal for a baseband filtered QPSK signal. Notice that every time a 180-deg phase change is encountered, the amplitude dips to zero. However, 90-deg phase changes result in much lower amplitude variations.

In most wireless communication systems, the output of the modulator is too small to be directly applied to the antenna. Therefore, the modulator output signal is sent to a transmit power amplifier for further amplification. In order to get the highest efficiency possible from the transmit power amplifier, it is common to run the circuit in nonlinear, Class C operation. In Class C operation, the amplifier consumes little power when no signal is applied to it. When a signal is applied to the amplifier, the amplifier is driven into limiting (sometimes called compression or saturation). Most of the power applied to a hardlimited amplifier is directly transferred to the load, resulting in maximum efficiency.

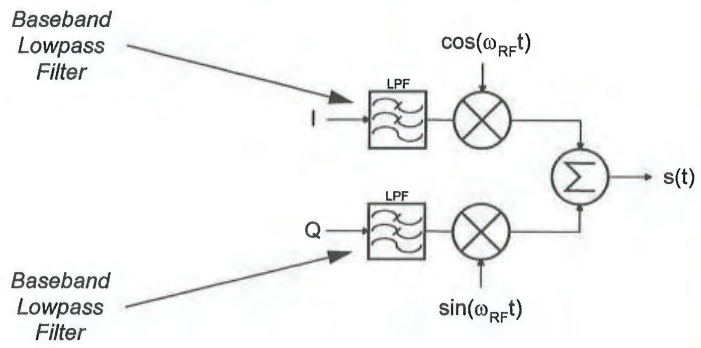


Figure 4.30 Using premodulation lowpass filters to limit QPSK transmitted bandwidth.

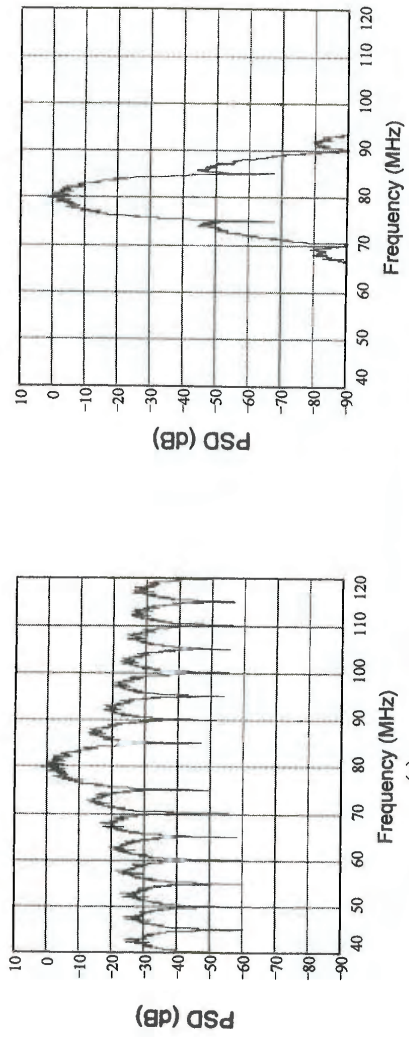


Figure 4.31 QPSK transmit power spectral density: (a) unfiltered (10-Mbps data rate, 80-MHz carrier), (b) baseband filtered (10-Mbps data rate, 80-MHz carrier, 2.5-MHz Gaussian lowpass premodulation filter).

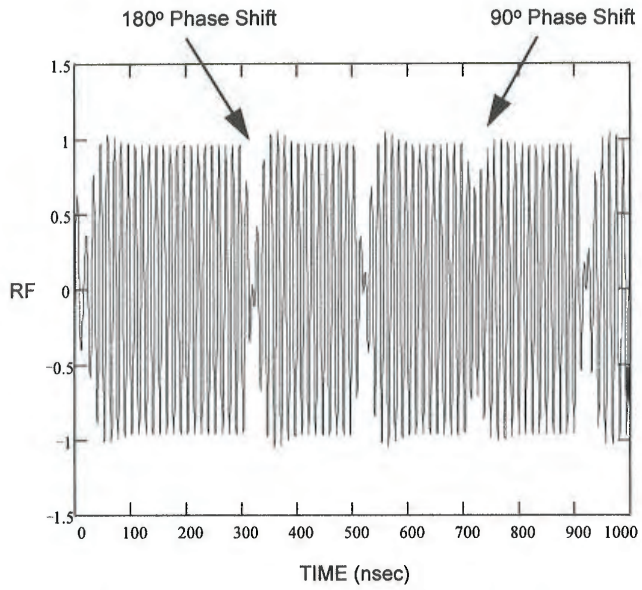


Figure 4.32 Time domain waveform of baseband-filtered QPSK transmit signal.

When the filtered QPSK signal is passed through a limited amplifier, much of the amplitude variations in the time domain envelope are reduced. However, it is these very amplitude variations that resulted in the narrow transmitted spectrum. Figure 4.33 shows the spectrum of a QPSK signal that has been first filtered and then limited. As the figure shows, the limiting action of the amplifier causes the spectrum to regrow. This is called *spectral regrowth*. In general, modulation schemes that have nonconstant envelopes suffer from spectral regrowth when passed through a hardlimited amplifier (i.e., an amplifier in saturation or compression). For this reason, these types of modulations generally require linear amplifiers in the transmit path.

4.2.3 Offset or Staggered Phase-Shift Keying (OQPSK)

As we have just seen, large amplitude variations in the QPSK waveform result whenever the phase is changed by 180 deg. We might expect that if we could make

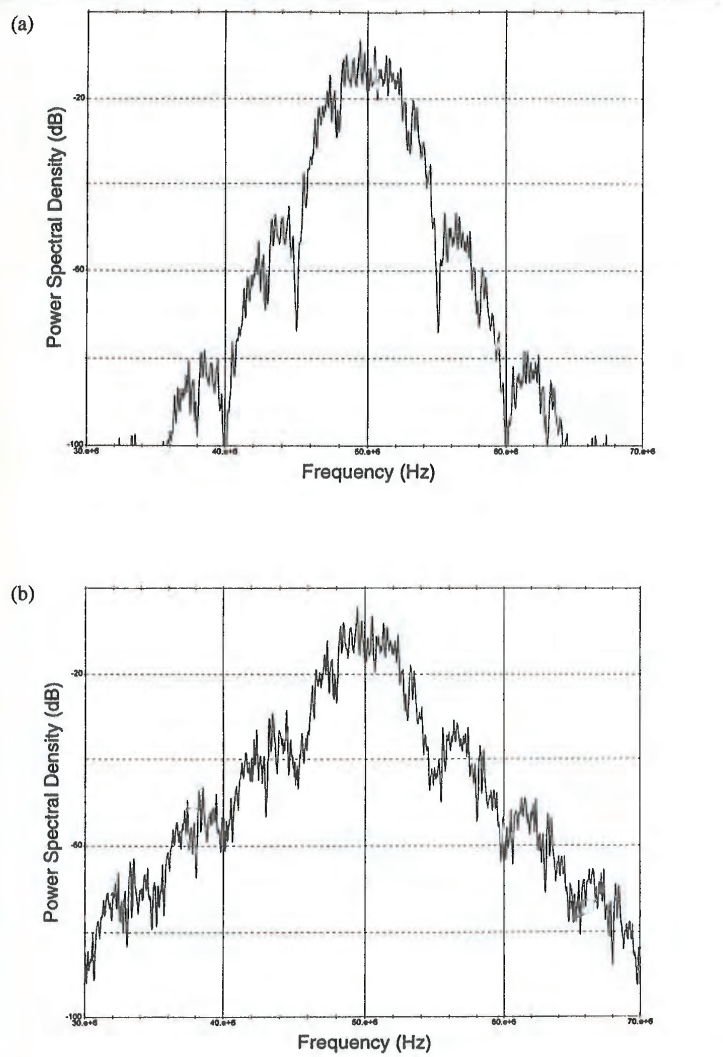


Figure 4.33 QPSK transmit output spectrum with baseband filtering: (a) without limiting, (b) with limiting (amplifier 10 dB into compression).

the phase transitions more gradual, then the QPSK waveform might be more tolerant to limiting. In offset, or staggered, QPSK (OQPSK), we do just that. OQPSK is identical to QPSK, except that the Q data stream is delayed by a single bit period (or 1/2 of a symbol period). This is easily done by modifying the data demultiplexer, as shown in Figure 4.34. As the figure shows, we simply remove the extra flip-flop from the I data stream. This has the same effect as delaying the Q data stream. Figure 4.35 shows the phase domain, I and Q data streams and corresponding phase angles for QPSK and OQPSK. The same two data streams are used in both cases, but notice how the action of the one bit period delay in QPSK causes no 180-deg phase transitions. OQPSK is sometimes used in satellite communications, but the real advantage of OQPSK is its use as a building block for other modulations, such as MSK (discussed in Section 4.3.1).

Figure 4.36 shows the time domain response of OQPSK. All of the phase transitions in OQPSK are 90 deg. Notice how the Q data stream is delayed by a bit period from its position in the example of QPSK.

Figure 4.37 shows a block diagram of a typical OQPSK modulator. The OQPSK modulator is identical to that of QPSK with the exception that an extra bit period delay is inserted into the Q data stream. In this figure, the demultiplexer is assumed to be identical to the one used in QPSK. Not surprisingly, the OQPSK demodulator is very similar to the QPSK demodulator. Figure 4.38 shows a block diagram of the demodulator, where an extra bit of delay is inserted in the I data stream to realign the data.

Due to the similarities with QPSK, both the power spectral density and the P_E of OQPSK are the same as for QPSK. However, when filtered and limited, OQPSK shows much smaller spectral regrowth than QPSK. Figure 4.39 shows a plot of a time domain response of a baseband-filtered OQPSK signal. As the plot shows, the amplitude fluctuations are greatly reduced. When this signal is passed through a

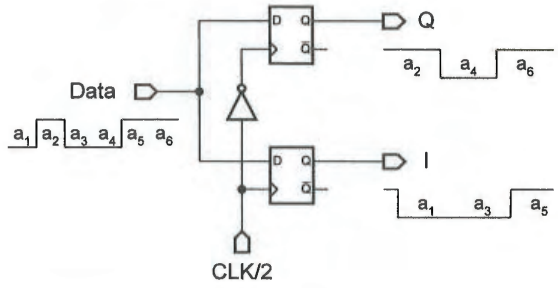


Figure 4.34 OQPSK data demultiplexer (serial-to-parallel).

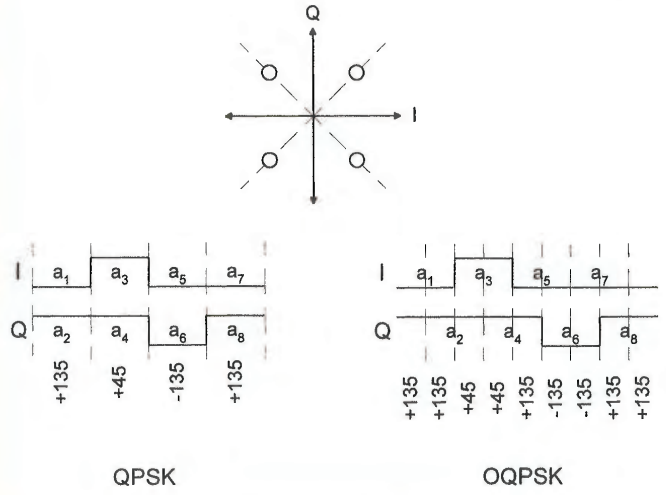


Figure 4.35 QPSK and OQPSK I/Q data waveforms with their corresponding phase angles.

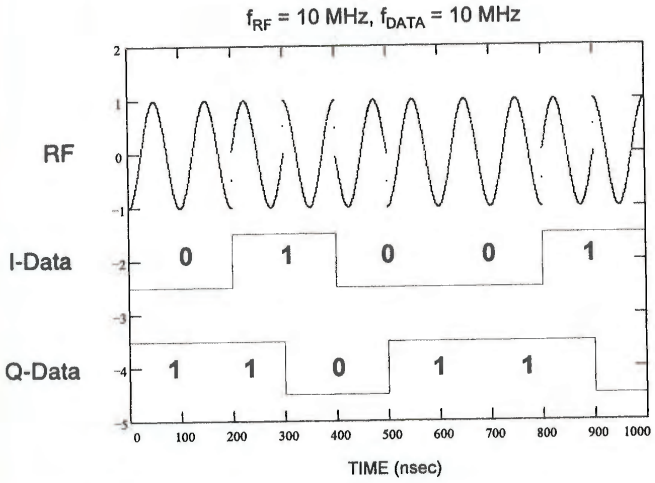


Figure 4.36 OQPSK time domain waveforms.

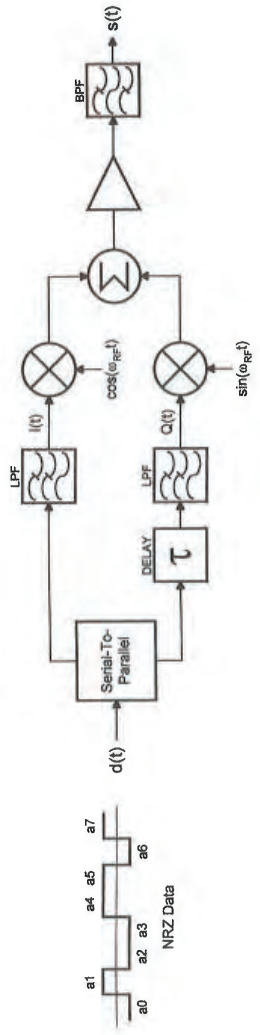


Figure 4.37 OQPSK modulator.

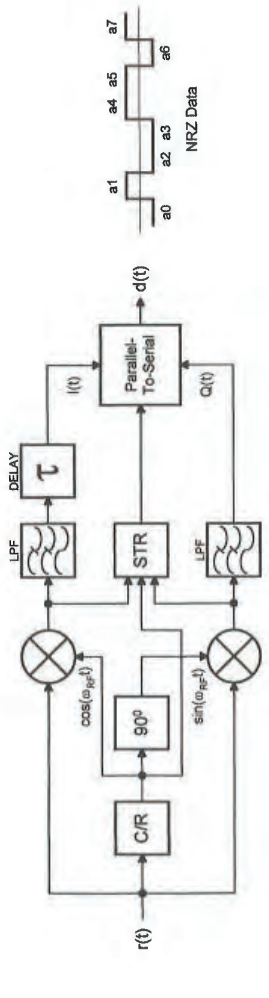


Figure 4.38 OQPSK demodulator.

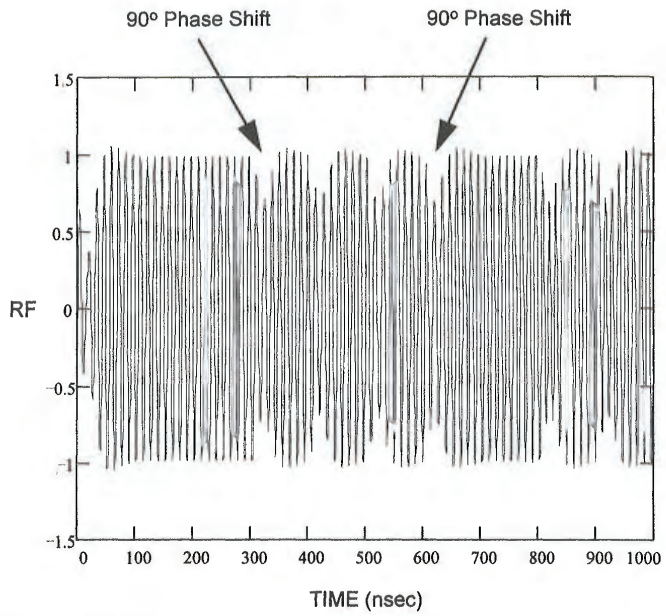


Figure 4.39 Time domain waveform of baseband-filtered OQPSK transmit signal.

limiting amplifier, its spectral regrowth is much lower than that of QPSK. Figure 4.40 shows a plot of this.

4.2.4 8-PSK and 16-PSK

The concept of using symbols can be extended further. By grouping three bits at a time into one symbol, we can transmit one of eight different phase angles, in reference to the unmodulated carrier. This type of phase modulation is known as 8-PSK. The phase domain of 8-PSK is shown in Figure 4.41. As the figure shows, each symbol lies on the unit circle in eight equally spaced phase angles, 45-deg apart. In order to maintain constant amplitude, the allowable values of I and Q are

$$I, Q = 0, \pm 1, \pm\frac{\sqrt{2}}{2} \tag{4.42}$$

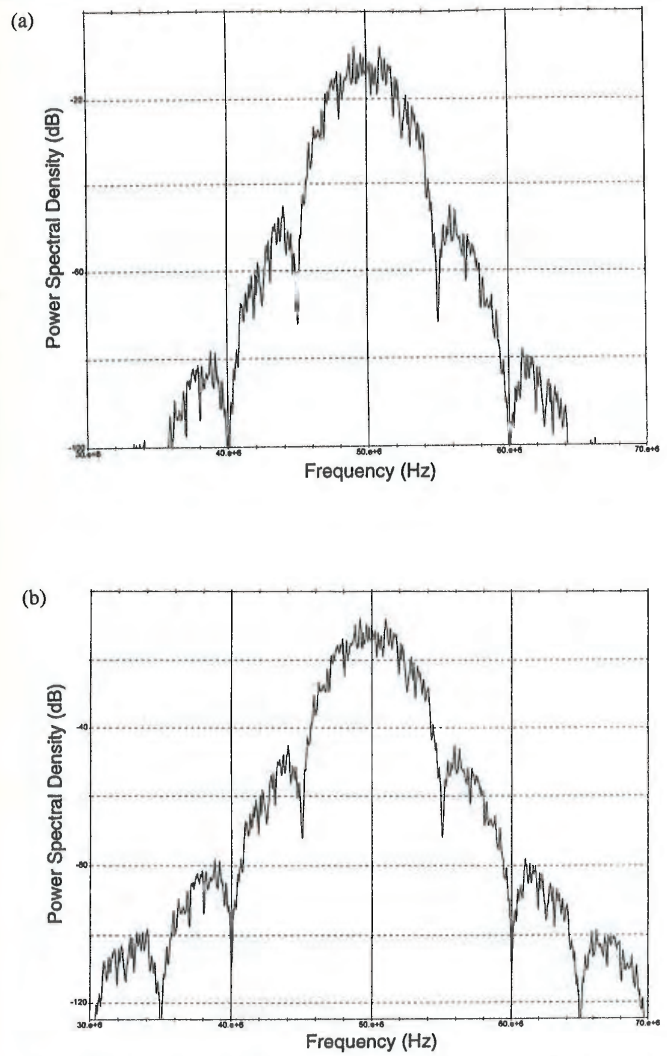


Figure 4.40 OQPSK transmit output spectrum with baseband filtering: (a) without limiting, (b) with limiting (amplifier 10 dB into compression).

Eight-PSK is sometimes used with special data encoding to form a modulation known as $\pi/4$-DQPSK (discussed in Section 4.4.2). 8-PSK is also used in point-to-point microwave communications.

Figure 4.42 shows the time domain response for an 8-PSK signal. The three upper traces correspond to the three demultiplexed bit streams taken from the original data. The modulated signal can vary in phase by any increment of 45 deg. Figure 4.43 shows a block diagram of an 8-PSK modulator. It is very similar to the QPSK modulator; however, the demultiplexer contains a data encoder that not only demultiplexes the original data in to groups of 3 bits, but also provides a set of outputs to drive the digital-to-analog-converters (DACs) with the allowable modulation levels given in (4.42).

Figure 4.44 shows a block diagram for an 8-PSK demodulator. The demodulator is basically the reverse of the 8-PSK modulator. Sine and cosine mixers are used to recover the I and Q signals corresponding to each phase angle. An analog-to-digital converter (ADC) is used to quantize the I and Q signals. The output of the ADC drives a decoder circuit that recovers the groups of three bits and reinterleaves them back into the original bit stream.

Once we exceed four phase angles, the bit error performance begins to deteriorate. Figure 4.45 shows the probability of symbol error versus E_b/N_0 for 2-PSK (i.e., BPSK), 4-PSK (i.e., QPSK), 8-PSK, and so on. Note that the figure shows symbol error probability, not bit error probability. Therefore, the P_E for BPSK is better than that of QPSK, given the same E_b/N_0. As the figure shows, more and more SNR is required to maintain the same P_E as the number of symbols increases.

Figure 4.46 shows a plot of the power spectral density for unfiltered 8-PSK. As in the case of both BPSK and QPSK, the sharp phase transitions result in the same classic sinx/x shape—with sharp nulls, large sidelobes and slow roll-off. However, the spectrum is only one-third as wide as that of BPSK, since groups of three bits are combined into one symbol, effectively lowering the symbol rate by 3.

Taking the concept one step further still, we take groups of four bits represented by a symbol. This is called 16-PSK and is shown in Figure 4.47. Each symbol is separated from the others by multiples of 22.5 deg. 16-PSK is used in some wireless local area network (WLAN) applications.

Eight-PSK and 16-PSK are generally considered as spectrally efficient modulations (having a spectral efficiency greater than 2 bits/sec/Hz). Quadrature amplitude modulation (QAM) falls into this class as well, though we do not cover it here. In QAM, we vary both the *phase* and *amplitude*. QAM modulations typically combine groups of 4 to 10 bits into individual symbols. QAM is considered one of the most spectrally efficient modulations in use today [5].

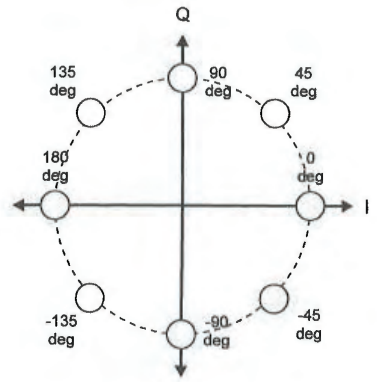


Figure 4.41 Phase domain of carrier for 8-PSK.

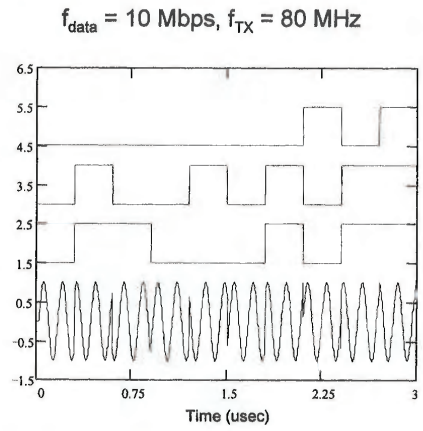


Figure 4.42 8-PSK time domain waveforms.

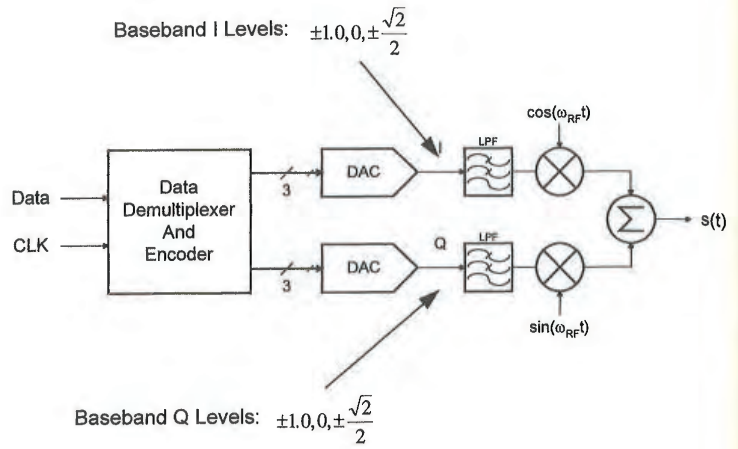


Figure 4.43 8-PSK modulator.

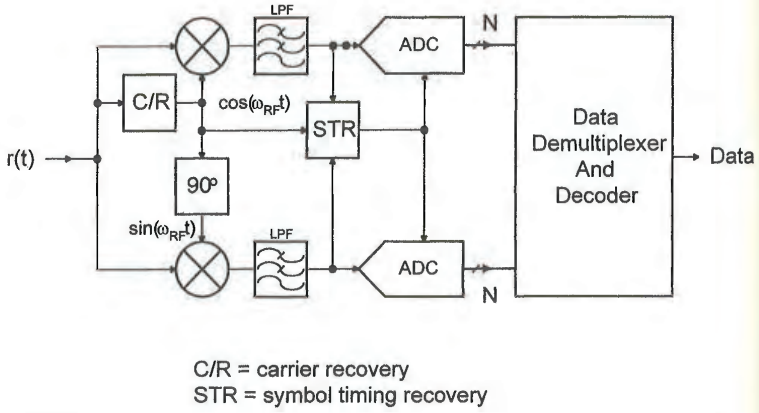


Figure 4.44 8-PSK demodulator.

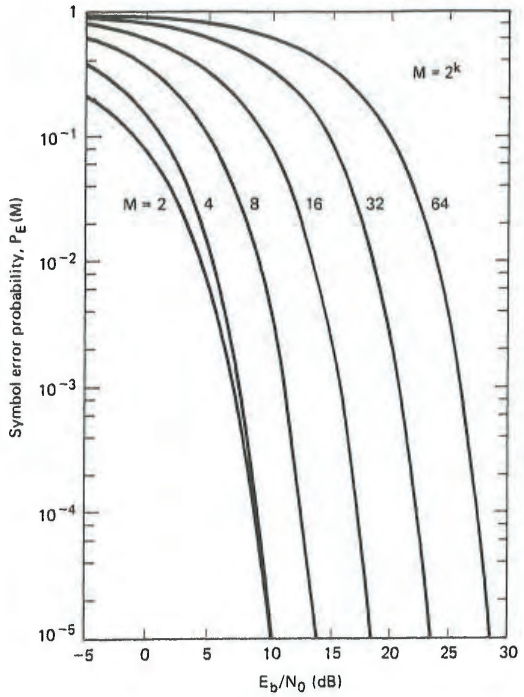


Figure 4.45 Probability of bit error for MPSK.

4.2.5 Nyquist Filtering as Applied to PSK Modulations

Earlier, we used (4.40), repeated here, to express the relation between E_b/N_0 and S/N

$$\frac{E_b}{N_0} = \frac{S}{N} \cdot \frac{B_n}{f_b} \tag{4.43}$$

where E_b is the energy per bit, N_0 is the noise power spectral density, S is the signal power, N is the noise power, B_n is the effective noise bandwidth of the receiver, and f_b is the bit rate.

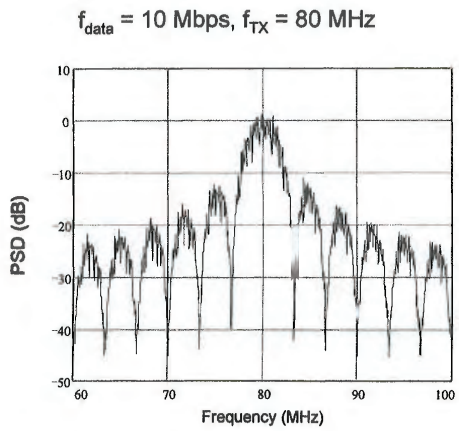


Figure 4.46 8-PSK transmit power spectral density (no baseband filtering).

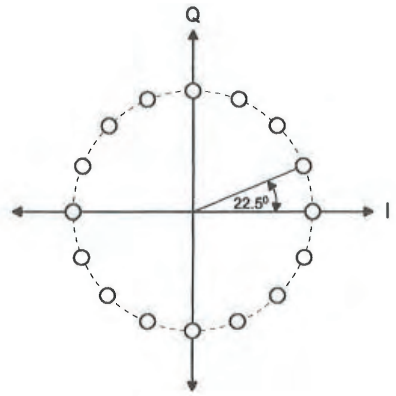


Figure 4.47 Phase domain of carrier for 16-PSK.

When the effective noise bandwidth is equal to the bit rate, E_b/N_0 equals S/N. However, we saw earlier that a modulation such as unfiltered BPSK occupies a much wider bandwidth than that equal to the bit rate. This means, in practice, the bandwidth of the receiver needs to be wider than the bit rate in order to avoid excessive ISI caused by passing a spectrally wide signal through too narrow a filter. If B_n gets larger, then (4.43) says that E_b/N_0 must increase in order to maintain the same P_E. This is because the wider bandwidth will allow more noise into the receiver, thus lowering the effective S/N (or SNR). Therefore, the narrower we can make the transmitted spectrum, the more sensitive we can make the receiver. However, too narrow a filter can cause excessive ISI, as we will see in the case of GMSK (Section 4.4.1).

A special class of filters, called Nyquist filters, solves the bandwidth problem without introducing excessive ISI into the waveform. Consider the frequency response of an ideal "brick-wall" lowpass filter shown in Figure 4.48(a). As the figure shows, the filter has unity gain within the passband and zero gain outside the passband. The impulse response can be found by simply taking the inverse Fourier transform of the filter's frequency response

$$h(t) = \int_{-\infty}^{\infty} H(f) \cdot e^{+j\omega t} df \tag{4.44}$$

$$h(t) = \int_{-f_b/2}^{f_b/2} 1 \cdot e^{+j\omega t} df = f_b \cdot \frac{\sin(\pi f_b t)}{(\pi f_b t)} \tag{4.45}$$

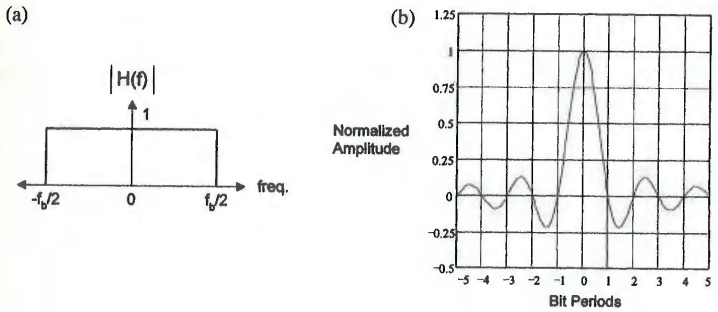


Figure 4.48 Frequency response and impulse response of an ideal brick-wall lowpass filter: (a) frequency response, (b) impulse response.

We see that the impulse response of the filter, shown in Figure 4.48(b), has the familiar $\sin x/x$ shape. A point of considerable interest is the fact the impulse response is equal to zero at the center of all bit periods except for $t = 0$. This can be easily seen by plugging $t = mT_b$ into the equation

$$h(mT_b) = f_b \cdot \frac{\sin(\pi f_b mT_b)}{(\pi f_b mT_b)} \tag{4.46}$$

Recall that $T_b = 1/f_b$. This gives

$$h(mT_b) = f_b \cdot \frac{\sin(m\pi)}{(m\pi)} \tag{4.47}$$

Thus,

$$h(0 \cdot T_b) = f_b \tag{4.48}$$

$$h(mT_b) = 0; \text{ for } m \neq 0 \tag{4.49}$$

Nyquist [10] proposed that the shape of each data pulse be such that the signal be bandlimited and yet there be no ISI due to preceding or following pulses at each sampling (or decision making) instant. He proposed three different methods for doing this. Filters, which produce pulse shapes with no ISI at each sampling instant, are referred to as Nyquist I type filters. Nyquist II filters, which will not be discussed here, employ filters whose pulse shape is such that pulses can be sent at a rate of $2f_b$ within a bandwidth of f_b and still be unambiguously interpreted. Nyquist III filters, also not discussed here, have pulse shapes such that the total area in all pulses not at $t = 0$ are zero.

Passing the sampled-data signal (actually ideal impulses) through an ideal brick-wall lowpass creates a pulse shape that meets Nyquist's first criterion. These types of filters are often called Nyquist filters.

Unfortunately, we can never actually realize the perfect brick-wall lowpass response. First of all, the impulse response for the brick-wall filter is noncausal. This is due to the fact that the impulse response is infinite in both directions in time. Secondly, the $\sin x/x$ shape tends to fall off rather slowly in time. We can, however, compromise slightly on the frequency response by allowing some of the energy in the filter to spill past the band edges of the filter.

Figure 4.49 shows one possible way in which this can be done. In this approach, we allow the edge of the band to taper off to zero following a cosine function. The filter response is equal to 1/2 at $\pm f_b/2$, but does not equal to zero until the frequency is equal to $\pm(1 + r) f_b/2$, where r is a constant. The quantity r is called the *roll-off factor*, and can range between 0 through 1. In addition, we allow the filter to begin

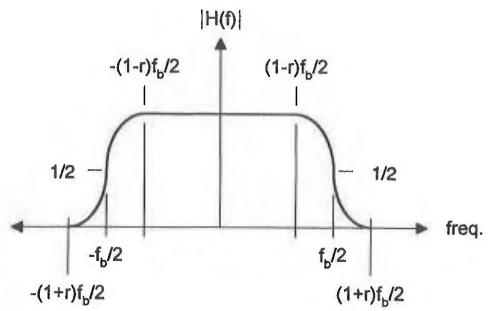


Figure 4.49 Frequency response for realizable raised cosine (Nyquist) lowpass filter.

to roll off starting at $\pm(1 - r) f_b/2$. Within the band $-(1 - r) f_b/2 < f < +(1 - r) f_b/2$, the filter has unity gain. By shaping the spectrum in this way, we can maintain the filter's characteristic of no ISI at the sampling instants of other bit periods, yet roll off the impulse response much more quickly in time. A filter with such a response is called a raised cosine filter. Notice that when $r = 0$, the filter response is equal to that of an ideal brick-wall lowpass filter. When $r = 1$, the frequency response begins to roll off at $f = 0$, with no flat-topped unity gain region.

To find the impulse response of the raised cosine filter, we take the inverse Fourier transform of the frequency response. The frequency response of the filter can be divided into five different regions

$$H(f) = H_1(f) + H_2(f) + H_3(f) + H_4(f) + H_5(f) \tag{4.50}$$

Regions $H_1(f)$ and $H_5(f)$ lie outside the filter band and are equal to zero. Regions $H_2(f)$ and $H_4(f)$ are on the cosine-shaped portion of the band, while region $H_3(f)$ has a frequency response equal to unity. The frequency response of the filter can be written as

$$|H(f)| = 1; \text{ for } 0 \leq |f| < \frac{f_b}{2}(1 - r) \tag{4.51}$$

$$|H(f)| = \cos^2\left\{\frac{1}{4f_b r} \cdot [2\pi f - \pi(1 - r) f_b\right\}; \text{ for } -\frac{f_b}{2}(1 - r) \leq |f| < +\frac{f_b}{2}(1 - r) \tag{4.52}$$

$$|H(f)| = 0; \text{ for } |f| < \frac{f_b}{2}(1 + r) \tag{4.53}$$

Letting $r = 0.5$, we can find the impulse response of the filter by taking the inverse Fourier transform. The result is

$$h(t) = \int_{-\infty}^{\infty} H(f) \cdot e^{+j\omega t} df = f_b \cdot \frac{\cos(\pi f_b t)}{1 - 4f_b^2 t^2} \cdot \frac{\sin(\pi f_b t)}{(\pi f_b t)} \tag{4.54}$$

Figure 4.50 shows a plot of raised cosine filter impulse responses for $r = 0$, 0.5, 1.0 over ±3 bit periods. Notice that $r = 0$ is for the ideal brick-wall filter response, $r = 1.0$ is for the response where the bandwidth begins to taper off immediately, and $r = 0.5$ is in between these two extremes. By looking at the plot, we see that $r = 0.5$ results in a pulse shape that is almost zero by time it gets out to ±3 bit periods.

Figure 4.51 shows a block diagram of an FIR filter, which can be used to construct a raised cosine filter. The input to the filter is really an analog variable. Therefore each delay is a pure analog delay. Such an analog filter is often fabricated

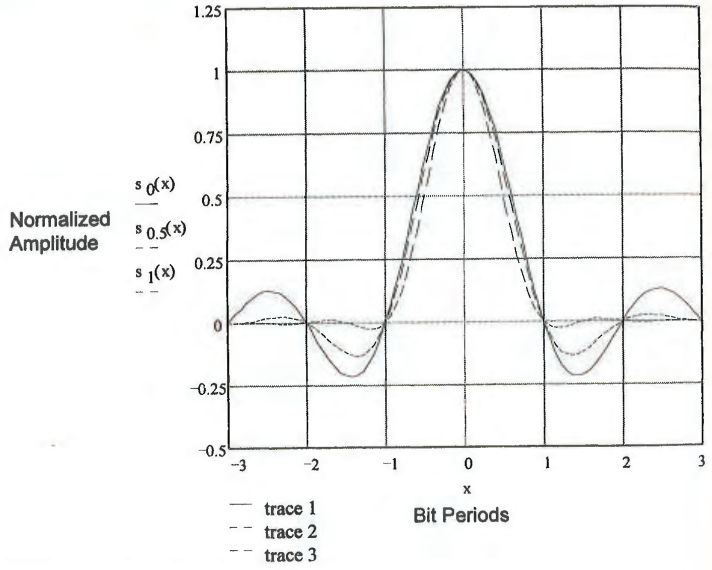


Figure 4.50 Raised cosine impulse response for different values of roll-off factor (r).

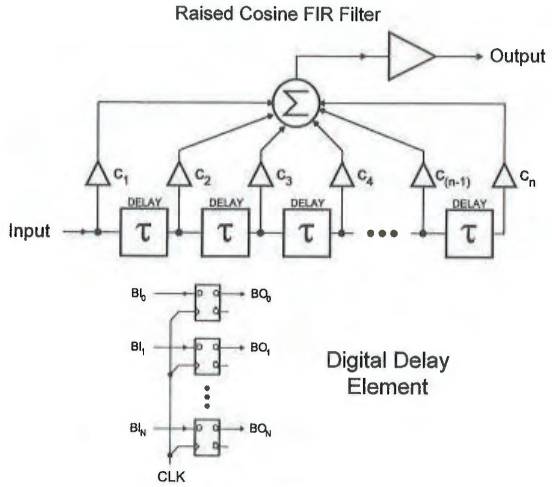


Figure 4.51 FIR filter realization of raised cosine filter.

in a surface acoustic wave (SAW) filter. The adders and tap-weight multipliers are implemented by the spacing and configuration of the metallization on the surface of the SAW substrate.

In many cases, the raised cosine filter is realized in digital form. The input to the filter is first digitized by an ADC and converted into a parallel combination of bits. Each adder and multiplier is implemented as a digital circuit. Each delay consists of a collection of parallel registers, as shown in the inset. The digital output of the filter is then converted back to analog form by a DAC, whose output is then filtered by an analog filter.

When the filter is implemented in digital form, a general purpose DSP processor is often used. The DSP processor performs the filter function by recursively using a single multiplier and accumulator to perform the tap-weight multiplication and summation. This approach performs the same function shown in Figure 4.51, but greatly reduces the amount of digital circuitry required.

Figure 4.52 shows an "eye" diagram of a raised cosine filter with $r = 0.5$. Notice that there is no ISI at the sampling instants. The overall pulse shape is sinusoidal in appearance, suggesting a narrow power spectral density. This is in fact the case. Figure 4.53 shows the power spectral densities for an unfiltered BPSK signal and a BPSK signal whose modulation signal has been passed through an

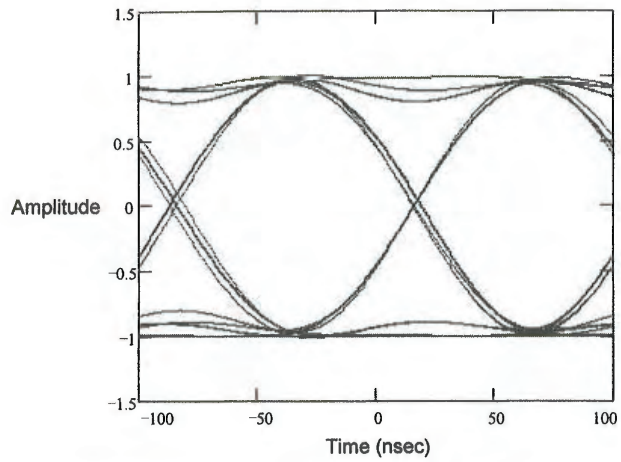


Figure 4.52 Eye diagram of output of raised cosine filter (no x/sinx equalization).

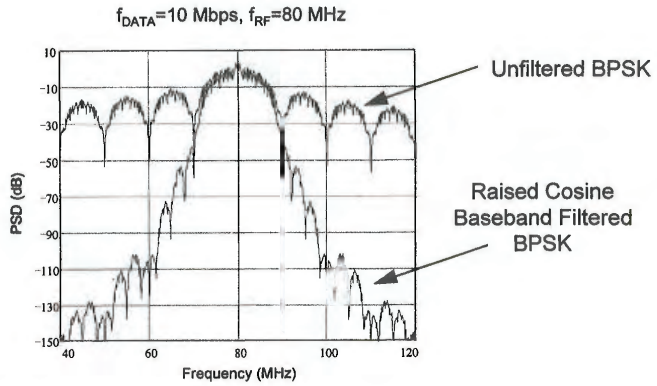


Figure 4.53 Transmit output spectrum of BPSK with and without raised cosine baseband filtering.

$r = 0.5$ raised cosine filter. As the figure shows, the spectrum of the filtered signal falls off dramatically.

The impulse response of a filter is the filter output when the input is an ideal impulse function (i.e., delta function). However, the output of a DAC driven by a digital signal is not a train of impulses, but a flat-topped succession of pulses that form a stairstep-looking response. This type of signal is often called a "boxcar" response due to its resemblance to a series of boxcars passing by on a freight train. Figure 4.54 shows a plot of the output of an ideal raised cosine filter and the output of an actual raised cosine filter. The boxcar output of the DAC causes the signal to prematurely roll off in frequency, instead of maintaining its flat-frequency response.

We can compensate the roll-off of the DAC within the raised cosine filter itself. This can be done by visualizing the filter output as a stream of impulses convolved with an ideal rectangular pulse. The result is the desired flat-topped boxcar output. This is shown in Figure 4.55. Recall that convolution in the time domain results in multiplication in the frequency domain. This implies

$$s_{Box}(t) = p(t) * s_{RC}(t) \tag{4.55}$$

where $s_{Box}(t)$ is the time domain boxcar output, $p(t)$ is the time domain pulse shape of the ideal rectangular pulse and $s_{RC}(t)$ is the sampled data impulse train of the raised cosine filter output. In the frequency domain, we can write

$$S_{Box}(f) = P(f) \cdot S_{RC}(f) \tag{4.56}$$

where

$$S_{RC}(f) = \frac{1}{T_{samp}} \cdot \sum_{n=-\infty}^{\infty} S\left(f - \frac{n}{T_{samp}}\right) \tag{4.57}$$

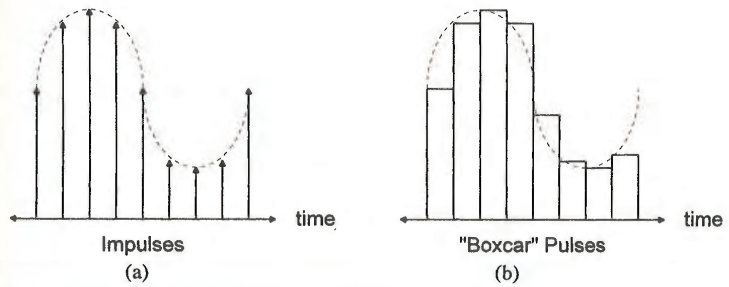


Figure 4.54 Sampled data output of raised cosine filter: (a) ideal filter output (unit impulses), (b) actual filter output (sample-and-hold, or boxcar).

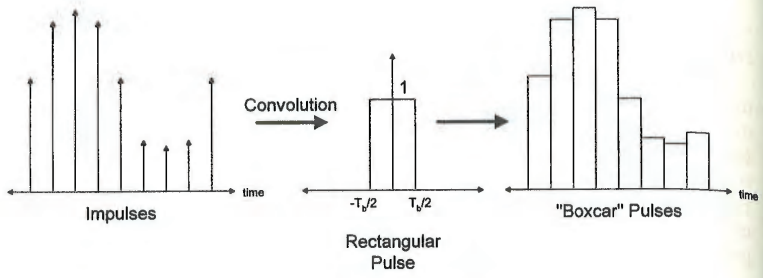


Figure 4.55 Modeling the boxcar output of the raised cosine filter.

Here, $S(f)$ is the spectrum of the raised cosine filter's impulse response and $S_{RC}(f)$ is the spectrum of the impulse train output of the raised cosine filter. The Fourier transform of the ideal rectangular pulse is given by

$$P(f) = T_{samp} \cdot \frac{\sin(\pi f T_{samp})}{(\pi f T_{samp})} \tag{4.58}$$

This results in the following boxcar spectrum

$$S_{Box}(f) = \left[T_{samp} \cdot \frac{\sin(\pi f T_{samp})}{(\pi f T_{samp})}\right] \cdot \frac{1}{T_{samp}} \sum_{n=-\infty}^{\infty} S_{RC}\left(f - \frac{n}{T_{samp}}\right) \tag{4.59}$$

$$S_{Box}(f) = \left[\frac{\sin(\pi f T_{samp})}{(\pi f T_{samp})}\right] \cdot \sum_{n=-\infty}^{\infty} S_{RC}\left(f - \frac{n}{T_{samp}}\right) \tag{4.60}$$

The above equations show that the flat-topped boxcar output of the DAC is equivalent to having the ideal raised cosine filter output multiplied by a sinx/x shaped roll-off. At the Nyquist frequency of the filter (i.e., one-half the sample rate), the sin$x$/x term causes significant roll-off. This can be calculated by letting $f = f_{samp}/2$

$$S_{Box}(f_{samp}/2) = \left[\frac{\sin\left(\pi \frac{f_{samp}}{2} T_{samp}\right)}{\left(\pi \frac{f_{samp}}{2} T_{samp}\right)}\right] \cdot S_{RC}(f_{samp}/2) = \frac{2}{\pi} \cdot S_{RC}(f_{samp}/2) \tag{4.61}$$

So, we see the boxcar effect reduces the output at the Nyquist frequency by about 3.9 dB.

This roll-off can be compensated for by "pre-emphasizing" the frequency response of the filter, before taking the inverse Fourier transform to find the impulse response. We do this by multiplying the frequency response by an $x/sinx$ term as follows:

$$|H(f)| = \frac{(\pi f T_b)}{\sin(\pi f T_b)}; \text{ for } 0 \le |f| < \frac{f_b}{2}(1 - r) \tag{4.62}$$

$$|H(f)| = \frac{(\pi f T_b)}{\sin(\pi f T_b)} \cos^2\left\{\frac{1}{4 f_b r} \cdot [2\pi f - \pi(1 - r) f_b]\right\}; \text{ for } -\frac{f_b}{2}(1 - r) \le |f| < +\frac{f_b}{2}(1 - r) \tag{4.63}$$

$$|H(f)| = 0; \text{ for } |f| < \frac{f_b}{2}(1 + r) \tag{4.64}$$

The desired impulse response is then computed from the above equations. Figure 4.56 shows a plot of an equalized eye diagram for a raised cosine filter. As the plot shows, the sections of the waveforms between the sampling instants tend to be much more pronounced.

4.3 FREQUENCY MODULATION

Up to now, we have discussed modulations where the phase of the carrier is abruptly changed to a new value in response to the input data signal. We have seen that this can result in a wide transmitted spectrum and can also result in spectral regrowth if the signal does not have a constant envelope. We have also seen that we can use multiple phase angles to represent groups of bits and thus reduce the effective symbol rate at the transmitter.

Instead of breaking the phase angle into small pieces, perhaps we could smoothly change the phase angle from one value to another. If we do this in such a way so as to remain on the unit circle, we will have a transmitted signal with constant amplitude. A constant amplitude signal would be impervious to the effects of filtering/hardlimiting and spectral regrowth.

Recall that frequency is the derivative of phase, or

$$\omega(t) = \frac{d\varphi(t)}{dt} \tag{4.65}$$

Smooth, continuous phase modulation corresponds to frequency modulation. In this section, we will show how frequency can be used as a modulation parameter.

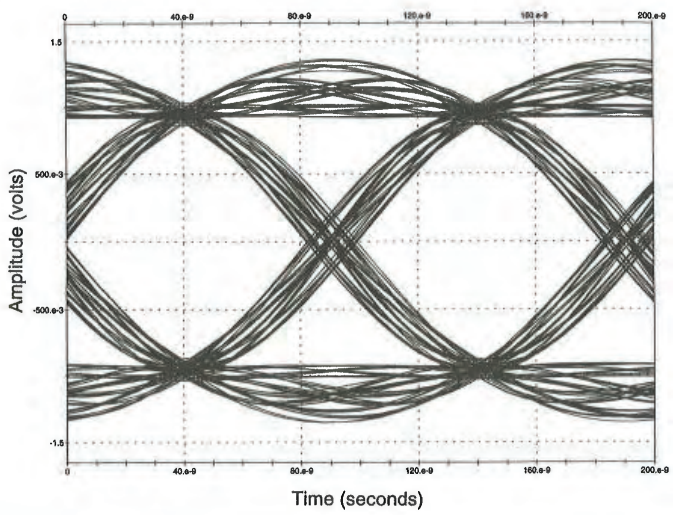


Figure 4.56 Raised cosine filter output with $x/\sin x$ equalization.

4.3.1 Minimum Shift Keying

Minimum shift keying (MSK) can be thought of as either a phase or a frequency modulation. Figure 4.57 shows the phase domain plot for MSK. In MSK, the carrier changes phase by ±90 deg over the course of a bit period. For example, suppose the carrier phase at the beginning of the data period is 0 deg. If the next data bit is a 1, then the phase smoothly advances by 90 deg by the end of the bit period. If the data bit remains a 1, then the phase advances another 90 deg to 180 deg. Instead, if the next bit were a 0, then the phase would smoothly retard by 90 deg to –90 deg. Again, if the data were to remain at 0, then phase would continue to retard itself by another 90 deg. Since the phase will continue to advance or retard itself over the course of each bit period, the derivative of the phase, or frequency, is varied. In this way, MSK can also be considered a frequency modulation. MSK is used in satellite communications, military tactical radio, and extremely low frequency (ELF) underwater communications.

Figure 4.58 shows how the phase of MSK can be plotted over time. This is known as a phase trellis. As the figure shows, the phase angle changes smoothly by 90 deg over each bit period. We can write the MSK signal as

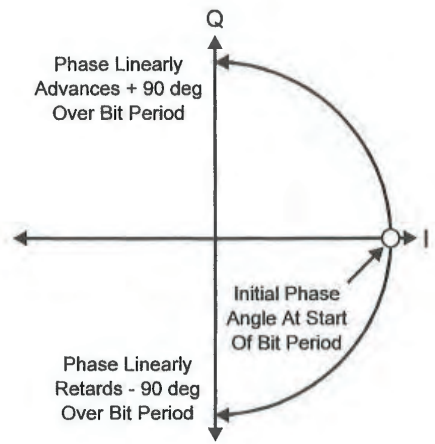


Figure 4.57 Phase domain of carrier for MSK.

$$s(t) = A \cdot \cos\left[\omega_{RF}t + \int_{-\infty}^{t} d(t) \cdot \frac{\pi}{2} dt\right] = A \cdot \cos[\omega_{RF}t + \frac{\pi}{2} \cdot (\pm 1) \cdot t + \varphi_0] \quad (4.66)$$

where $d(t)$ gives rise to the ±1 term in the above equation.

To find the effective frequency difference, assume that $d(t)$ changes from one bit period to another. In one case, the frequency will be higher, and when the input changes, the frequency will be lower. We can write this as

$$\Delta f^{+} = \frac{\Delta\omega}{2\pi} = \frac{\Delta\varphi(t)}{\Delta t} \cdot \frac{1}{2\pi} = \frac{\pi/2}{T_b} \cdot \frac{1}{2\pi} = \frac{f_b}{4} \quad (4.67)$$

$$\Delta f^{-} = \frac{\Delta\omega}{2\pi} = \frac{\Delta\varphi(t)}{\Delta t} \cdot \frac{1}{2\pi} = \frac{-\pi/2}{T_b} \cdot \frac{1}{2\pi} = -\frac{f_b}{4} \quad (4.68)$$

where $f_b = 1/T_b$ where f_b is the bit rate and T_b is the bit period. Since we are interested in the frequency difference, we subtract the two terms above to get

$$\Delta f^{+} - \Delta f^{-} = \left(\frac{f_b}{4}\right) - \left(-\frac{f_b}{4}\right) = \frac{f_b}{2} \quad (4.69)$$

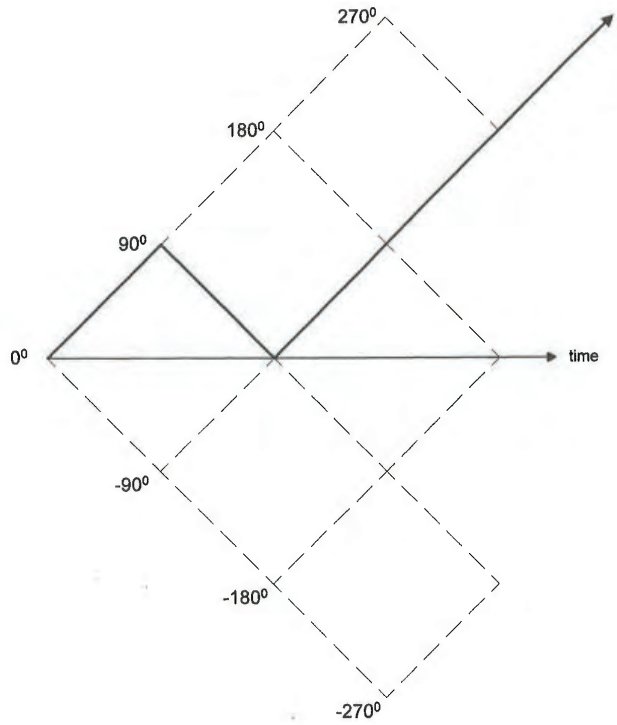


Figure 4.58 MSK phase trellis.

So, we see in MSK the frequency difference that corresponds to a 90-deg phase advance or lag over a bit period is equal to one-half the bit rate. In fact, the word minimum corresponds to the minimum frequency difference between the "mark" and "space" frequencies, which can be coherently demodulated.

We have stated earlier that the quadrature modulator is a sort of "universal" modulator. We might ask ourselves if there is a set of waveforms we can apply to the *I* and *Q* inputs that result in the desired constant envelope phase shift required for MSK. The answer is yes, and these waveforms are shown in Figure 4.59. As the figure shows, if the data stream does not change, then the *I* and *Q* modulation signals are simply two sinusoids 90 deg out of phase with one another. We can

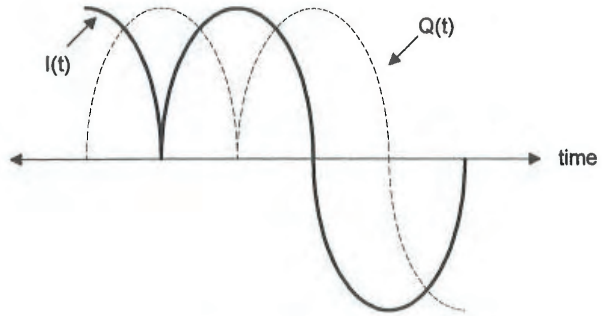


Figure 4.59 *I* and *Q* waveforms corresponding to MSK phase trellis shown in Figure 4.58.

think of this as a rectangular-to-polar conversion process. Simply look at any point on the unit circle and determine what values of *I* and *Q* are required for that phase angle. If the next bit is a 1, find a point on the unit circle that corresponds to 90-deg more of phase angle. This will dictate what the *I* and *Q* modulation signals should look like. As the data stream changes value, the *I* and *Q* signals will look like half-period sinusoids.

This is better illustrated in Figure 4.60. At the bottom of the plot, the *I* and *Q* values are shown as rectangular data pulses. After passing through special pulse-shaping filters, the *I* and *Q* signals become "half-sinusoidal" in nature. These are shown overlaid on the data streams. Notice that the corresponding output signal is indeed constant envelope, with its frequency changing from a slightly higher to a slightly lower value.

A block diagram of an MSK modulator is shown in Figure 4.61 [11]. Notice that the MSK modulator is identical to an OQPSK modulator, except for the special pulse-shaping filters. Figure 4.62 shows a detailed block diagram of each pulse-shaping filter, which is implemented as a counter-based finite impulse response (FIR) filter using read-only memories (ROMs). In this case, however, the filter has only one bit as an input, which greatly reduces complexity. The input to the filter simply reverses the polarity of the output signal at the end of each symbol period. Figure 4.63 shows a block diagram of the MSK demodulator. The MSK demodulator is identical to an OQPSK modulator. In general, the lowpass filters shown in the diagram are implemented as *matched filters*. Matched filters are special filters that are shaped to allow the signal itself to pass with little or no attenuation and reject all other waveforms (including noise).

The careful reader will notice that if a data stream of logic 1s or logic 0s are fed into the MSK modulator, the phase will not rotate about the unit circle. Instead,

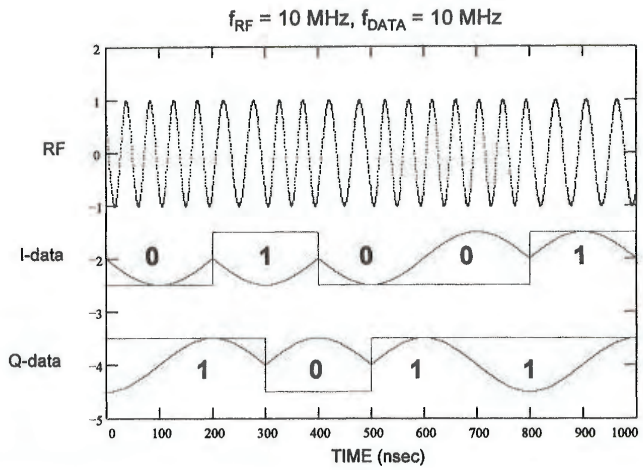


Figure 4.60 MSK time domain waveforms.

the phase bounces back and forth 90 deg. This really isn't a problem in most applications, since the incoming data stream is generally very close to a random bit stream anyway and so long as the demodulator "undoes" what the modulator does, the original bit stream is recovered. However, in applications that require true FM-like modulation, the I and Q bit streams can be differentially encoded.

As we have said earlier, MSK can be viewed as either frequency or phase modulation. This implies that we should be able to produce an MSK signal via standard FM circuitry. This is in fact the case, as Figure 4.64 shows. In the figure, the bit stream is fed directly into the modulation input of a VCO. We have already seen that the MSK signal has total frequency deviation equal to one-half the bit rate. Recall that the modulation index for an FM signal is given by

$$h \equiv \beta = \frac{\Delta f}{f_m} \tag{4.70}$$

where the variable, h, is more commonly used as the index of modulation for digital signals and β is the more familiar FM (analog) index of modulation. Δf is the total frequency difference and f_m is the modulation frequency. For MSK, we set the level of the modulation signal into the VCO to produce a modulation index, h, of 0.5.

Returning to Figure 4.64, the oscillator output is fed to a transmit power amplifier and filtered before being sent to the antenna. Because MSK can be viewed

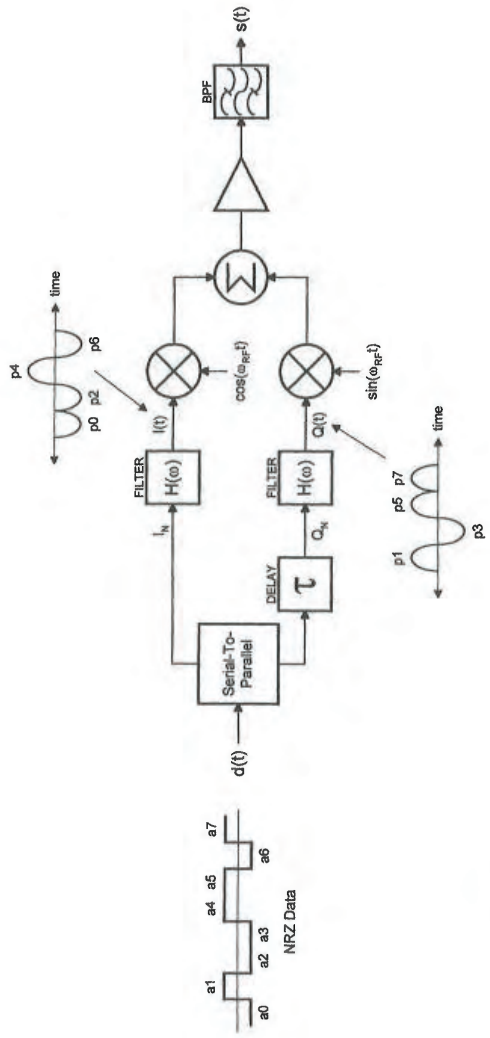


Figure 4.61 MSK modulator.

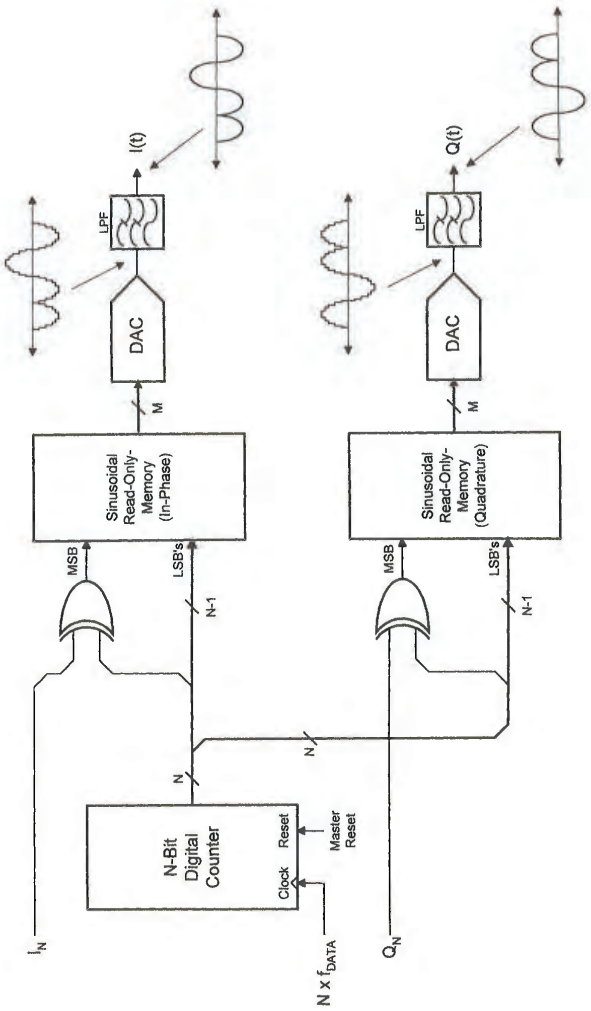


Figure 4.62 Implementation of MSK modulator using digital counters, ROMs, and D/A converters.

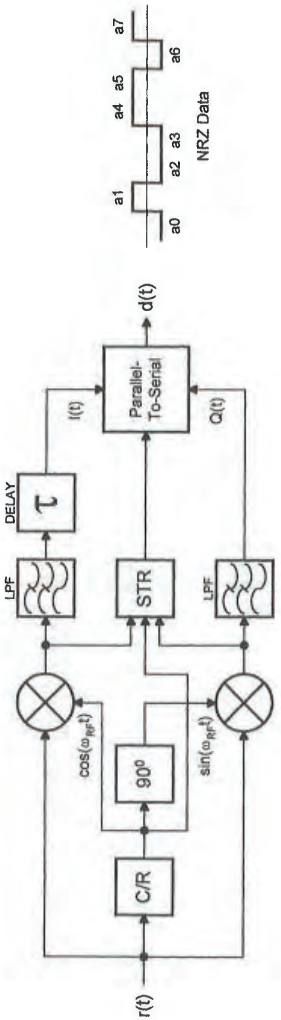


Figure 4.63 MSK demodulator.

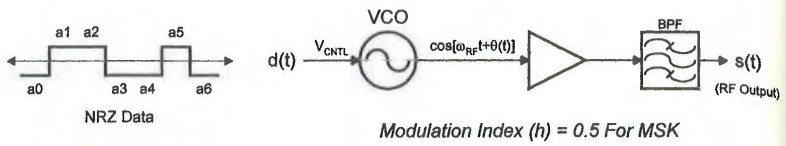


Figure 4.64 Using a voltage-controlled oscillator (VCO) as an MSK modulator.

as a type of frequency modulation, it is sometimes known as *fast frequency shift keying* (FFSK). Using the standard FM modulator is acceptable when MSK is noncoherently demodulated. However, it can be difficult to maintain the required index of modulation with an open-loop VCO approach due to variations of modulation sensitivity in the VCO. Treating MSK as a type of *I*/*Q* phase modulation ensures that the modulation index is fixed to 0.5, thus providing for fully coherent demodulation.

Figure 4.65 shows the time domain response of MSK. The modulated carrier is shown at the top of the plot. As stated earlier, the phase changes linearly 90 deg over the course of a bit period. This phase change is shown in the second trace.

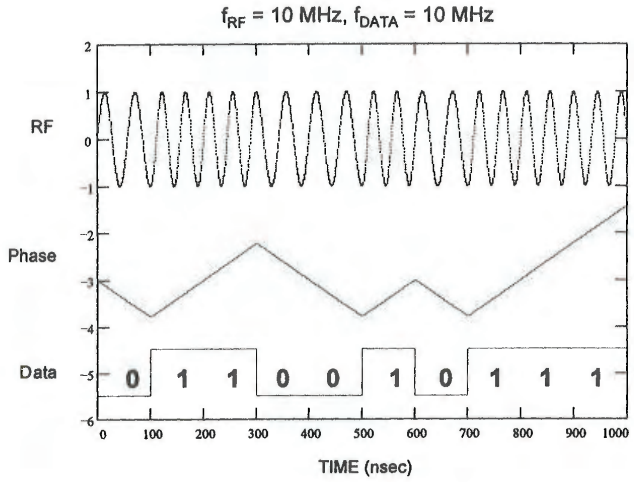


Figure 4.65 Time domain waveforms from VCO-based MSK modulator.

Notice that the relative slope is the same for a logic 0 as it is for a logic 1, only with a negative slope. The lower portion of the plot shows the input data stream.

A plot of the power spectral density of an MSK signal, compared to that of an unfiltered QPSK signal, is shown in Figure 4.66 for a carrier frequency of 80 MHz and a data rate of 10 Mbps. Notice how the MSK spectrum falls off much more rapidly than the QPSK spectrum. This is due to the fact that the MSK signal is continuous in-phase. Although the phase is continuous in MSK, the frequency (i.e., the slope of the phase) is not. It is this discontinuous frequency that results in a somewhat wider power spectral density with sharp nulls.

Figure 4.67 shows a plot of bit error probability versus E_b/N_0. Notice that the curve is the same as that of BPSK and QPSK. The reason for this is that MSK can be viewed as QPSK (actually OQPSK due to the one bit period offset) with special pulse-shaping filters, acting over two bit periods. The *I* and *Q* modulation signals are each half-sinusoidal pulses two bit periods long, which can assume either positive or negative values. This makes MSK another form of antipodal modulation. So long as we integrate over two bit periods in the demodulator, we can achieve the same P_E as for other antipodal modulations. In effect, we have knowledge of where the phase has been and where it is going if we look at the signal over two bit periods. This knowledge can be used to predict the value of the next signal, and hence lower the P_E of the signal.

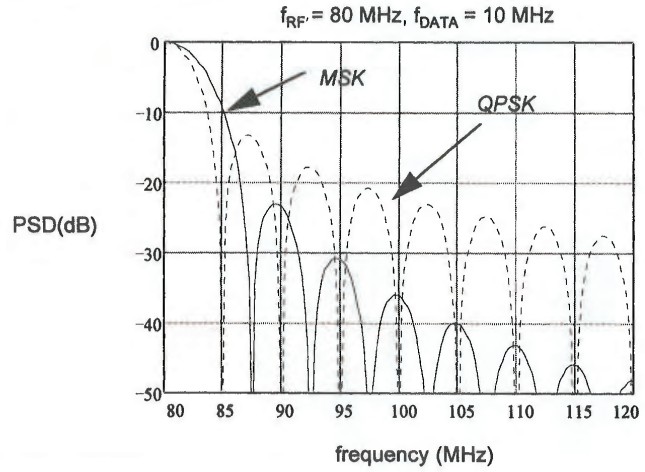


Figure 4.66 MSK versus QPSK power spectral density.

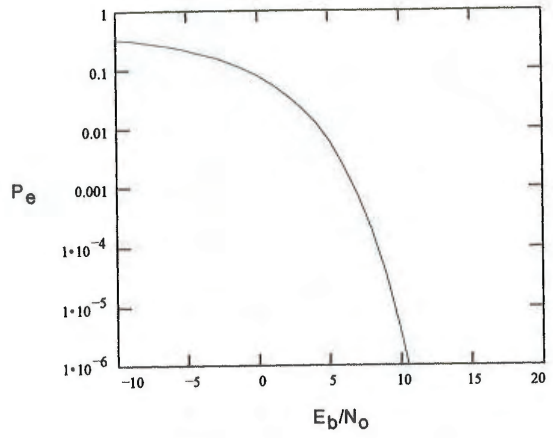


Figure 4.67 MSK probability of bit error.

Figure 4.68 shows a plot of an MSK signal in the time domain. As theory suggests, MSK is a true constant envelope modulation. No amplitude fluctuations can be seen in the waveform. Figure 4.69 shows a plot of the MSK signal before and after hardlimiting. As the figure shows, there is essentially no spectral regrowth. This feature of MSK makes it very attractive in applications where saturated power amplifiers are employed.

4.3.2 Frequency Shift Keying

MSK can be considered as either frequency or phase modulation. The only caveat is that we make the modulation index, h, equal to 0.5. Recall, that $h = 0.5$ is the minimum value of modulation index that we can use for coherent modulation. If we remove the restriction that $h = 0.5$, we can classify the type of modulation as a form of frequency shift keying (FSK). FSK has been used as a means of transmitting digital information for years. This is primarily due to the simplicity of the modulation and demodulation circuitry. Figure 4.70 shows block diagrams for a noncoherent FSK modulator and demodulator. As Figure 4.70(a) shows, FSK modulation is accomplished by applying the appropriate-level signal to the VCO's modulation input. The amplitude of the modulation signal is adjusted for the desired index of modulation. In general, the received FSK signal is passed through a limiting IF amplifier and fed to a noncoherent demodulator such as a frequency discriminator

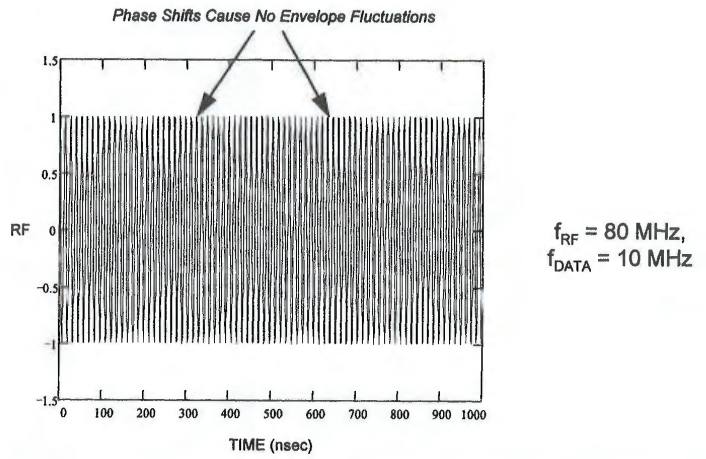


Figure 4.68 Time domain waveform of baseband-filtered MSK transmit signal (using 5-MHz Gaussian lowpass filter).

or quadrature detector. The image of the detected signal is filtered out and fed to a bit slicer or comparator to recover the original bit stream.

Figure 4.71 shows the time domain response for an FSK signal with $h = 1.0$. Notice that the modulated signal is constant amplitude, but the difference between the mark and space frequencies is larger due to the higher modulation index. With an $h = 1.0$, the total phase variation over one bit period is equal to 180 deg. This results in much larger—though still linear—phase variations over a bit period. The input bit stream is shown at the bottom of the figure.

In order to coherently demodulate the FSK signal, we need reference waveforms of the two possible transmitted frequencies. A block diagram of a coherent demodulator is shown in Figure 4.72. The received signal is applied to two double-balanced mixers. One mixer is driven by an LO at a frequency corresponding to a logical input of 0 on the modulator input. The other mixer is driven by the frequency that corresponds to a logic 1 on the modulator input. The corresponding outputs of the mixer are then integrated over a bit period and sampled. The outputs of the sample-and-hold circuits go to a decision circuit that decides which of the two signals was larger. Coherent FSK modulation is not often used in practice due to the difficulty in generating two reference frequencies close together at the receiver.

Figure 4.73 shows a plot of bit error probability of FSK versus E_b/N_o for both coherent and noncoherent signals. Not surprisingly, the noncoherent signal requires a higher SNR than the coherent signal in order to maintain the same P_E.

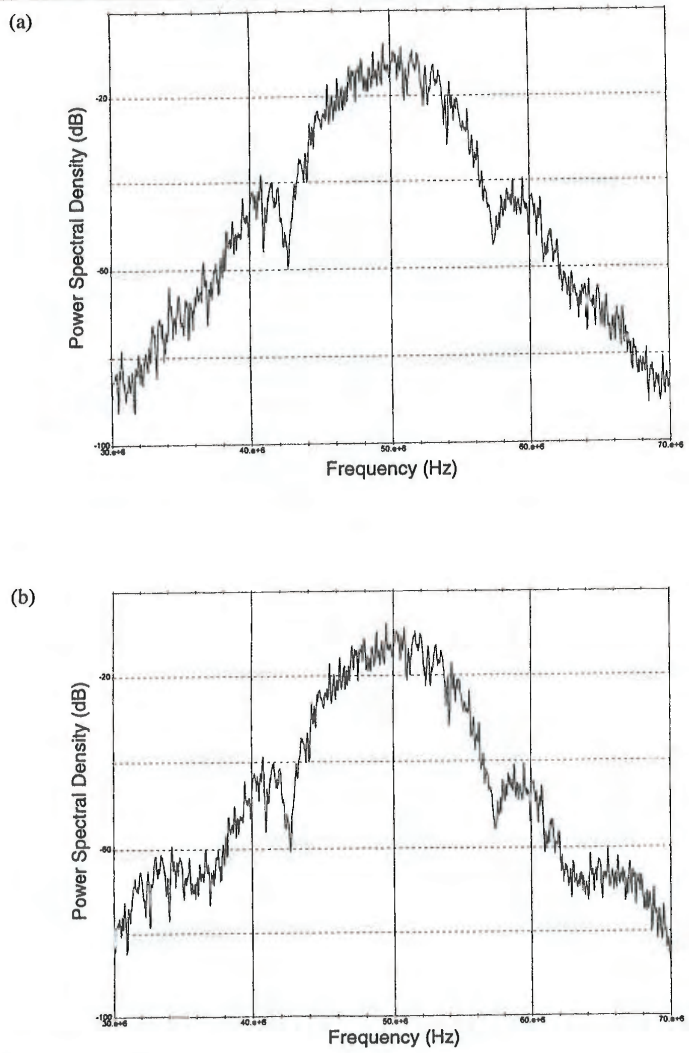


Figure 4.69 MSK transmit output spectrum with baseband filtering: (a) without limiting, (b) with limiting (amplifier 10 dB into compression).

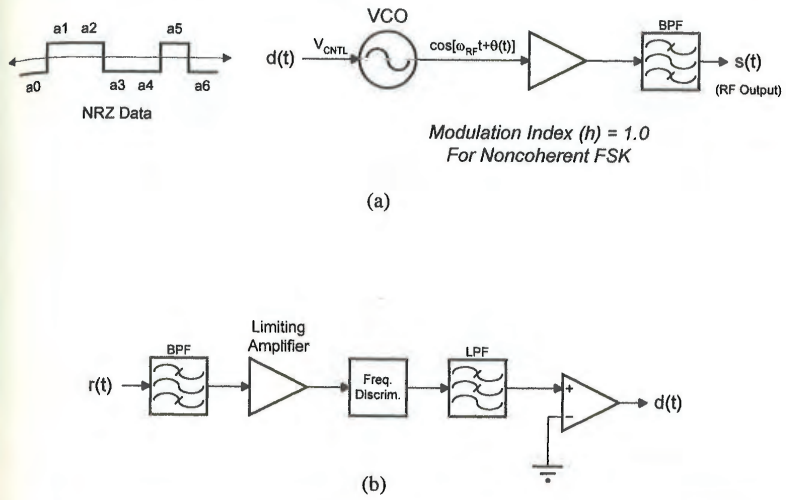


Figure 4.70 Noncoherent FSK modulator and demodulator: (a) FSK modulator, (b) FSK demodulator.

Figure 4.74 shows a plot of the power spectral density of an FSK signal with $h = 1.0$. Here, the carrier frequency is 80 MHz and the data rate is 10 Mbps. Notice that the spectrum is very similar to that of MSK, except that the main lobe is wider due to the larger modulation index. Also notice that FSK with $h = 1.0$ produces a spur at both the mark and space frequencies (though only one of the two spurs can be seen because of the single-sided nature of the plot). This spur has been used in some applications as a means to recover the carrier frequency.

Up to now, we have only been talking about using binary FSK modulation. As in the case of PSK, we can represent groups of bits by using different symbols. This is done in FSK by using more than two different frequencies. For example, we might choose to use four closely spaced frequencies as our symbol and thus reduce the data rate by one-half. Figure 4.75 shows a diagram of this. This is called 4-level FSK, or 4-FSK. As in the case of QPSK, the input bit stream is demultiplexed into two independent bit streams at half the original data rate. Each group of two bits, or symbols, is then represented by one of four different frequencies.

Figure 4.76 shows a block diagram of a 4-FSK modulator. An encoder is used to demultiplex the data and produce a corresponding digital word that is used to drive the DAC. The DAC output is then applied through a filter to the VCO. Figure

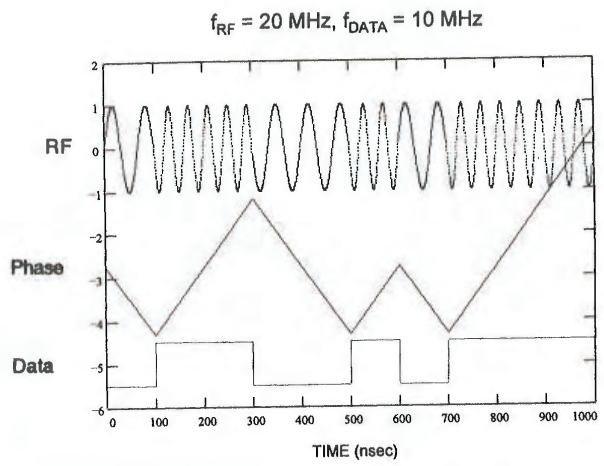


Figure 4.71 FSK time domain waveforms.

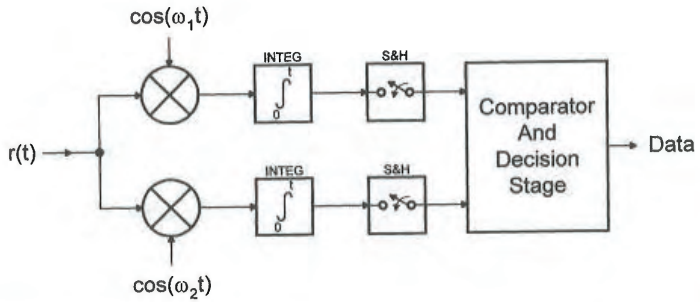


Figure 4.72 Coherent FSK demodulator.

4.77 shows a plot of a 4-FSK modulated signal with $h = 2.0$ having a carrier frequency of 80 MHz and a bit rate of 10 Mbps (corresponding to a symbol rate of 5 Msps). In this case, the maximum frequency deviation, Δf, is the frequency difference between the two outermost symbols. Notice the relatively flat-topped spectral response and the individual spurs at each of the corresponding symbol frequencies.

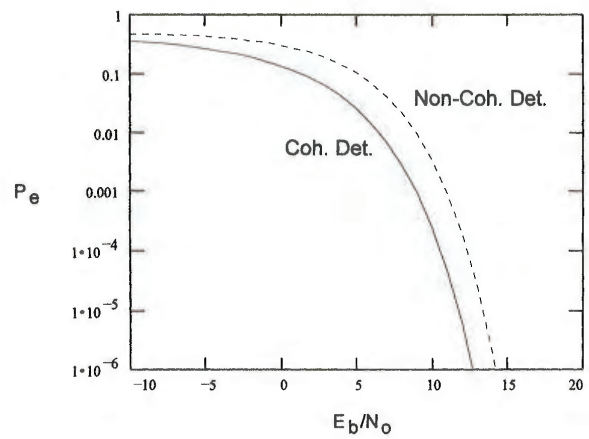


Figure 4.73 FSK probability of bit error.

Four-FSK can be demodulated by simply filtering and detecting each of the corresponding symbol frequencies. Figure 4.78 shows this brute-force approach. This approach is useful where the relative spacing of each symbol frequency is widely separated. This approach is commonly used in touch-tone phone decoders at audio frequencies. A disadvantage with this approach is that we essentially need as many different IF strips as we have symbol frequencies. If the IF center frequency is high and the spacing between each symbol frequency is small, it can be difficult to maintain adequate separation between each symbol frequency due to insufficient filter stop-band performance.

A more modern, and higher speed, approach is shown in Figure 4.79. In this approach, a single frequency discriminator is used. The output of the frequency discriminator produces a multilevel analog signal. The four different levels corresponding to each of the symbols are detected by means of an ADC. In the past, the relatively high complexity of ADCs made this approach less attractive. However, with modern IC techniques, it is possible to fabricate the ADC in the same IC as the decoder, and other portions of the receiver.

Figure 4.80 (a) shows a plot of the symbol error probability of a multilevel FSK signal versus E_b/N_0 for coherent demodulation. In the Figure, M is an integer that represents the number of symbols. Figure 4.80(b) shows the same plot for noncoherent demodulation. Comparing the two plots reveals that there is very little difference in the symbol error probabilities when the number of symbols is large and E_b/N_0 is high. This is an advantage of multilevel FSK.

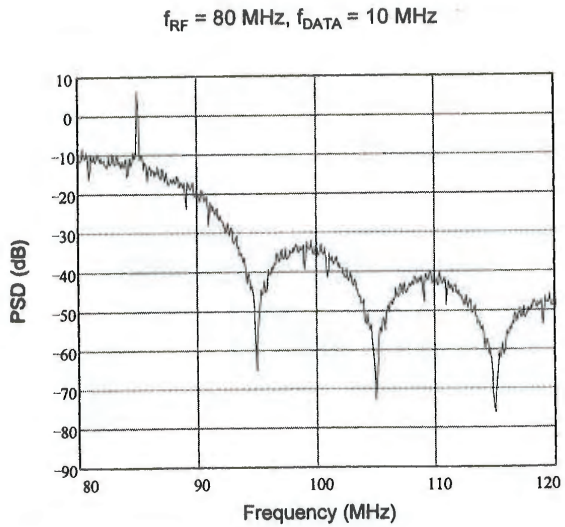


Figure 4.74 FSK (h = 1.0) power spectral density.

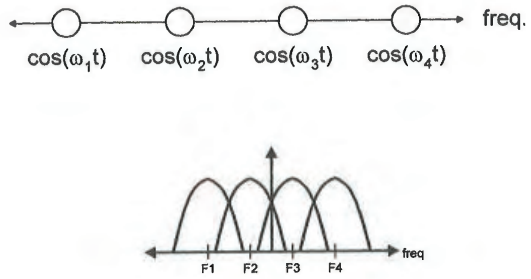


Figure 4.75 4-level FSK (4-FSK).

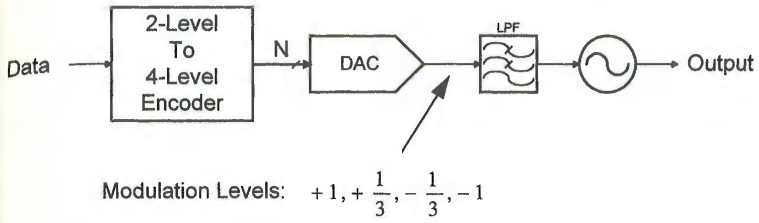


Figure 4.76 4-FSK modulator.

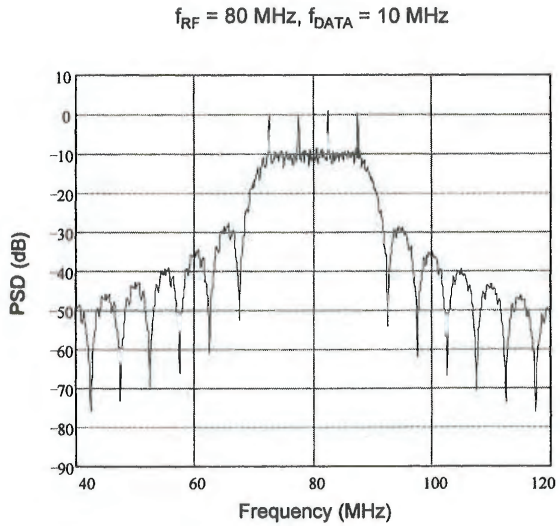


Figure 4.77 4-FSK power spectral density (h = 2.0).

4.4 RELATED MODULATIONS

The basic principles of digital phase and frequency modulation have been covered in the preceding sections. A number of variations are possible on each of these modulations, which can produce a number of additional benefits.

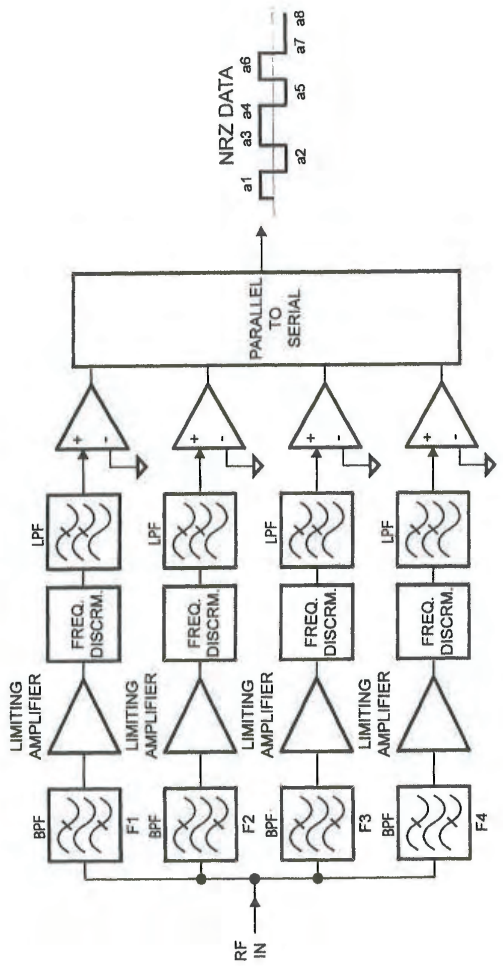


Figure 4.78 Brute force 4-level FSK demodulator (noncoherent).

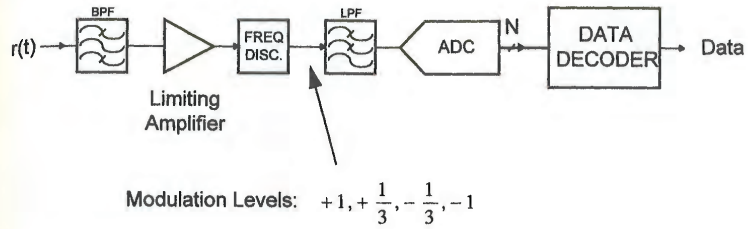


Figure 4.79 Modern 4-level FSK demodulator (noncoherent).

4.4.1 Gaussian-Filtered MSK (GMSK)

We have shown that baseband filtering of the modulation signal can result in a much narrower transmitted spectrum. However, as we saw in the case of QPSK, care must be taken when the signal is nonlinearly amplified to avoid spectral regrowth. The best way to avoid spectral regrowth is to use a constant envelope modulation, such as MSK. However, even then, the transmitted spectrum of MSK is still too wide for many applications.

One way to narrow the transmitted spectrum of MSK is by filtering the modulation signal before applying it to the VCO. This can be done by passing the modulation signal through a lowpass filter. It is important, however, that the lowpass filter have a well-behaved time domain response. A class of filters that has a well-behaved time domain response is Gaussian filters. The frequency response of a Gaussian lowpass filter is Gaussian in nature, following the following relation

$$H(\omega) = \tau \cdot \sqrt{2\pi} \cdot e^{-\frac{1}{2}(\tau\omega)^2} \tag{4.71}$$

This is the familiar $\exp(-x^2)$ shape of the Gaussian, or normal, probability density function (i.e., the classic "bell-shaped" curve). In the equation, ω is the frequency in rads/sec and τ is a constant.

A peculiar property of a filter with a Gaussian frequency response is that its impulse, or time domain, response is Gaussian as well. This can be seen by taking the inverse Fourier transform of the spectrum

$$h(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega) \cdot e^{-j\omega t} d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left[\tau \cdot \sqrt{2\pi} \cdot e^{-\frac{1}{2}(\tau\omega)^2}\right] \cdot e^{-j\omega t} d\omega \tag{4.72}$$

Solving the integral yields