IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re *Inter Partes* Review of:	)
U.S. Patent No. 10,833,908	)
Issued: Nov. 10, 2020	)
Application No.: 16/902,740	)
Filing Date: Jun. 16, 2020	)

For: **Channel Probing Signal For A Broadband Communication System**

DECLARATION OF INGRID HSIEH-YEE, PH.D.

TABLE OF CONTENTS

I. Introduction 4

A. Qualifications and Professional Experience 4

B. Scope of This Declaration 11

C. Person of Ordinary Skills in the Arts 13

D. Evidence Considered in Forming My Opinions 14

II. Authentication and Public Availability of *Kaiser* (Ex. 1007) 16

A. Authentication 16

B. IEEE Database Record for *Kaiser* 18

C. Linda Hall Library Copy 20

D. Linda Hall Library Records 21

E. Copyright Registration Record 24

F. Usage Record 25

G. Summary of My Opinion on *Kaiser* 26

III. Authenticity and Public Availability of *Gibson* (Ex. 1008) 27

A. Authentication 27

B. Internet Archive Digitized Copy 28

C. Library of Congress Records 29

D. Stanford University Record 34

E. Copyright Record 36

F. Usage Record 37

Ford Motor Co.
Exhibit 1030
Page 002

G. Summary of My Opinion on *Gibson* 37

IV. Conclusion..... 38

Ford Motor Co.
Exhibit 1030
Page 003

I. Introduction

I, Ingrid Hsieh-Yee, Ph.D., do hereby declare as follows:

1. I have personal knowledge of all of the matters about which I testify in this declaration.

2. I have been retained as an independent expert witness on behalf of Ford Motor Company ("FORD") for proceedings before the United States Patent and Trademark Office ("PTO") regarding the authenticity and public availability of certain prior art references in relation to US Patent No. 10,833,908 ("the '908 patent").

3. I am being compensated for my work in this matter at my customary hourly rate. I am also being reimbursed for any reasonable expenses associated with my work and testimony in this investigation. My compensation is not contingent on the results of my study, the substance of my opinions, or the outcome of this matter.

A. Qualifications and Professional Experience

4. My complete qualifications and professional experience are described in my academic curriculum vitae (**Appendix A**). The following is a brief summary of my relevant qualifications and professional experience.

5. In 2022 I was conferred the title of Professor Emerita after 32 years as a professor in the Department of Library and Information Science of the Catholic University of America. I have experience working in an academic library, a medical

Ford Motor Co.
Exhibit 1030
Page 004

library, and a legislative library. I hold a Ph.D. in Library and Information Studies from the University of Wisconsin-Madison and a Masters in Library and Information Studies from the University of Wisconsin-Madison.

6. I am an expert on library cataloging and classification and have published two editions on this subject, *Organizing Audiovisual and Electronic Resources for Access*: *A Cataloging Guide* (2000 and 2006). I taught a variety of courses, including Cataloging and Classification, Advanced Cataloging and Classification, Organization of Internet Resources, Organization of Information, Digital Content Creation and Management, Metadata, Internet Searches and Web Design, Information Literacy Instruction, Advanced Information Retrieval and Analysis Strategies, and The Information Professions in Society. I am familiar with metadata schema design and implementation. In my teaching I covered the design and implementation of metadata in libraries, databases, search engines, digital repositories, digital libraries, and digital archives. I also covered how the ways we organize information resources affect the discovery and access to digital resources on the Internet. My research interests cover cataloging and classification, information organization, metadata, information retrieval, information architecture, digital collections, scholarly communication, social media, user interaction with information systems, and others.

Ford Motor Co.
Exhibit 1030
Page 005

7. I am very familiar with a library cataloging encoding standard known as the "Machine-Readable Cataloging" standard, also known as "MARC," which became the national standard for sharing bibliographic data in the United States by 1971 and the international standard by 1973. MARC is the primary communications protocol for the transfer and storage of bibliographic metadata in libraries. Experts in my field reasonably rely upon MARC records when forming their opinions.

8. A MARC record consists of several fields, each of which contains specific data about the work. Each field is identified by a standardized, unique, three-digit code corresponding to the type of data that follows. **Appendix B** is a true and correct copy of Parts VII to X of "Understanding MARC Bibliographic: Machine-Readable Cataloging" (http://www.loc.gov/marc/umb/um07to10.html) from Library of Congress. *Understanding MARC Bibliographic* was originally published by the Follett Software Co. and copyrighted in 1988. Editions in 1989, 1990, 1994, 1998, 2000, 2003 and 2009 include changes. The latest edition was copyrighted in 2009. This document explains commonly-used MARC fields. For example, the personal author of the work is recorded in Field 100, the title is recorded in Field 245, publisher information is recorded in Field 260, the physical volume and characteristics of a publication are recorded in Field 300, and topical subjects are recorded in the 650 fields.

Ford Motor Co.
Exhibit 1030
Page 006

9. The Online Computer Library Center (OCLC) is the largest bibliographic network of the world, with more than 554 million records and thousands of member institutions (many of which are libraries of some type) in more than 100 countries. OCLC was founded in 1967 to promote and support library cooperation. According to the "Third Article, Amended Articles of Incorporation of OCLC Online Computer Library Center, Inc.," OCLC was created "to establish, maintain and operate a computerized library network and to promote the evolution of library use, of libraries themselves, and of librarianship, and to provide processes and products for the benefit of library users and libraries, including such objectives as increasing availability of library resources to individual library patrons and reducing the rate of rise of library per-unit costs, all for the fundamental public purpose of furthering ease of access to and use of the ever-expanding body of worldwide scientific, literary and educational knowledge and information" (source: https://www.oclc.org/content/dam/oclc/membership/articles-of-incorporation.pdf).

10. OCLC members can contribute original cataloging records in MARC to the system or derive cataloging records from existing records, an activity referred to as "copy cataloging." When an OCLC participating institution acquires a work, it can create an original MARC record for the work in OCLC's Connexion system (a system for catalogers to create and share MARC records), and the system will automatically generate a code for the date of record creation in the yymmdd format,

Ford Motor Co.
Exhibit 1030
Page 007

and the creating library's OCLC symbol is recorded in subfield "a" of the 040 field. Once the MARC record is in Connexion, it becomes available to other OCLC members for adoption to their local online catalogs (*i.e.*, copy cataloging).

11. After a MARC record is created in Connexion, it also becomes searchable and viewable on WorldCat, which is a free web portal for users to explore the collections of more than 10,000 libraries worldwide. The record in WorldCat, however, is not presented in MARC fields. Instead, the data elements are labeled to help users interpret the record. Thus, the information stored in MARC records in Connexion is available to the interested public through the free and user-friendly WorldCat web portal.

12. WorldCat (http://www.worldcat.org) is "the world's largest network of library content and services" and its features are summarized in "about WorldCat" (https://www.worldcat.org/about). WorldCat began operation in January 1998 as a free web portal to more than 10,000 library collections worldwide. Through WorldCat, users can search for information resources in their local libraries and libraries around the world. WorldCat allows users to search for books, CDs, videos, and many new types of digital content, such as audiobooks, in many languages. Users can also retrieve research materials and article citations with links to their full text. After an item is retrieved, WorldCat helps users identify a library nearby that holds the item or all the libraries that hold the item. WorldCat is an efficient way to

Ford Motor Co.
Exhibit 1030
Page 008

explore library collections on the Internet. WorldCat is publicly accessible, and there is no fee to perform searches on WorldCat.

13. Library online catalogs gained acceptance in the early 1980s and many libraries migrated their systems to the World Wide Web in the mid-1990s. Library online catalogs are based on MARC records that represent their collections in order to help the public understand what materials are publicly accessible in those libraries. Most libraries with online catalogs have made their catalogs freely available on the Web. These online catalogs offer user-friendly search interfaces. Strong user interest in keyword searches and the popularity of Google have led to the "googlization" of library search systems. As a result, many library catalogs now provide a single search box for users to conduct keyword searches, with additional support for searches by author, title, subject terms, and other data elements such as ISBN (International Standard Book Number). Library catalogs these days also offer features for users to narrow their search results by language, year, format, and other elements. Many libraries display MARC records on their online catalogs with labels for the data elements to help the public interpret MARC records. Many libraries also offer the option to display MARC records in MARC fields.

14. Libraries create MARC records for works they acquire to facilitate the discovery and access to works in their collections, which may include books, serials, motion pictures, and publications in other formats. Monograph cataloging is fairly

Ford Motor Co.
Exhibit 1030
Page 009

common in libraries, and most libraries make a newly cataloged monograph available to the public soon after the cataloging work is completed, usually within a week. Libraries can create original cataloging records or use an existing record in OCLC to create a copy cataloging record. As soon as the cataloging record is completed, it is added to the library's online catalog for users. If the record is an original record, it is also entered into the OCLC system. If it is a copy cataloging record, the library's holding symbol is attached to the existing original record in the OCLC system to facilitate searching and interlibrary loan.

15. Conference proceedings can be cataloged as monographs or serials, depending on a library's policy. Libraries usually choose to catalog conference proceedings as monographs because records for monographs are easier for users to discover and understand, and access to individual conference proceedings becomes easier. Once a record for a conference proceedings is found, users can see the title, conference name, subjects, and related contributors clearly identified in the record and the call number will enable users to access the desired title quickly. Conference proceedings cataloged as monographs are treated as regular monographs, and the newly cataloged monographs typically become publicly accessible soon after the cataloging work is completed, usually within a week.

Ford Motor Co.
Exhibit 1030
Page 010

16. I am personally familiar with many online catalogs, databases, and search engines. In preparing for this declaration, I used the following authoritative information systems to search for records:

- Google Scholar (https://scholar.google.com)
- Internet Archive (https://archive.org)
- Online catalog of Stanford University Library (https://searchworks.stanford.edu/)
- Online catalog of the Library of Congress (https://catalog.loc.gov)
- Online catalog of the Linda Hall Library (https://catalog.lindahall.org/)
- Public catalog of the US Copyright Office (https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First)
- WorldCat (free for public access on the Web) (https://www.worldcat.org)

B. Scope of This Declaration

17. I have been asked to offer an opinion on the authenticity and public availability date of the following documents:

> Kaiser, S., & Fazel, K., (1997, September), "A flexible spread-spectrum multi-carrier multiple-access system for multi-media applications," in *Waves of the Year 2000+ : PIMRC '97 : The Eighth IEEE International Symposium on Personal, Indoor and Mobile Radio*

Ford Motor Co.
Exhibit 1030
Page 011

Communications, September 1-4, 1997, Marina Congress Center, Helsinki, Finland, vol. 1, pp. 100-104, **Ex. 1007** ("*Kaiser*");

Gibson, J. D., (1999), *The Mobile Communications Handbook*, 2nd ed., CRC Press ; IEEE Press, **Ex. 1008** ("*Gibson*").

18. I am rendering my expert opinion on the authenticity of the documents referenced herein and on whether they were published. I will provide my opinion on when and how these documents were publicly disseminated or otherwise made publicly available to the extent that persons interested and ordinarily skilled in the subject matter or art, exercising reasonable diligence, could have located and retrieved these documents.

19. I am informed by counsel that an item is considered authentic if there is sufficient evidence to support a finding that the item is what it is claimed to be. I am also informed that authenticity can be established based on the contents of the documents themselves, such as the appearance, content, substance, internal patterns, protocols followed or other distinctive characteristics of the item.

20. I am informed by counsel that a given reference is "published" if it is publicly accessible upon a satisfactory showing that such a document has been disseminated or otherwise made publicly available to the extent that persons interested and ordinarily skilled in the subject matter or art could locate it exercising

Ford Motor Co.
Exhibit 1030
Page 012

reasonable diligence and obtain the document. I have also been informed by counsel that materials available in a library constitute "printed publications" if they are cataloged and indexed according to general library practices and protocols that make the references available and accessible to members of the interested public.

21. I have been informed by counsel that the '908 patent was filed in the United States on June 16, 2020 and that Patent Owner alleges that the '908 patent is entitled to a priority date of no later than January 29, 2004. I have assessed public accessibility of the documents referenced below as of this date.

C. Person of Ordinary Skills in the Arts

22. I am told by counsel that the subject matter of this proceeding relates to wireless communication systems.

23. I have been informed by counsel that a "person of ordinary skill in the art at the time of the inventions" (POSA) is a hypothetical person who is presumed to be familiar with the relevant field and its literature at the time of the inventions. This hypothetical person is also a person of ordinary creativity, capable of understanding the scientific principles and literature applicable to the pertinent field.

24. I am told by counsel that a person of ordinary skill in this subject matter or art in the early 2004 timeframe would have had a bachelor's degree in electrical engineering, computer engineering, computer science, or an equivalent field, or an

Ford Motor Co.
Exhibit 1030
Page 013

advanced degree in those fields, as well as at least 3-5 years of academic or industry experience in mobile wireless communications, or comparable industry experience.

25. It is my opinion that such a person would have been actively engaged in learning about the field, possibly through formal instruction through the bibliographic resources. By at least early 2004 such a person would have had access to a vast array of print and electronic resources, including at least the documents referenced below.

D. Evidence Considered in Forming My Opinions

26. In the preparation of this declaration, I have reviewed the documents referenced below and any other documents I reference herein, and each of these is a type of material that experts in my field would reasonably rely upon when forming their opinions:

(1) The document referenced above in Section I.B;

(2) IEEE metadata record for *Kaiser*, available at https://ieeexplore.ieee.org/abstract/document/624371, and the front matter of *Waves of the Year 2000+ : PIMRC '97* that contains *Kaiser*, available at https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=626981, accessed and obtained on March 18, 2023, **Appendix 1007A**;

(3) Library copy of *Kaiser*, obtained from the Linda Hall Library, **Appendix 1007B**;

(4) Bibliographic and MARC records for *Waves of the Year 2000+ :*

Ford Motor Co.
Exhibit 1030
Page 014

PIMRC '97 : The Eighth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, September 1-4, 1997, Marina Congress Center, Helsinki, Finland that contains *Kaiser*, available at https://catalog.lindahall.org/permalink/01LINDAHALL_INST/19lda7s/alma992781953405961, accessed and obtained on March 18, 2023, **Appendix 1007C**;

(5) Copyright registration record for *Waves of the year 2000+ : PIMRC '97 : the 8th IEEE international symposium on personal, indoor and mobile radio communications, September 1-4, 1997, Marina Congress Center, Helsinki, Finland : technical program : proceedings* that contains *Kaiser*, available at https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?v1=6&ti=1,6&Search%5FArg=International%20Symposium%20on%20Personal%2C%20Indoor%20and%20Mobile%20Radio%20Communications&Search%5FCode=TALL&CNT=25&PID=CwLYnVyxip0rzeUTEZg8Mlggv7Mhw&SEQ=20230318124753&SID=5, accessed and obtained on March 18, 2023, **Appendix 1007D**;

(6) Citations to *Kaiser*, obtained from Google Scholar, **Appendix 1007E**;

(7) Internet Archive copy of *Gibson* and record for *Gibson*, available at https://archive.org/details/mobilecommunicat0000unse_s8j1, accessed and obtained on March 19, 2023, **Appendix 1008A**;

(8) Bibliographic and MARC records for *Gibson*, available at https://lccn.loc.gov/98046558 from the online catalog of the Library

Ford Motor Co.
Exhibit 1030
Page 015

of Congress, accessed and obtained on March 19, 2023, **Appendix 1008B**;

(9) Bibliographic and MARC records for *Gibson*, available at https://searchworks.stanford.edu/view/4047173 from the online catalog of the library of Stanford University, accessed and obtained on March 19, 2023, **Appendix 1008C**;

(10) Copyright registration record for *Gibson*, available at https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?Search_Arg=waves+of+the+year+2000%2B&Search_Code=TALL&PID=mo8BsLvX_HsIg9Ptpwbxq9VVZNW7L&SEQ=20230326091405&CNT=25&HIST=1 from the public catalog of the US Copyright Office, accessed and obtained on March 19, 2023, **Appendix 1008D**;

(11) Citations to *Gibson*, obtained from Google Scholar, **Appendix 1008E**.

II. Authentication and Public Availability of *Kaiser* (Ex. 1007)

A. Authentication

27. **Ex. 1007** is a true and correct copy of "A flexible spread-spectrum multi-carrier multiple-access system for multi-media applications," ("Kaiser"), by S. Kaiser and K. Fazel, in *Waves of the Year 2000+ : PIMRC '97 : The Eighth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, September 1-4, 1997, Marina Congress Center, Helsinki, Finland*, vol. 1, pp. 100-104, that I obtained from the Institute of Electrical and Electronics Engineers (IEEE).

Ford Motor Co.
Exhibit 1030
Page 016

When I began preparing this declaration I searched Google Scholar for records for *Kaiser*, and the link from the retrieved record led me to the IEEE metadata record for this article in the IEEE Xplore Digital Library. I then used the PDF link to purchase this article from the publisher. This IEEE copy is presented as **Ex. 1007** in this declaration.

28. **Ex. 1007** is a true and correct copy of *Kaiser* that I obtained from IEEE Xplore, the digital library created and maintained by the Institute of Electrical and Electronics Engineers. Page 1 (internal page 100) shows the title, authors, and abstract of this article. The bottom of this page shows the article appears in a publication with an ISBN (International Standard Book Number) of "0-7803-3871-5" and has a "1997" copyright date with the IEEE as the copyright holder. **Ex. 1007** shows *Kaiser* has a total of five pages, including six equations, five figures, one table and 15 references.

29. The IEEE copy of *Kaiser* (**Ex. 1007**) is in a condition that creates no suspicion about its authenticity. Specifically, this copy is not missing any intermediate pages, the text on each page appears to flow seamlessly from one page to the next, and there are no visible alterations to the document. Moreover, **Ex. 1007** was found in the custody of a digital library owned and managed by the publisher, a place where, if authentic, it would likely be found. Accordingly, I see no reason to question the authenticity of the IEEE copy of *Kaiser* (**Ex. 1007**).

Ford Motor Co.
Exhibit 1030
Page 017

B. IEEE Database Record for *Kaiser*

30. **Appendix 1007A** is a true and correct copy of the metadata record for *Kaiser* obtained from the IEEE Xplore Digital Library. I personally located, identified, and obtained this record. This is the type of record experts in my field would reasonably rely upon when forming their opinions.

31. Page 1 of **Appendix 1007A** shows the title and authors of the *Kaiser* article. It also shows IEEE as the publisher, and indicates *Kaiser* has 41 paper citations, 12 patent citations, and 203 full text views. The abstract in this record is identical to the abstract in the IEEE copy (**Ex. 1007**). Page 2 of the IEEE metadata record shows *Kaiser* was published in the "Proceedings of 8th International Symposium on Personal, Indoor and Mobile Radio Communications – PIMRC '97," the conference date is "01-04 September 1997" and the conference location is "Helsinki, Finland." The article DOI (digital object identifier) is "10.1109/PIMRC.1997.624371" and the "Date Added to IEEE Xplore" field shows this metadata record was added to this digital library on "06 August 2002." The record also shows the conference proceedings has an ISBN of "0-7803-3871-5." The title, authors, abstract, and ISBN in this record match the information contained in the publisher's copy (**Ex. 1007**), confirming this record represents *Kaiser*.

32. The IEEE metadata record for *Kaiser* includes a direct link to the conference, and the link leads to an IEEE record for this conference (available at

Ford Motor Co.
Exhibit 1030
Page 018

https://ieeexplore.ieee.org/xpl/conhome/4928/proceeding). A screen capture of the top of this record is included as page 4 in **Appendix 1007A**. The front matter of this conference is freely available to the public at https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=626981, and is included in **Appendix 1007A** as pages 5 to 24.

33. Page 5 of **Appendix 1007A** is the title page that shows "Waves of the Year 2000+, The 8th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications" is the title. The title page also shows the time, location, and sponsor of the conference, and the contents of the Technical Program Proceedings in three volumes. Page 6 is the copyright page that shows a "1997" copyright date and the Institute of Electrical and Electronics Engineers, Inc. as the copyright holder. It also shows 97TH8271 as the IEEE catalog number, 0-7803-3871-5 (softbound) and 0-7803-3872-3 (microfiche) casebound) as the ISBNs, and 97-70761 as the Library of Congress number. This page also shows the IEEE Service Center is located at Piscataway, NJ. Pages 7 to 24 (internal page I to XVIII) are the table of contents and page 8 (internal page II) shows *Kaiser* is included in the "C5 Multicarrier CDMA" section and identified as C5.1, running from pages 100 to 104.

34. Data in the IEEE metadata record for *Kaiser* and the front matter of the print version of the conference proceedings containing *Kaiser* inform my opinion that *Kaiser* was published in 1997, and IEEE Xplore indexed the article and added

Ford Motor Co.
Exhibit 1030
Page 019

the metadata record into the IEEE Xplore Digital Library on August 6, 2002. The IEEE record made *Kaiser* discoverable in the IEEE Xplore system from that date on. Interested users in August 2002 would have been able to locate the metadata record for *Kaiser* and purchase the article from IEEE or obtain it from an institution that subscribed to IEEE Xplore Digital Library.

C. Linda Hall Library Copy

35. To determine if *Kaiser* was available in libraries, I searched WorldCat by the title of the conference proceedings and found a record at https://www.worldcat.org/title/37841842 that identified Linda Hall Library as one of the holding libraries. After the holdings information was verified in the library's online catalog, I requested a copy of Kaiser from their Document Delivery Service. This library copy is presented as **Appendix 1007B** in this declaration.

36. **Appendix 1007B** is a true and correct copy of *Kaiser* obtained from Linda Hall Library. Pages 1 to 5 (internal pages 100 to 104) are the full text of this article. I have closely compared this copy with the publisher copy (**Ex. 1007**) and found them to have the same content.

37. Page 6 is the front cover of the conference proceedings containing Kaiser, and it indicates this is volume 1 of 3. Page 7 is the copyright page that shows a hand-written call number of "TK5103.2 .I33 1997 v. 1 c. 2" and four identifiers for this publication. Page 8 shows greetings from two conference organizers, and the

Ford Motor Co.
Exhibit 1030
Page 020

upper left corner shows a hand-written note of "2.10.98" to indicate this publication was added to the Linda Hall Library Collection on February 10, 1998. The knowledge of this annotation practice is based on my personal experience with this library's collection and staff. Page 9 is the title page of volume 1 of 3. Pages 10 to 27 (internal pages I to XVIII) are the table of contents. I have closely compared pages 9 to 27 of the conference front matter in the Linda Hall Library copy (**Appendix 1007B**) with their corresponding pages included in the IEEE copy (**Appendix 1007A**) and found them to have the same content.

38. The Linda Hall Library copy of *Kaiser* (**Appendix 1007B**) is in a condition that creates no suspicion about its authenticity. Specifically, this copy is not missing any intermediate pages, the text on each page appears to flow seamlessly from one page to the next, and there are no visible alterations to the document. In addition, the front matter from this copy is the same as the front matter from IEEE Xplore (**Appendix 1007A**). Moreover, **Appendix 1007B** was found in the custody of a research library, a place where, if authentic, it would likely be found. Accordingly, I see no reason to question the authenticity of the Linda Hall Library copy of *Kaiser* (**Appendix 1007B**).

D. Linda Hall Library Records

39. **Appendix 1007C** is a true and correct copy of the bibliographic and MARC records for the conference proceedings containing Kaiser. I personally

Ford Motor Co.
Exhibit 1030
Page 021

located, identified and obtained these records from the online catalog of the Linda Hall Library. These are the type of records experts in my field would reasonably rely upon when forming their opinion.

40. Pages 1 to three are the bibliographic record that identifies this publication as a conference proceeding, and shows that it is available at "Linda Hall Library Closed Stacks – CSA Materials (TK5103.2 .I33 1997)." The Location Items area shows the library holds three volumes of this conference proceedings and all of them are "loanable" (meaning available for user request). Pages 2 and 3 show title, variant titles, creators, subjects, identifiers, publisher, creation date, format, and general notes about this publication.

41. Page 4 is the MARC record and the first six digits of Field 008 show the record was created on "971024" (*i.e.*, October 24, 1997). Subfields "a" and "c" of Field 040 show "LHL" created the MARC record. According to the Directory of OCLC Members (https://www.oclc.org/en/contacts/libraries.html), "LHL" is the OCLC symbol for Linda Hall Library. Field 005 is a machine-generated field that records the last time this library modified this MARC record, and it shows "19971024143200" (i.e., October 24, 1997, 2 p.m., 32 minutes) as the last date and time of record modification. Data from Fields 008, 040 and 005 inform my opinion that Linda Hall Library cataloged the conference proceedings containing Kaiser and completed the MARC record on October 24, 1997. This MARC record became

Ford Motor Co.
Exhibit 1030
Page 022

searchable in the library's online catalog from that date on.

42. The date note in the Linda Hall Library copy of Kaiser (Appendix 1007B) shows copy 2 was added to the library collection on February 10, 1998, but the MARC record became searchable on October 24, 1997. In consultation with the library staff, I learned that copy 1 became missing, so copy 2 was added. It is therefore my opinion that the conference proceedings containing Kaiser was cataloged and became discoverable as of October 24, 1997, and copy 2 of this publication was added to the library collection on February 10, 1998.

43. Field 010 of the MARC record shows the Library of Congress control number, Fields 020 show the ISBNs, Field 111 shows the conference name as the author, Field 245 shows the title, Fields 246 show four variant titles are included to help users discover this publication, Field 260 shows publisher information and copyright date, Field 300 shows this work contains three volumes with illustrations, Field 500 show the IEEE catalog number for this work, Field 504 shows the work includes bibliographical references and indexes, and Field 710 shows the IEEE Communications Society is included to help users discover this conference proceedings.

44. Subjects of this work are represented in Field 090 by a Library of Congress Classification (LCC) number of "TK5103.2" that represents general works on "Wireless communication systems. Mobile communication systems." and in

Ford Motor Co.
Exhibit 1030
Page 023

Fields 650 by four Library of Congress subject headings, with "Wireless communication systems," "Mobile communication systems," "Personal communication service systems," and "Internetworking (Telecommunication)" as the main topics, each followed by a form subheading "Congresses" encoded in subfield "v" to indicate the main topics are treated at a conference.

45. This MARC record (**Appendix 1007C**) shows that Linda Hall Library cataloged the conference proceedings containing *Kaiser* and completed the MARC record on October 24, 1997, making this record searchable in their online catalog from that date on. The Linda Hall Library copy of *Kaiser* (**Appendix 1007B**) shows that copy 2 of this conference proceedings was added to the library collection on February 10, 1998. A customary library practice is to process a newly cataloged work promptly for user access, usually within a week after the cataloging work is completed. It is therefore my opinion that the conference proceedings containing *Kaiser* would have become publicly accessible as early as October 31, 1997 (for copy 1), and no later than February 17, 1998 (for copy 2).

E. Copyright Registration Record

46. **Appendix 1007D** is a true and correct copy of the copyright registration record for *Waves of the year 2000+ : PIMRC '97 : the 8th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications* that contains *Kaiser*. I obtained this record from the public catalog of the U.S. Copyright Office

Ford Motor Co.
Exhibit 1030
Page 024

by first accessing their website at http://www.copyright.gov/, selecting the "Learn More" link under "Search Copyright Records," and following the link for "Official Public Catalog (1978 to present)" under "Online Records Collections" to perform a title search to retrieve the record. I personally located, identified and obtained this record. This is the type of record experts in my field would reasonably rely upon when forming their opinions.

47. The "Type of Work" field identifies the publication as "Text." The application title, title, copyright claimant and other title match the information presented in the front matter of this conference proceedings included in the IEEE copy (**Appendix 1007A**) and the Linda Hall Library copy (**Appendix 1007B**). The "Date of Creation" field shows this work was created in 1997, and the "Date of Publication" field shows "1997-09-01" (*i.e.*, September 1, 1997). The "Registration Number / Date" field shows "TX0004659742 / 1997-11-03" indicating this conference proceedings was registered with the Copyright Office on November 3, 1997. This copyright registration record informs my opinion that the conference proceedings containing *Kaiser* was published on September 1, 1997, and registered with the U.S. Copyright Office on November 3, 1997.

F. Usage Record

48. Actual usage of a publication is reflected by the papers that make reference to it. The citation history on Google Scholar shows that *Kaiser* has been

Ford Motor Co.
Exhibit 1030
Page 025

cited 106 times. **Appendix 1007E** is a list of citations from 1999 and 2000 to demonstrate early usage. The earliest citation is an IEEE article published in February 1999.

G. Summary of My Opinion on *Kaiser*

49. The IEEE copy of *Kaiser* (**Ex. 1007**) shows it was published in 1997, and the IEEE record for this article (**Appendix 1007A**) shows the metadata record was entered into the IEEE Xplore Digital Library on August 6, 2002, making *Kaiser* publicly discoverable from that date on. Because IEEE Xplore is freely accessible on the Internet, interested users would have been able to search the system to discover *Kaiser*. After that, they could obtain free access to the article through an institution with a subscription to IEEE Xplore Digital Library, or purchase the article from IEEE themselves. The copyright registration record for the conference proceedings containing Kaiser (**Appendix 1007D**) shows the conference proceedings was published on September 1, 1997. The MARC record from the Linda Hall Library (**Appendix 1007C**) shows the conference proceedings containing *Kaiser* became discoverable in their online catalogs by October 24, 1997, and the Linda Hall Library copy of Kaiser (**Appendix 1007B**) shows copy 2 was added to their collection on February 10, 1998, meaning this copy would have become publicly accessible no later than February 17, 1998. The citation history shows the earliest citation to *Kaiser* is an IEEE publication that appeared in February 1999. All

Ford Motor Co.
Exhibit 1030
Page 026

the records I have reviewed support my opinion that the conference proceedings containing *Kaiser* became discoverable by October 24, 1997, and copies of the proceedings (and *Kaiser* contained therein) would have been publicly accessible as early as October 31, 1997, and no later than February 17, 1998. These public availability dates are several years before the January 2004 critical date.

III. Authenticity and Public Availability of *Gibson* (Ex. 1008)

A. Authentication

50. **Ex. 1008** is a true and correct copy of *The Mobile Communications Handbook*, 2nd ed., ("*Gibson*"), by J. D. Gibson, published by CRC Press/IEEE Press in 1999, that I obtained from counsel. When I was asked to prepare this declaration, I searched WorldCat by the title of this book for records, and found a record at https://www.worldcat.org/title/39985639 that informed me the Library of Congress is one of the holding libraries. I then verified the holdings information in the online catalog of the Library of Congress to obtain the bibliographic and MARC records for this work. These records are discussed below.

51. **Ex. 1008** is a scanned copy of *Gibson*. I understand from counsel that Volkswagen Group of America submitted this scanned copy of *Gibson* with its *inter partes* review petition for the '908 patent. Page 1 is the cover that shows the title, edition number, editor, and publishers. Page 2 shows the spine and page 3 is the back cover that shows "8597" as the CRC Press catalog number for this book, "PC5772"

Ford Motor Co.
Exhibit 1030
Page 027

as the IEEE order number for this book, and "0-8493-8597-0" as the ISBN (International Standard Book Number) assigned to this work. Page 4 is the half title page, page 5 is the series page, page 6 is the title page that shows the same information as the front cover, and page 7 is the copyright page that shows a "1999" copyright and CRC Press LLC as the copyright holder. This page also shows a "Library of Congress Cataloging-in-Publication data" indicating a CIP record was created for this work before its publication. Pages 13 and 14 are the table of contents that shows *Gibson* has a total of 36 chapters, with the first 14 chapters in Section I Basic Principles, and Chapters 15 to 36 in Section II Wireless, and an index that is 19 pages long. The book includes many equations, figures, and color maps, and each chapter has its own page sequence.

B. Internet Archive Digitized Copy

52. To authenticate **Ex. 1008**, I located a digitized version of *Gibson* on Internet Archive that is available at https://archive.org/details/mobilecommunicat0000unse_s8j1. Screen captures of portions of this digitized version on Internet Archive are included as **Appendix 1008A** in this declaration. Page 1 shows the front cover of *Gibson* on the Internet Archive platform and indicates users can borrow the book for one hour at a time or purchase the book at Better World Books. Page 2 shows a "withdrawn" stamp indicating the book came from a college in Plano Texas. Page 3 shows this version

Ford Motor Co.
Exhibit 1030
Page 028

was digitized by Internet Archive in 2022, and the page on the right is the half title page. Page 4 shows the series page and the title page. Page 5 shows the copyright page and the preface. Page 6 shows the back cover. Pages 7 and 8 show the metadata record for this digitized version of *Gibson* that shows the book was originally published in 1999 by CRC Press and IEEE Press of Boca Raton, Fla, and the digitized copy was added to Internet Archive on "2022-04-07" (*i.e.*, April 7, 2022). I have closely compared **Ex. 1008** with the digitized copy on Internet Archive (**Appendix 1008A**), and concluded that these two copies have the same content.

53. **Ex. 1008** is in a condition that creates no suspicion about its authenticity. Specifically, **Ex. 1008** is not missing any intermediate pages (within their sequences), the text on each page appears to flow seamlessly from one page to the next, and there are no visible alterations to the document. Furthermore, **Ex. 1008** has the same content as the digitized copy on Internet Archive (**Appendix 1008A**), confirming that the pages in **Ex. 1008** are from a published and publicly accessible work. I therefore see no reason to question the authenticity of **Ex. 1008**.

C. Library of Congress Records

54. As evidence of the public accessibility of *Gibson* in libraries, I obtained records for this book from the Library of Congress online catalog. **Appendix 1008B** contains a true and correct copy of the bibliographic and MARC records for the second edition of *The Mobile Communications Handbook* by Jerry D. Gibson. I

Ford Motor Co.
Exhibit 1030
Page 029

personally identified, located, and obtained these records from the online catalog of the Library of Congress. They are the type of records experts in my field would reasonably rely upon when forming their opinion.

55. Pages 1 and 2 of **Appendix 1008B** are the bibliographic record that shows the title, edition statement, authors, publisher, and physical description of *Gibson*. It also shows the ISBN of this book is "0849385970" and the LCCN (Library of Congress control number) for this book is "98046558." The "Item Available" area shows the book's call number is "TK6570.M6 M5934 1999 Copy 1," the status of the book is "Not Charged" (meaning available for user request), and users can request this book at the "Jefferson or Adams Building Reading Rooms."

56. Page 3 is the MARC record that shows the date of record creation and revision. The first six digits of Field 008 show the MARC record was created on "980924" (*i.e.*, September 24, 1998), and the "s1999" code following this creation date indicates this publication has a single publication date of 1999. Field 955 is a field for recording local cataloging activities, and it shows the MARC record was created on "09-24-98," and a copy of the book was sent to CIP on "03-12-99." Subfields "a" and "c" of Field 040 shows that "DLC" created the original record. According to the *Directory of OCLC Members* (https://www.oclc.org/en/contacts/libraries.html), "DLC" is the OCLC library symbol for the Library of Congress. Fields 008, 040 and 955 inform my opinion that

Ford Motor Co.
Exhibit 1030
Page 030

the original MARC record was created as a Cataloging-in-Publication (CIP) record based on the information supplied by the publisher. After the CIP record was completed, the CIP record was included in the OCLC Connexion system for OCLC members to use for copy cataloging and collection management purposes. The CIP record was also entered into the online catalog at the Library of Congress to facilitate the discovery of this book. In addition, the CIP record was sent to the publisher for inclusion in the published work. The block of CIP data on the copyright page of *Gibson* (**Ex. 1008** and **Appendix 1008A**) reflects this practice.

57. Participants of the Cataloging-in-Publication (CIP) Program at the Library of Congress are obligated to send a copy of the work when it is published, and the catalogers at the Library of Congress will use the published work to verify and enrich the original CIP record. Field 955 shows that CIP verification took place on "03-12-99" (meaning March 12, 1999), and the cataloging work is then concluded. A customary practice of libraries is to process a newly cataloged monograph quickly to make it publicly accessible, usually within a week after the cataloging work is completed. Considering the large amount of works processed by the Library of Congress and my personal experience with their collections, it is my opinion that *Gibson* would have been publicly accessible one to three months after the cataloging work was completed in March 1999. This means that a copy of *Gibson* would have been publicly accessible at the Library of Congress as early as April 12,

Ford Motor Co.
Exhibit 1030
Page 031

1999, and no later than June 12, 1999.

58. The Cataloging in Publication Program (CIP) of the Library of Congress is responsible for cataloging books in advance of publication to alert the library community to forthcoming publications and facilitate acquisition. The initial CIP record is created based on information submitted by the publisher. The Library of Congress distributes CIP records in MARC format to large libraries, bibliographic services, and book vendors worldwide.

59. CIP records are also entered into OCLC Connexion system, the cataloging system containing more than 554 million MARC records, for OCLC member libraries to use it for copy cataloging and acquisition purposes. After the book is published, the publisher is obligated to send a copy to the CIP Program, and Library of Congress catalogers use the newly received publication to verify and update data in the original CIP record.

60. Field 010 of the MARC record (**Appendix 1008B**) shows the Library of Congress control number assigned to *Gibson* and Field 020 shows the book's ISBN. Field 245 shows the title and author statement, Field 250 shows the edition statement, Field 260 shows the publisher and copyright date, Field 300 shows physical description, Field 500 shows two identifiers, Field 504 shows the presence of bibliographic references and an index, and Field 700 shows the editor is included to help users discover this work. The LCCN, ISBN, title, publisher, publication

Ford Motor Co.
Exhibit 1030
Page 032

place, publication date, physical description, and the notes match the information presented in **Ex. 1008** and the digitized copy on Internet Archive (identified in **Appendix 1008A)** and confirm this MARC record represents *Gibson*.

61. Subjects of this book are represented in Field 050 by a Library of Congress Classification (LCC) number of "TK6570.M6" that represents "Mobile communication systems" and in Field 082 by a Dewey Decimal Classification (DDC) number of "621.3845" that represents "Radiotelephony." Subjects are also represented in Field 650 by a Library of Congress subject heading, "Mobile communication systems." These subject representations enable users to conduct subject searches to find this book by *Gibson*.

62. This MARC record (**Appendix 1008B**) shows that the Library of Congress created a CIP record *Gibson* on September 24, 1998, making it discoverable from that date on in their online catalog. Users would have been able to conduct subject searches by the LCC number, DDC number and the Library of Congress subject heading in the MARC record to discover this book. They would also have been able to search for *Gibson* by the author, title, ISBN, and LCCN. After *Gibson* was published in 1999 and the cataloging work was completed at Library of Congress in March 1999, the copy would have become publicly accessible no later than June 1999.

Ford Motor Co.
Exhibit 1030
Page 033

D. Stanford University Record

63. My research in WorldCat on *Gibson* also identified Stanford University Library as another holding library. **Appendix 1008C** is a true and correct copy of the bibliographic and MARC records for *Gibson*. I personally located, identified and obtained these records. These are the type of material experts in my field would reasonably rely upon when forming their opinions.

64. Pages 1 to 4 of **Appendix 1008C** are the bibliographic record for *Gibson* that shows the editor, edition statement, imprint, physical description, series, bibliographic note, summary, subject, publication date, notes and ISBNs assigned to this work. The "At the library" area on page 2 shows this work is stored at "SAL3 (off-campus storage) No public access" and its call number is "TK6570 .M6 M5934 1999" and users can request access to this book.

65. Pages 5 and 6 are the MARC record. Fields 008 and subfields (a" and "c" of Field 040 show "DLC" (*i.e.*, Library of Congress) created the original MARC record on "980924" (*i.e.*, September 24, 1998), and subfield "d" of Field 040 shows "CSt" modified the Library of Congress MARC record later. According to the *MARC Code List for Organizations* (https://www.loc.gov/marc/organizations/), "CSt" is the MARC code for Stanford University. Field 916 is a local field recording their cataloging activity, and it shows "Date Cataloged 19990325" (meaning the cataloging work was completed on March 25, 1999). Data from Fields 008, 040 and

Ford Motor Co.
Exhibit 1030
Page 034

916 inform my opinion that Library of Congress created a CIP record on September 24, 1998, before the book was published, and the Stanford University Library used the CIP record for copy cataloging and completed their MARC record on March 25, 1999. This MARC record made *Gibson* discoverable in the online catalog of the Stanford University Library from that date on. A customary library practice is to process a newly cataloged monograph promptly for public access, usually within a week after the cataloging work is completed. It is therefore my opinion that the Stanford copy of *Gibson* would have been publicly accessible no later than early April of 1999.

66. Because the Stanford MARC record (**Appendix 1008C**) is based on the Library of Congress MARC record (**Appendix 1008B**), these two records have the same essential data elements for this book, including the Library of Congress control number in Field 010, the ISBN for the paper version in Field 020, the Library of Congress Classification number in Field 050, the Dewey Decimal Classification number in Field 082, title in Field 245, edition information in Field 250, physical description in Field 300, bibliographical reference and index note in Field 504, the Library of Congress subject heading in Field 650, and the editor in Field 700. The differences are that the Stanford MARC record does not include IEEE Press in Field 260, includes a series statement in Fields 490 and 830, has a Field 530 to alert users that a PDF version of this book is "available on the Internet to registered users" and

Ford Motor Co.
Exhibit 1030
Page 035

shows a detailed content note and a publisher summary.

67. The Stanford MARC record (**Appendix 1008C**) shows *Gibson* was published in 1999 and cataloged by this library on March 25, 1999, making it discoverable in their online catalog from that date on. Interested users would have been able to discover this work by the Library of Congress control number, ISBN, title, series title, editor, Library of Congress classification number, and the Library of Congress subject heading.

E. Copyright Record

68. **Appendix 1008D** is a true and correct copy of the copyright registration record for *Gibson*. I obtained this record from the public catalog of the U.S. Copyright Office by first accessing their website at http://www.copyright.gov/, selecting the "Learn More" link under "Search Copyright Records," and following the link for "Official Public Catalog (1978 to present)" under "Online Records Collections" to perform a title search to retrieve the record. I personally located, identified and obtained this record. This is the type of record experts in my field would reasonably rely upon when forming their opinions.

69. The "Type of Work" field identifies the publication as "Text." The title, edition, imprint, series, and copyright claimant information match the information presented in **Ex. 1008** and the copy identified in **Appendix 1008A**. The "Date of Creation" field shows this work was created in 1998, and the "Date of Publication"

Ford Motor Co.
Exhibit 1030
Page 036

field shows "1999-02-23" (*i.e.*, February 23, 1999). The "Registration Number / Date" field shows "TX0004939558 / 1999-03-08" indicating *Gibson* was registered with the Copyright Office on March 8, 1999. This record informs my opinion that *Gibson* was published on February 23, 1999, and registered with the U.S. Copyright Office on March 8, 1999.

F. Usage Record

70. Actual usage of a publication is reflected by the papers that make reference to it. The citation history on Google Scholar shows that *Gibson* has been cited many times. **Appendix 1008E** is a list of citations from 2001 to demonstrate early usage. The earliest citation was published in April 2001.

G. Summary of My Opinion on *Gibson*

71. The digitized copy of *Gibson* on Internet Archive (identified by **Appendix 1008A**) shows that **Ex. 1008** is authentic. The copyright registration record (**Appendix 1008D**) shows *Gibson* was published on February 23, 1999. Library of Congress MARC record (**Appendix 1008B**) shows the library created a MARC record on September 24, 1998, before the book was published, and this record made *Gibson* discoverable in the online catalog of the Library of Congress from that date on. After the book was published in 1999, Library of Congress completed the cataloging work on March 12, 1999. Their copy of *Gibson* would have become publicly accessible as early as April 1999 and no later than June 1999.

Ford Motor Co.
Exhibit 1030
Page 037

The Stanford MARC record (**Appendix 1008C**) shows their copy cataloging record was based on the Library of Congress MARC record, and the cataloging work was completed on March 25, 1999, making the record searchable in their online catalog from that date on. The Stanford copy of *Gibson* would have become publicly accessible by April 1, 1999. All the records and copies I have reviewed inform my opinion that *Gibson* was published in 1999, and cataloging records were completed in March 1999, making this work discoverable in library catalogs no later than March 1999. The physical copies would have become publicly accessible no later than April of 1999. Citation history (**Appendix 1008E**) shows the earliest citation to *Gibson* appeared in April 2001. It is therefore my opinion that *Gibson* was publicly available well before the January 2004 critical date.

IV. Conclusion

72. In signing this declaration, I recognize that the declaration will be filed as evidence in a Petition for *Inter Partes* Review before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I also recognize that I may be subject to deposition in the case. If deposition is required of me, I will appear at a reasonable time and place to be agreed upon.

73. I hereby declare that all statements made herein on my own knowledge are true and that all statements made on information and belief are believed to be true, and further, that these statements were made with the knowledge that willful

Ford Motor Co.
Exhibit 1030
Page 038

false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Date: 3-27-2023 Executed: Ingrid Hsieh-Yee

Ingrid Hsieh-Yee, Ph.D.

Ford Motor Co.
Exhibit 1030
Page 039

Appendix A

Ingrid Hsieh-Yee

Professor Emerita
Dept. of Library and Information Science
Catholic University of America
Washington, D.C. 20064
E-mail: hsiehyee@cua.edu

Education

Ph.D. 1990. Library and Information Studies, University of Wisconsin-Madison
Minors: Sociology and Psychology

M.A. 1984. Library and Information Studies, University of Wisconsin-Madison.

M.A. 1983. Comparative Literature, University of Wisconsin-Madison.

B.A. 1980. Foreign Languages and Literature, National Taiwan University.

Work Experience

Professor, School/Dept. of Library and Information Science, Catholic University of America, 1990-2022 (Assistant Professor, 1990-1996; Associate Professor, 1997-2004; Professor 2004-2022; Professor Emerita 2022)

Co-Chair, Dept. of Library and Information Science, Catholic University of America, June 2015-August 2016.

Acting Dean, School of Library and Information Science, Catholic University of America, January 2010-June 2012.

Cataloger, Dept. of Legislative Reference Library, Annapolis, Maryland, 1989-1990.

Lecturer, School of Library and Information Studies, University of Wisconsin-Madison, 1988.

Teaching Assistant, School of Library and Information Studies, University of Wisconsin-Madison, 1986-1988.

Cataloger, Health Sciences Library, University of Wisconsin-Madison, 1984-1986.

Areas of Teaching and Research Interests

Information Organization and Access; Metadata; Cataloging & Classification; Information Architecture; Information Retrieval; Digital Collections; Social Media; Scholarly

Communication; Information Behavior; Health Informatics; Human Computer Interaction; Usability Studies

Grants & Honors

Cultural Heritage Information Management Project. IMLS grant. Amount: $498,741. Period: Aug. 2012 to July 2015. Co-PI with Dr. Youngok Choi.

D.C. Health Information Technology (HIT4): Building Capacity & Providing Access in Our Nation's Capital. Dept. of Labor H2B Training Grant. Grant amount: $4,175,500. Grant period: Nov. 2011 to Dec. 2015. Partner with the Metropolitan School of Professional Studies of the Catholic University of America, Children's National Medical Center, D.C. Department of Employment Services, Holy Cross Hospital, Howard University, Center for Urban Progress, Providence Hospital, and Sibley Memorial Hospital.

Capital Health Careers Project. Department of Labor Healthcare Sector and Other High Growth and Emerging Industries Grant. Grant amount: $4,953,999. Grant period: March 2010 – February 2013. Awarded to a group of healthcare organizations and educational institutions in Washington, D.C. Providence Health Foundation of Providence Hospital (Lead institution). Part of the grant supported the development of a Master's degree program in Information Technology with a concentration in Health Information Technology offered by the School of Library and Information Science.

The Washington D.C. School Librarians Project. IMLS grant. Grant amount: $412,660. Grant period: Aug. 2007 – June 2011. The School partnered with the District of Columbia Public Schools (DCPS) and the District of Columbia Library Association to educate and mentor school media specialists for the DCPS system. PI, Jan. 2010 to June 2011.

SIG Member of the Year, American Society for Information Science and Technology (2009).

Most Outstanding Paper of *OCLC Systems & Services* (2001).

ALISE Research Grant (2001).

Most Outstanding Paper of *OCLC Systems & Services* (2000).

Research Grant from ERIC (1999-2000).

Best Research Paper Award; Association for Library and Information Science Education (1998).

Research Grants, Catholic University of America. 1991, 1992, 1993, 1996, 1998, 1999, 2004, 2005, 2006, 2007, 2013-14.

Cooperative Faculty Research Grant, Consortium of Universities in the Washington Metropolitan Area (1993-1994).

Cooperative Research Grant, Council on Library Resources (1993-1994).

Journal of the American Society for Information Science Best Paper Award (1993).

ASIS/ISI Information Science Doctoral Dissertation Scholarship (1989).

HEA Title IIB Fellowship (Dept. of Education) (1989)

Chinese-American Librarians Association Scholarship (1987).

Beta Phi Mu (1985).

Vilas Fellowship, University of Wisconsin-Madison. 1984

Publications

Hsieh-Yee, I. (2021). Can We Trust Social Media? *Internet Reference Services Quarterly*, 1-15. Published online: 28 July, 2021. 10.1080/10875301.2021.1947433.

Bailey, T., & Hsieh-Yee, I. (2020). Combating the Sharing of False Information: History, Framework & Literacy Strategies, *Internet Reference Services Quarterly*, 24 (1-2), 9-30.

Choi, Y., and Hsieh-Yee, I. (2010). Finding Images in an OPAC: Analysis of User Queries, Subject Headings, and Description Notes. *Canadian Journal of Information and Library Science,* 34(3): 271 – 295.

Hsieh-Yee, I. (2008). Educating Cataloging Professionals in a Changing Information Environment. *Journal of Education for Library and Information Science,* 46(2): 93-106.

Vellucci, S. L., Hsieh-Yee, I., and Moen, W.E. (2007). The Metadata Education and Research Information Commons (MERIC): A Collaborative Teaching and Research Initiative. *Education for Information,* 25(3&4): 169-178.

NISO Framework for Guidelines for Building Good Digital Collections. 3rd ed. Baltimore, MD: National Information Standards Organization, 2007. Also available online: http://www.niso.org/framework/framework3.pdf (NISO Working Group members: Priscilla Caplan (chair), Grace Agnew, Murtha Baca, Tony Gill, Carl Fleischhauer, Ingrid Hsieh-Yee, Jill Koelling, and Christie Stephenson.)

Choi, Y., Hsieh-Yee, I., and Kules, B. (2007). Retrieval Effectiveness of TOC and LCSH. *Proceedings of the Joint Conference on Digital Libraries*, pp. 233-234.

Vellucci, S., and Hsieh-Yee, I. (2007). They Didn't Teach Me That in Library School! Building

a Digital Teaching Commons to Enhance Metadata Teaching, Learning and Research. *Proceedings of the National Conference of the Association of College and Research Libraries, Baltimore, MD*, pp. 26-31.

Mitchell, Vanessa, and Ingrid Hsieh-Yee. (2007). Converting Ulrich's Subject Headings to FAST Headings: A Feasibility Study. *Cataloging & Classification Quarterly,* 45(1): 59-85.

Hsieh-Yee, I., Tang, R., and Zhang, S. (2007). User Perceptions of a Federated Search System. *IEEE Technical Committee on Digital Libraries Bulletin,* Summer 3(2) (URL = http://www.ieee-tcdl.org).

Tang, R., Hsieh-Yee, I., and Zhang, S. (2007). User Perceptions of MetaLib Combined Search: An Investigation of How Users Make Sense of Federated Searching." *Internet Reference Services Quarterly,* 12(12): 211-236.

Hsieh-Yee, I., Tang, R., and Zhang, S. (2006). User Perceptions of a Federated Search System. *Proceedings of the Joint Conference on Digital Libraries*, *June 11-15, 2006, Chapel Hill*, p. 338.

Hsieh-Yee, I. (2006). *Organizing Audiovisual and Electronic Resources for Access: A Cataloging Guide*. 2nd ed. Westport, Conn.: Libraries Unlimited.

NISO A Framework of Guidance for Building Good Digital Collections. 2nd ed. Bethesda, MD: National Information Standards Organization, 2004. Framework Advisory Group: Grace Agnew, Liz Bishoff, Priscilla Caplan (Chair), Rebecca Gunther and Ingrid Hsieh-Yee.

Hsieh-Yee, I. (2004). Cataloging and Metadata Education in North American LIS Programs. *Library Resources & Technical Services,* 48(1): 59-68.

Hsieh-Yee, I. (2004). Cataloging and Metadata Education. In Gary E. Gorman (Ed.), *International Yearbook of Library and Information Management 2003: Metadata Applications and Management*, (pp.204-234). London: Facet Publishing.

Yee, P. L., Hsieh-Yee, I., Pierce, G.R., Grome, R., and Schantz, L. (2004). Self-Evaluative Intrusive Thoughts Impede Successful Searching on the Internet. *Computers in Human Behavior,* 20(1): 85-101.

Hsieh-Yee, I. (2003). Cataloging and Metadata Education: A Proposal for Preparing Cataloging Professionals of the 21st Century. A report submitted to the ALCTS-Education Task Force in response to Action Item 5.1 of the *Bibliographic control of Web Resources: A Library of Congress Action Plan*. Approved by the Association for Library Collections and Technical Services. Web version available since April 2003 at http://lcweb.loc.gov/catdir/bibcontrol/CatalogingandMetadataEducation.pdf.

Hsieh-Yee, I. (2002). Cataloging and Metadata Education: Asserting a Central Role in Information Organization. *Cataloging & Classification Quarterly* 34(½): 203-222.

Hsieh-Yee, I., and Smith, M. (2001). The CORC Experience: Survey of Founding Libraries, Part I. *OCLC Systems & Services,* 17: 133-140. (Received "The Most Outstanding Paper of OCLC Systems & Services in 2001" award.)

Hsieh-Yee, I., and Smith, M. (2001). The CORC Experience: Survey of Founding Libraries, Part II, Automated Tools and Usage. *OCLC Systems & Services,* 17: 166-177. (Received "The Most Outstanding Paper of OCLC Systems & Services in 2001" award.)

Hsieh-Yee, I. (2001). ERIC User Services: Changes and Evaluation for the Future. *Government Information Quarterly,* 18: 31-42.

Hsieh-Yee, Ingrid. (2001). Research on Web Search Behavior. *Library and Information Science Research,* 23: 167-185.

Logan, E., and Hsieh-Yee, I. (2001). Library and Information Science Education in the Nineties. *Annual Review of Information Science and Technology,* 35: 425-477.

Hsieh-Yee, I. (Ed.) (2001). *Library and Information Science Research*, 23 (2). A special issue in honor of the retirement of Douglas L. Zweizig.

Hsieh-Yee, I. (2000). *ERIC User Services: Evaluation in a Decentralized Environment*. Washington, D.C.: Dept. of Education.

Hsieh-Yee, Ingrid. (2000). *Organizing Audiovisual and Electronic Resources for Access: A Cataloging Guide*. Littleton, CO: Libraries Unlimited.

Hsieh-Yee, I. (2000). Organizing Internet Resources: Teaching Cataloging Standards and Beyond. *OCLC Systems & Services,* 16: 130-143. (Received "The Most Outstanding Paper of OCLC Systems & Services in 2000" award.)

Hsieh-Yee, I. (1998). The Retrieval Power of Selected Search Engines: How Well Do They Address General Reference Questions and Subject Questions? *Reference Librarian,* 60: 27-47.

Hsieh-Yee, I. (1998). Search Tactics of Web Users in Searching for Texts, Graphics, Known Items and Subjects: A Search Simulation Study. *Reference Librarian,* 60: 61-85. (Received the 1997 Best ALISE Research Paper Award.)

Hsieh-Yee, I. (1997). Access to OCLC and Internet Resources: LIS Educators' Views and Teaching Practices. *RQ,* 36: 569-86.

Hsieh-Yee, I. (1997). Teaching Online and CD-ROM Resources: LIS Educators' Views and Practices. *Journal of Education for Library and Information Science,* 38: 14-34.

Hsieh-Yee, I. (1996). The Cataloging Practices of Special Libraries and Their Relationship with OCLC. *Special Libraries,* 87: 10-20.

Hsieh-Yee, I. (1996). Modifying Cataloging Practice and OCLC Infrastructure for Effective Organization of Internet Resources. In *Proceedings of the OCLC Internet Cataloging Colloquium*. [Online]. Available: http://www.oclc.org/oclc/man/colloq/hsieh.htm

Hsieh-Yee, I. (1996). Student Use of Online Catalogs and Other Information Channels. *College & Research Libraries,* 57: 161-175.

Hsieh-Yee, I. (1995). Ten entries in James S. C. Hu (Ed.), *Encyclopedia of Library & Information Science*, 913, 1028-29, 1036, 1037, 1145-46, 1514, 1575, 1763-64, 2216-27, 2378-79. Taipei, Taiwan: Sino-American Publishing. (Topics include "Advanced Technology/Libraries," "Information Ethics," "Instruction on Cataloging and Classification," "Instruction on Reference Services.")

Hsieh-Yee, I. (1993). Effects of Search Experience and Subject Knowledge on Online Search Behavior: Measuring the Search Tactics of Novice and Experienced Searchers. *Journal of the American Society for Information Science,* 44: 161-174. (Received the 1993 Best JASIS Paper Award.)

Works in Progress

Hsieh-Yee, I. Actionable lessons for debunking false information: A comparative analysis of literacy lessons in communication and library and information science

Hsieh-Yee, I. Representing diverse voices and communities in cataloging and indexing: The evolution of controlled vocabularies

Hsieh-Yee, I. Information Organization and Retrieval in Practice (a monograph scheduled for completion and publication in 2023)

Presentations

Bailey, T. & Hsieh-Yee, I. (February 2020). The Phenomena of Sharing Misinformation and the Need for Information Literacy. Presented at the 2020 Bridging the Spectrum Symposium, Washington, D.C., February 14, 2020.

Hsieh-Yee, I. and Fragan-Fly, J. (May 2018) Trends, Design & Strategies for Digital Scholarship Services. Presented at the 2018 Maryland/Delaware Library Association Conference, Cambridge, MD.

Hsieh-Yee, I. (February, 2018) Research Data Management: What It Takes to Succeed. Presented at the 10th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. (February, 2017) *Research Data Management: New Competencies and Opportunities for Information Professionals*. Presented at the 9th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. and Lawton, P. (February, 2017) *Enhancing Catholic Portal Searches with User Terms and LCSH*. Presented at the 9th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. (2016, October) *Visualizing Data for Information*. Presented at the 2016 Virginia Library Association Conference, Hot Springs, VA.

Hsieh-Yee, I. (2016, August) *Religious Materials Toolbox for Archivists: Solutions to Problems Facing the Profession*. Presented at Archives * Records 2016, Atlanta, GA.

Hsieh-Yee, I. and Lawton, P. (2016, March) *Enhancing Retrieval of Catholic Materials with LCSH Knowledge Structure*. Presented at the 2016 Catholic Library Association Conference, San Diego, CA.

Fagan-Fry, J. and Hsieh-Yee, I. (2016, February) *Approaches to Digital Scholarship at Top Universities around the World: Scholarly Publishing in the Digital Age.* Presented at the 8th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee and Fagan-Fry, J. (2016, January) *Innovative Services for Digital Scholarship at Top 100 Research Libraries of the World*. Poster presented at the 2016 Annual Conference of the Association for Library and Information Science Education, Boston, Mass.

Hsieh-Yee, I. and Lawton, P. (2015, June). *Crowdsourcing terms for CRRA portal themes.* Poster presented at the third CRRA symposium and annual meeting, Bringing the created toward the Creator: Liturgical art and design since Vatican II. Catholic Theological Union, Chicago, Illinois.

Hsieh-Yee, I. and Lawton, P. (2015, February). *Crowdsourcing terms for thematic exploration in the Catholic Portal.* Poster presented at the 7th Annual Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I., James, R., and Fagan-Fry, J. (2015, February). *Support for digital scholarship at top university libraries of the world*. Poster presented at the 7th Annual Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I., Zhang, S., Lin, K., and Cherry, S. (2015, February). *Thus said the end users: Summon experience and support for research workflows*. Poster presented at the 7th Annual Bridging the Spectrum Symposium, Washington, D.C.

Yontz, E., Hsieh-Yee, I., & Houston, S. (2015, February). *Healthy Heroes Summer Reading Club: Developing healthy youth at public libraries.* 11th Annual Jean Mills Health Symposium, Greenville, North Carolina.

Yontz, E., Hsieh-Yee, I., and Houston, S. (2015, January). *Healthy youth and libraries: A pilot study.* Association for Library & Information Science Education (ALISE) Annual Conference, Chicago, Illinois.

Hsieh-Yee, I. (2014, May). *Linking CRRA resources to portal themes via authority files*. Presented at the Catholic Research Resources Alliance 2014 Membership Meeting, Marquette, WI.

Hsieh-Yee, I. (2014, April). *Enhancing subject access to CRRA resources*. Presented at the 2014 Catholic Library Association Conference, Pittsburgh, PA.

Hsieh-Yee, I. (2014, January). *Health Information Technology Program: Educational entrepreneurship in action*. Presented at the 2014 annual Conference of the Association for Library and Information Science Education, Philadelphia, PA

Hsieh-Yee, I., Zhang, S., Lin, K., and Cherry, S. (2014, January). *Discovering information through Summon: An analysis of user search strategies and search success*. Paper presented at the 6th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. (2012, December). *National Digital Stewardship Alliance and SLIS at CUA: An Educational Partnership*. Paper presented at Best Practices Exchange: Acquiring, Preserving, and Providing Access to Government Information in the Digital Era, Annapolis, MD

Choi, Y. and Hsieh-Yee, I. (2010, January). *Finding Images in an OPAC: Analysis of User Queries, Subject Headings, and Description Notes*. Paper presented at 2nd Annual Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. and Coogan, J. (2010, January). *Google Scholar vs. Academic Search Premier: What Libraries and Searchers Need to Know*. Paper presented at 2nd Annual Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (2009, November). *Information Science Education: An LIS School's Perspective*. Paper presented at Annual Meeting of the American Society for Information Science and Technology, Vancouver, British Columbia, Canada.

Hsieh-Yee, I., Menard, E., Ya-Ning Chen, A., Shu-Jiun Chen, S., Kalfatovic, M. R., Wisser. K. M. (2009, November). *Information Organization in Libraries, Archives and Museums: Converging Practices and Collaboration Opportunities*. Presented at Annual Meeting of the American Society for Information Science and Technology, Vancouver, British Columbia, Canada. (Organizer and moderator of this panel.)

Hsieh-Yee, I. and Coogan, J. (2009, July). *Catching up to Google Scholar: The Retrieval Power of Academic Search Premier and Google Scholar*. Poster presented at American Library Association Conference, Chicago, Illinois.

Hsieh-Yee, I., with the CUA Scholarly Communications Project Team. (2009, January). *Digital Scholarship@CUA: Developing an Institutional Repository for CUA*. Poster presented at 1st Annual Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Wise, M., Cylke, K., and Hsieh-Yee, I. **(**2009, January). *Digital Talking Books: Meeting the Needs of the Blind and the Handicapped*. Paper presented at the Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (2009, January). *User Expectations of MERIC*. Presented at the Information Organization Competencies for the 21st Century Discussion Session of the 2009 Conference of the Association for Library and Information Science Education, Denver, Colorado.

Choi, Y., and Hsieh-Yee, I. (2008,November). *Subject Access for Images in an OPAC.* Annual Meeting of the American Society for Information Science and Technology, Columbus, Ohio. (Also co-organized a panel on Retrieving and Using Visual Resources: Challenges and Opportunities for Research and Education.)

Hsieh-Yee, I. (2008, June). *Educating Cataloging Professionals in a Changing Information Environment*. National Taiwan University, Taipei, Taiwan.

Vellucci, S. L., Moen, W.E., Hsieh-Yee, I., Marson, B., and Wisser, K. (2008, January) *Building a Metadata Education and Research Community through MERIC (Metadata Education and Research Information Commons): Demo and Stakeholder Input*. A panel presented at the 2008 Conference of the Association for Library and Information Science Education, Philadelphia, Pennsylvania.

Hsieh-Yee, I., Choi,Y. and Kules, B. (2007, October). *Searching for Books and Images in OPAC: Effects of LCSH, TOC and Subject Domains.* A poster presented at the American Society for Information Science and Technology Annual Meeting, Milwaukee, Wisconsin.

Hsieh-Yee, I. and Coogan, J. (2007, August) *Google Scholar vs. Academic Search Premier: A Comparative Analysis.* Presented to the Faculty and Staff of the University of the District

of Columbia.

Hsieh-Yee, I. and Coogan, J. (2007, June). *Google Scholar vs. Academic Search Premier: A Comparative Analysis*. Presented to the Washington Research Library Consortium Community, Catholic University of America, Washington, D.C.

Hsieh-Yee, I., Choi, Y., and Kules, B.. (2007, June). *What Users Need for Subject Access: Table of Contents or Subject Headings?* A poster presented at the 2007 American Library Association Annual Conference, Washington, D.C., June 2007.

Choi, Y., Hsieh-Yee, I., and Kules, B. (2007, June). *Retrieval Effectiveness of TOC and LCSH.* A paper presented at the Joint Conference on Digital Libraries 2007, Vancouver, Canada.

Vellucci, S. L., Hsieh-Yee, I., and Moen, W.E. (2007, May). *If We Build It, Will They Come? Building a Community of Practice for Metadata Stakeholders.* A poster presented at the Rutgers University Research Day, Bridgeton, New Jersey.

Hsieh-Yee, I. (2007, May). *Federated Searching: User Experience & Perceptions.* International Conference on Information Organization & Retrieval, National Taiwan University, Taipei, Taiwan.

Hsieh-Yee, I. (2007, May). *Search Performance of Google Scholar and Academic Search Premier.* International Conference on Information Organization & Retrieval, National Taiwan University, Taipei, Taiwan.

Hsieh-Yee, I. (2007, May) *MERIC: Building a Digital Commons for Metadata Education & Research.* International Conference on Information Organization & Retrieval, National Taiwan University, Taipei, Taiwan.

Hsieh-Yee, I., and Coogan, J. (2007, March/April). *A Comparative Analysis of Google Scholar and Academic Search Premier.* Poster presented at the Association of College & Research Libraries 13th National Conference, Baltimore, Maryland.

Vellucci, S. L. and Hsieh-Yee, I. (2007, March/April) *They Didn't Teach Me That in Library School! Building a Digital Teaching Commons to Enhance Metadata Teaching, Learning and Research*. On-site presentation and Webcast by Elluminate. A contributed paper presented at the Association of College & Research Libraries 13th National Conference, Baltimore, Maryland. The acceptance rate for contributed paper was 20%. This paper was one of 10 conference papers chosen for live webcast during the conference.

Moen, W., Hsieh-Yee, I. and Vellucci, S.L. (2007, January) *A DSpace Foundation for a Teaching & Research Commons: The Metadata Education and Research Information Commons.* A poster session presented at the Open Repositories Conference 2007, San Antonio, Texas.

Tang, R., Hsieh-Yee, I., and Zhang, S. (2006, November) *User Perception of MetaLib Combined Search*. Paper presented at the Annual Meeting of the American Society for Information Science and Technology, Austin, Texas, Nov. 2006.

Hsieh-Yee, I. (2006, November). *Federated Searching: User Perceptions, System Design, and Library Instructions.* Paper presented at the Annual Meeting of the American Society for Information Science and Technology, Austin, Texas. (Panel organizer, moderator, presenter).

Hsieh-Yee, I. (2006, November). *Building a Digital Teaching Commons to Enhance Teaching and Learning: The MERIC Experience and Challenges.* Paper presented at the Annual Meeting of the American Society for Information Science and Technology, Austin, Texas. (Panel organizer, moderator, presenter)

Hsieh-Yee, I. (2006, September). *Search Performance of Google Scholar and Academic Search Premier.* Paper presented at the ERIC Publishers Meeting, Washington, D.C.

Hsieh-Yee, I., Zhang, S., and Rong Tang, R. (2006, June). *User Perceptions of a Federated Search System*. Poster presented at Joint Conference on Digital Libraries, Chapel Hill, North Carolina.

Hsieh-Yee, I. and Zhang, S. (2006, June). *Preparing Users for Federated Search: Implications of a MetaLib User Perceptions Study.* Paper presented at the 2006 Ex Libris User Groups of North America Conference, Knoxville, Tennessee.

Hsieh-Yee, I. (2006, January). *MERIC Organizations and Navigation*. Paper presented at the 2006 ALISE Annual Conference, San Antonio, Texas.

Hsieh-Yee, I. (2006, January). *Metadata and Cataloging Education: Recommended Competencies.* Paper presented at the 2006 ALISE Annual Conference, San Antonio, Texas.

Hsieh-Yee, I. (2005, November). *Digital Library Evaluation: Progress & Next Steps.* Presentation at the Annual Meeting of the American Society for Information Science & Technology, Charlotte, North Carolina.

Hsieh-Yee, I. (2005, August). *Providing Access to Digital Content: Issues for DL Managers.* Presentation at MDK12 Digital Library Steering Committee Meeting, Columbia, Maryland.

Hsieh-Yee, I. (2005, April). *Enhancing Teaching and Learning: The Role of School Library Media Specialists.* Presentation at Meeting of the Baltimore County Public School System School Media Specialists, Baltimore, Maryland.

Hsieh-Yee, I. (2005, January). *Subject Access and Users: Insights & Inspirations from Marcia J. Bates.* Paper presented at the Historical Perspectives SIG, 2005 Conference of the Association for Library and Information Science Education, Boston, Massachusetts.

Hsieh-Yee, I. (2005, January). *Electronic Resource Management: Practice, Employer Expectations, & CE Interests.* Paper presented at Technical Services Education SIG, 2005 Conference of the Association for Library and Information Science Education, Boston, Massachusetts.

Hsieh-Yee, I. (2004, October). *Library Professionals for the Digital Age: Competencies & Preparation.* Paper presented at Bibliographic Access Management Team meeting, Library of Congress, Washington, D.C.

Hsieh-Yee, I. (2004, January). *Cataloging and metadata expertise for the digital era.* Presented at Preparing 21st Century Cataloging and Metadata Professionals: A Workshop for Educators and Trainers, San Diego and sponsored by ALCTS, ALISE, LC, and OCLC.

Hsieh-Yee, I. (2004, January). *Educating catalogers for the digital era.* Paper presented at the Technical Services SIG, 2004 Conference of the Association for Library and Information Science Education, San Diego.

Hsieh-Yee, I. (2003, July). *Cataloging Education for the 21st Century.* A presentation at the Library of Congress, Washington, D.C.

Hsieh-Yee, I. (2002, January) *Metadata Education and Research Priorities: A Delphi Study of Metadata Experts.* Presentation at the 2002 Conference of the Association for Library and Information Science Education, New Orleans.

Hsieh-Yee, I. (2001, November). *A Delphi Study of Metadata: Preliminary Findings.* Poster session at the 2001 Annual Meeting of the American Society for Information Science & Technology, Washington, D.C.

Hsieh-Yee, I. (2001, June). *Resources on Asian American Children: Analysis of Retrieval by Search Engines and WorldCat.* Presentation at the National Conference on Asian Pacific American Librarians, San Francisco.

Hsieh-Yee, I. (2001, January). *Delphi Study on Metadata: Project Design.* Presentation at Research Awards Session, Association for Library & Information Science Education, Washington, D.C.

Hsieh-Yee, I. (2000, May). *Web Search Behavior Research: Progress and Implications.* Presentation at the Symposium on Evaluating Library and Information Science Research, University of Wisconsin-Madison, Madison, Wisconsin.

Hsieh-Yee, I. (2000, March). *ERIC User Services: Evaluation in a Decentralized Environment.* Presentation at the National ERIC Joint Directors/Technical Meeting, Arlington, Virginia.

Hsieh-Yee, I. (2000, January). *Enhancing Learning with Web Technology.* Presentation at Faculty Conversations, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (2000, January). *From Surrogates to Objects: CUA's Approaches to Organizing Electronic Resources.* Paper presentation at the Annual Conference of the Association for Library and Information Science Education, San Antonio, Texas.

Yee, P., and Hsieh-Yee, I. (1997, November). *Individual Differences in Search Behavior on the WWW.* A poster session presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia, Pennsylvania.

Hsieh-Yee, I. (1997, April). *Research + Marketing + Preparation = Job!* Presented at the "Workshop on Resume and Interview Techniques," Special Libraries Association, Student Chapter, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (1997, February). *Creating CyberCatalogers: Education and Training.* Presentation at ALA's Midwinter Meeting, Washington, D.C.

Hsieh-Yee, I. (1997, February). *Search Tactics of Web Users in Searching for Texts, Graphics, Known Items and Subjects: A Search Simulation Study.* Presented at the Conference of the Association for Library and Information Science Education, Washington, D.C.

Hsieh-Yee, Ingrid. "Beginning Your Special Library/Information Center Career." Presented at SLA's "Career Day," Jan. 11, 1997, Catholic University of America.

Hsieh-Yee, I. (1996, September). *The Roles of Library and Information Scientists in Managing Electronic Information.* Presentation at Hamilton College, Clinton, New York.

Hsieh-Yee, I. (1996, May). *The Future of Cataloging as a Profession.* Presented at "The Cataloging Forum, Library of Congress, Washington, D.C.

Hsieh-Yee, I. (1994, October). *The Impact of the Internet on OPACs.* Presented at the Third Workshop on User Interfaces for OPACs, Library of Congress, Washington, D.C.

Reports

Hsieh-Yee, I., with Knowledge Management Competencies and Performance Action Group of the Federal Knowledge Management Initiative. "From Knowledge Management Competencies to Improved Organizational Performance." April 9, 2009.

Hsieh-Yee, I., with Knowledge Practices Action Group of the Federal Knowledge Management Initiative. "KM Practice in Government Agencies: Findings and Recommendations." April 9, 2009.

Hsieh-Yee, I. "Delphi Study on Metadata." 2001. Three quarterly reports submitted to the Association for Library and Information Science Education.

Hsieh-Yee, I. "College Students' Information Channels: Patterns of Use and Possible Factors in Channel Selection." 1995. Submitted to the Catholic University of America.

Hsieh-Yee, I. "The Information-Seeking Patterns of Scholars and Their Use of an Online Information System." 1994. Submitted to the Council on Library Resources.

Book Reviews

Review of *The Measurement and Evaluation of Library Services*, by Sharon L. Baker and F. Wilfrid Lancaster. *Information Processing and Management* 30 (1994): 450-52.

Review of *Subject Access to Films and Videos, by Sheila S. Intner and William E. Studwell; and Cataloging Unpublished Nonprint Materials*, by Verna Urbanski with Bao Chu Chang and Bernard L. Karon. *Information Processing and Management* 30 (1994): 449-50.

Review of *Automated Information Retrieval in Libraries: A Management Handbook*, by Vicki Anders. *Journal of Library and Information Science* 19 (1993): 98-100.

Review of *Full Text Databases*, by Carol Tenopir and Jung Soon Ro. *Information Processing and Management* 28 (1992): 667-68.

Review of *Descriptive Cataloging for the AACR2R And USMARC: A How-to-Do It Workbook*, by Larry Millsap and Terry Ellen Ferl. *Information Processing and Management* 28 (1992): 809-11.

Review of *MARC Manual: Understanding and Using MARC Records*, by Deborah J. Byrne. *Information Processing and Management* 28 (1992): 537-38.

Service

Professional Associations and Societies

- Library of Congress. RDA Training Program for the Profession. Co-authored with Tim Carlton. 2013-2014.
- 2014 Digital Preservation Outreach & Education Survey. Contributed to the design of the survey, 2014.
- National Digital Stewardship Alliance. Outreach Committee. 2011-2014.
- National Digital Stewardship Residency Program. Advisory Group, 2012-2013.
- FEDLINK Health Information Technology Advisory Council, 2011-2015.

- 2012 Joint Conference on Digital Libraries. Program Planning Committee, Pre-Conference Proposals Review Committee, 2012
- Catholic Research Resources Alliance. Five-Year Strategic Plan Task Force, 2011-2012
- Institute of Museum and Library Services. Grant reviewer. 2004, 2005, 2010.

- Association for Library and Information Science Education.
  * ALISE Bodan Wynar Research Paper Award Committee, 2015, 2016, 2017
  * ALISE Eugene Garfield Dissertation Award Competition, Jury, 2013, 2014
  * ALISE Research Grant Competition Committee. Chair, 2012
  * Pratt-Severn Faculty Innovation Award. Chair, 2009, 2010
  * ALISE Doctoral Poster Jury, 2012
  * "Information Organization Competencies for the 21st Century" Discussion session leader. 2009 Conference of the Association for Library and Information Science Education.
  * Assisted Technical Services SIG Convener in organizing a program, ""Building a Metadata Education and Research Community through MERIC (Metadata Education and Research Information Commons): Demo and Stakeholder Input" for the 2008 ALISE conference.
  * Association for Library Collections and Technical Services/Association for Library and Information Science Education (ALCTS/ALISE) Metadata Education and Research Information Center (MERIC) Advisory Board, Co-Chair (with Sherry Vellucci), 2005-2007. Chair, 2008-2009 (leading the effort to build MERIC, a repository and collaborative space for metadata educators, practitioners, and researchers)
  * Technical Services SIG, Convener, 2004-2005. Organized a program on "Electronic Resources Management: Current Practices, Employer Expectations, and Teaching Strategies" for the 2005 conference in Boston, Massachusetts.
  * Technical Services SIG, Convener, 2003-2004. Organized a program on "Organizing Information with Metadata: Desired Competencies and Teaching Innovations" for the 2004 conference.
  * Technical Services SIG, Convener, 1999-2000. Organized a program on "Teaching the Organization of Electronic Resources" for the 2000 conference.
  * Curriculum SIG, Co-convener (with Sibyl Moses), 1996-97. Organized a program on "Government Information Policy" for the 1997 conference.

- American Society for Information Science & Technology.
  * Reviewer, Conference program panel submissions and poster submissions, 2005, 2006, 2007, 2009, 2011, 2012, 2013, 2014, 2015, 2016, 2017
  * Nomination Committee, 2009-2011
  * Information Science Education Special Interest Group. American Society for Information Science and Technology. Chair-Elect, 2007-2008. Chair 2008-2009.
  * Committee on Information Science Education. 1999-2006.
  * Committee on Information Science Education. Organizing Committee for an orientation program for students at ASIS annual meetings, 1999-2001
  * Committee on Information Science Education. Sub-committee on Student Welfare (focusing on issues related to master's education), 1998-2001

* SIG ED. Organizing Committee for the "Seminar on Research and Career Development" for junior researchers. 1995-96 (chair), 1997-2001
* ISI Doctoral Dissertation Proposal Scholarship Jury, 1997; 2001, 2002
* Pratt-Severn Best Student Research Paper Award Jury. Chair. 1997
* 1998 Midyear Meeting (referee of contributed papers), 1997
* Organizer and moderator of the ASIS Doctoral Forum and the Doctoral Research Seminar 1994-1995
* SIG Human Computer Interaction. Chair-Elect, Chair, 1993-1995
* Doctoral Forum Award Jury, 1995
* Best Student Paper Award Jury, 1995

- American Library Association.
  * Committee on Accreditation, External Review Panelist, 2009- (site visiting team 2013-2014; site visiting team 2016-2017)
  * Association for Library Collections and Technical Services Task Force on Competencies and Education for a Career in Cataloging, member, 2008-2009
  * Facilitator for "What They Don't Teach in Library School: Competencies, Education and Employer Expectations for a Career in Cataloging," an Association for Library Collections and Technical Services Preconference, June 22, 2007 in Washington, D.C. Also a local liaison for bringing this program to the Catholic University of America.
  * Facilitator for a discussion on "Effect of Electronic Resources on Technical Services" at ALA's Midwinter Meeting held in Feb. 1997 in Washington, D.C.
  * International Relations Committee, Subcommittee Task Force for IFLA and China, 1994-1997

- Virginia Association of School Librarians. Scholarships and Awards Committee. 2010-2012

- Federal Knowledge Management Initiative, Knowledge Management Practices Action Group. Member. 2009 (leading the effort to build a knowledge management repository)
- Federal Knowledge Management Initiative, Knowledge Management Competencies & Learning Action Group. Member. 2009 (developing an action plan for helping government knowledge workers and government agencies to develop knowledge management competencies)
- National Center for Education Statistics. Technical Review Panel. 2008.
- External evaluator for a case of promotion to full professorship. University of Tennessee. 2008.
- National Information Standards Organization (NISO). Advisory Board, Revision of "IMLS Framework of Guidance for Building Good Digital Collections," 2004, 2007.
- Library of Congress, Bibliographic Control of Web Resources: A Library of Congress Action Plan. Principal Investigator of Action Item 5.1, focusing on cataloging and metadata education for students and new librarians, 2002-2003. (worked with the Association for Library Collections and Technical Services, Education Task Force)

- Chinese American Librarians Association
  * Chinese American Librarians Association Outstanding Library Leadership Award in Memory of Dr. Margaret Chang Fung, Award Committee, 2016-2017
  * Achievement Award Jury, 2000-2001
  * CALA Goal 2000 Task Force, 1997
  * Scholarship Committee, 1995, 1996-1997 (chair)
  * Board of Directors, 1994-1997
  * Publication Committee, 1993-1995
  * International Relations Committee, 1993-1996

- Sailor[SM] Assessment Advisory Group (An impact study of Sailor, Maryland's Public Information Network), 1995

- Editorial boards
  Journal of Library & Information Science. Editorial Board, 2012-
  Chinese American Librarians Association, *Occasional Papers Series*. Editorial Board, 2009-2016.
  Library Quarterly. Editorial Board, 2003-2008
  Bulletin of the Medical Library Association, 1994-97
  Newsletter editor for the Chinese American Librarians Association, 1989-92

- Referee for the following journals
  Information Processing and Management
  Journal of Digital Information
  Journal of Education for Library and Information Science
  Journal of Information Science
  Journal of Library & Information Science
  Journal of Library Metadata
  Journal of the American Society for Information Science & Technology
  Library and Information Science Research
  Library Quarterly

- Expert reviewer, "Digital Library" course, Evaluation module, University of North Carolina, Chapel Hill, 2007-2008.
- Expert reviewer, "Information Organization" course, University of Michigan, Ann Arbor. 2007.

Catholic University of America

- Academic Senate Committee on Appointments and Promotions, 2005-2008
- Academic Senate Committee on Committees and Rules, 2009-2012, 2019-2022
- Academic Senate representative (for School of Arts & Sciences), 2017-2020

- Academic Senate representative (for School of Library & Information Science) 2003-2012
- Academic Senate Library Committee, Interim Chair (2007), Member, 2008-2012
- Academic Senate Committee on Computing, 1995-2003

- President's Administrative Council, 2010-2012
- Deans' Council, 2010-2012
- Academic Leadership Group, 2010-2012
- Graduate Board, 2010-2012

- School of Arts & Sciences, Committee on Appointments and Promotions, 2015-2019
- School of Arts & Sciences, Academic Council, 2015-2016.
- School of Arts & Sciences, Ordinary Professor Group, 2013-
- Doctoral Dissertation Defense Committee, Chair, Dept. of Psychology, 2016, 2017, 2018, 2019, 2020, 2021, 2022
- Doctoral Dissertation Defense Committee, Chair, Dept. of Education, 2014, 2015, 2017, 2018, 2019
- Doctoral Dissertation Defense Committee, Chair, National Catholic School of Social Services, 2019, 2020, 2021
- Doctoral Dissertation Defense Committee, Chair, School of Nursing, 2006, 2008
- CUA Scholarly Communication Project Team, Member (2007), Chair, 2008-2009
- Dean Search Committee, 1992-1994, 1998-1999, 2002-2003, 2006-2007
- Fulbright Review Panel, 2006
- CUA Service Learning Advisory Board, 2001-2002
- CUA Faculty Conversations on Enhancing Teaching and Learning through Technology, Planning Group, 1999-2001
- CUA Initiative on Technology and Teaching, 1998-2001

Dept. of Library and Information Science

- Admissions Committee, 2007-2009, Chair 2010-2012, Member 2013-2015, Member 2018-2022
- Appointments and Promotions Committee, 1991-2022
- Student advisement, 1990-2022
- 4+1 Bachelor to Master program, Program Advisor, 2014-2022
- LIS Minor, Advisor, 2017-2022
- Curriculum Committee, 2020-21 (Member), 2021-2022 (Chair)
- Scholarship & Awards Committee, 2016-2018, 2020-21
- Comprehensive examination editor, 2020-21
- Symposium, Colloquium, Lecture Series Committee, May 2018-May 2020 (Chair)
- Symposium and Colloquium Committee, fall 2016-May 2018
- Bridging the Spectrum Symposium session moderator, 2018, 2019, 2020, 2021

- Organization of Information Core Course Outcomes Assessment, 2019, 2020
- The Information Professions in Society Core Course Outcomes Assessment, 2019, 2020
- Community Services Librarianship, Course of Study Review and Revision, 2020
- Organization of Information, Course of Study Review and Revision, 2019
- Organization of Information, Competency and Course Offerings Review, 2018-2019
- Accreditation presentation, Chair, June 2015-August 2016
- Interim Co-Chair, June 2015-August 2016.
- Blended/OWL Learning Committee, spring 2016-2018
- Technology Committee, fall 2016-2017
- Comprehensive examination editor, 2016-2017, reader (every year since 1990)
- LIS Advisory Board, 2015-2016 (chair); fall 2016- May 2018 (member)
- Committee on Planning and Assessment, 2015-2016 (chair)
- Senior Faculty Committee, 2014-2016.
- Accreditation Steering Committee, 2014-2016 (Chair, 2015-2016)
- Accreditation Students Standard Committee, co-chair, 2014-2016
- Accreditation Mission, Goals, and Objectives Standard Committee, co-chair, 2014-2016
- Accreditation Curriculum Standard, member 2014-2-16
- Accreditation Administration and Finance Standard, member 2014-2016
- Cultural Heritage Information Management Project (IMLS-funded), Co-PI, 2012-2015
- Cultural Heritage Information Management Forum (scheduled for June 2015), Co-Organizer, 2013-2015
- Health Information Technology Interim Review Committee, 2015 (chair)
- Health Sciences Librarianship Advisory Group, 2015- (chair)
- Comprehensive examination editors, 2013-2014, 2016-2017
- National Digital Stewardship Alliance liaison, 2011-2014
- Advisory Board, Chair 2010-2012
- Academic Honesty Committee, Chair, 2008-2012
- Blended Learning Committee, 2010-2012
- Colloquium Committee, 2010-2012
- Comprehensive Examination Administration, 2010-2012
- Cultural Heritage Information Management Advisory Committee, 2010-2012 (chair), 2013-
- Curriculum Committee, 1991-2003, 2007-2009, Chair 2010-2012, member 2013-
- Curriculum Subcommittee on Comprehensive Examination, Chair 2009-2012
- Health Information Technology Advisory Board, Chair 2010-2012. Member 2013-
- Health Sciences Advisory Committee, 2009, Chair 2010-2012. Member 2013-
- HIT Expert Forum, Chair 2012. Member 2013-2015
- Health Information Technology Student Group Advisor, 2011-2012
- State Council for Higher Education of Virginia, SLIS Representative, 2010-2012
- Symposium Planning Committee, 2010-2012
- Website Management Team, Chair, 2010-2012
- Urban School Librarianship Project (IMLS-Funded), PI, 2007-2011 (chair, 2010-11)

- Failing Grades Committee, 1995-1997 (chair), 2000-2001 (chair), 2004-2005 (chair), 2007 (chair)-2011
- Faculty Search Committee, 1994-1998, 2002-2004, 2006 (chair), Fall 2007-2009, Chair fall 2009-2012
- Recruitment Committee, Chair 2010-2012
- Strategic Planning Committee, Chair 2010-2012
- Technology Committee, 2010-2012
- Accreditation Advisory Committee, 2007-2009
- Accreditation Coordinating Committee, 2007-2009
- Accreditation Steering Committee, 2007-2009
- SLIS Advisory Group, 2007-2009
- Accreditation Curriculum Standard Committee, Co-chair, 2007-2009
- Accreditation Faculty Standard Committee, Co-chair, 2007-2009
- LSC 551 Information Organization Review Team, Co-chair, 2008-2009, 2015-2016.
- Curriculum Subcommittee on Portfolios, 2009
- LSC 555 Information Systems in Libraries and Information Centers Review Team, contributor, 2008-2009
- Redesign of LSC 730 Use and Users of Libraries and Information. 2009-
- Development of a metadata institute that was taught as LSC 715 Organization of Internet Resources in 2008. The institute is being revised and will be offered in 2010 under a new course title.
- Development of lesson plans, assignments, and evaluation rubrics for LSC 606, Cataloging and Classification, for the School's NCATE accreditation. 2008
- Howard and Mathilde Rovelstad Scholarship Committee, Chair, 2004-2007
- Assistant Dean Search Committee, Chair, Fall 2007
- Liaison to the Association for Library Collections and Technical Services to bring its preconference program, Cataloging Education and Employer Expectations, to CUA during the 2007 American Library Association Annual Meeting in Washington, D.C.
- Organizer of the colloquium presentation and reception for Tamar Sadeh of Ex Libris on PRIMO June 2007
- Practicum review and design (work with potential supervisors, such as the American Indian Museum internship description revision) 2006-
- Comprehensive examinations (edits, proctoring, and grading), 1990-
- SLIS Web site redesign: Comments and suggestions. Fall 2007
- Conducted surveys of current students and alumni in preparation for the 2005 re-accreditation, 2004-2005
- Technology Committee, 1992-1999 (chair, 1996-1998), 2002-2003 (member)
- Colloquia Committee 1997-1999, 2002-2003.
- Advisor of the CUA Student Chapter of the American Society for Information Science and Technology, 2002-2003
- Visiting Professor Search Committee, 1999, 2000, 2001
- Leader, Participation in the CORC experiment, 1999-2000
- Advisor of the Special Libraries Association Student Chapter, 1993-1999; the group was recognized for outstanding leadership by SLA in 1999.

- COA planning Committee, Task Force on Electronic Presentation of SLIS Reports (team leader) 1997-1998
- COA Planning Committee, Subcommittee on Technology 1996-1998
- NLM practicum coordinator, 1997-1998
- Computer Literacy Workshops: Assisted with the development and evaluation of the workshops, 1996-1998
- Leader, Participation in the InterCat project, 1995-1997

Appendix B

MARC 21 Reference Materials

Part VII: A Summary of Commonly Used MARC 21 Fields
Part VIII: A List of Other Fields Often Seen in MARC Records
Part IX: The Leader
Part X: Field 008 for Books

Part VII:

A Summary of Commonly Used MARC 21 Fields

This is a summary of the MARC 21 tags used most frequently by libraries in entering their own bibliographic records. For full listings of all MARC 21 tags, indicators, and subfield codes, see *MARC 21 Format for Bibliographic Data*.

In the explanations on these pages:

Tags -- The tags (3-digit numbers) are followed by the names of the fields they represent. In this summary, and in the *MARC 21 Format for Bibliographic Data*, if a tag can appear more than once in one bibliographic record, it is labeled repeatable (R). If it can only be used once, it is labeled non-repeatable (NR). For example, a catalog record can have several subjects, so the tags for subject added entries (6XX) are labeled repeatable (R).

Indicators -- The use of indicators is explained in fields where they are used. Indicators are one-digit numbers. Beginning with the 010 field, in every field -- following the tag -- are two character positions, one for Indicator 1 and one for Indicator 2. The indicators are not actually defined in all fields, however. And it is possible that a 2nd indicator will be used, while the 1st indicator remains undefined (or vice versa). When an indicator is undefined, the character position will be represented by the character # (for blank space).

Subfield codes -- All the data in each field (beginning with the 010 field) is divided into subfields, each of which is preceded by a delimiter-subfield code combination. The most common subfield codes used with each tag are shown. Each subfield code is preceded by the character $, signifying a delimiter. The name of the subfield follows the code.

In general, every field MUST have a subfield 'a' (**$a**). One exception that is often seen is in Field 020 (ISBN), when the ISBN information (subfield **$a**) is unavailable but the price (subfield **$c**) is known. Some subfields are repeatable. In this summary, repeatability is noted for only the more common repeatable subfields.

Examples: Examples follow the explanation for each field. For clarity, one space has

Ford Motor Co.
Exhibit 1030
Page 063

been placed between the tag and the first indicator, one space has been placed between the second indicator and the first delimiter- subfield code, and one space has been inserted between the delimiter-subfield code and the subfield data.

010 Library of Congress Control Number -- (LCCN)
(NR, or Not Repeatable

Indicators undefined.
Subfield used most often:
$a -- Library of Congress control number

```
Example:    010 ##  $a ###86000988#
```

020 International Standard Book Number -- (ISBN)
(R, or Repeatable)

Indicators undefined.
Subfields used most often:
$a -- International Standard Book Number
$c -- Terms of availability (often a price)
$z -- Cancelled/invalid ISBN (R)

```
Example:    020 ##  $a 0877547637
```

040 Cataloging source -- (NR)

Indicators undefined.
Subfields used most often:
$a -- Original cataloging agency
$c -- Transcribing agency
$d -- Modifying agency (R)

```
Example:    040 ## $a DLC
                   $c DLC
                   $d gwhs
```

100 Main entry -- Personal name -- (primary author)
(NR; there can be only one main entry)

Ford Motor Co.
Exhibit 1030
Page 064

Indicator 1: Type of personal name entry element
0 -- Forename
1 -- Surname (this is the most common form)
3 -- Family name
Indicator 2 undefined.
Indicator 2 became obsolete in 1990. Older records may display 0 or 1
Subfields used most often:
$a -- Personal name
$b -- Numeration
$c -- Titles and other words associated with a name (R)
$q -- Fuller form of name
$d -- Dates associated with a name (generally, year of birth)

```
Example:    100 1# $a Gregory, Ruth W.
                   $q (Ruth Wilhelme),
                   $d 1910-
```

130 Main entry -- Uniform title -- (NR)

Indicator 1: Nonfiling characters
0-9 -- Number of nonfiling characters present (for initial articles, including spaces)
Indicator 2 undefined.
Indicator 2 became obsolete in 1990. (See 100 above.)
Subfields used most often:
$a -- Uniform title
$p -- Name of part/section of a work (R)
$l -- Language of a work
$s -- Version
$f -- Date of a work

```
Example:    130 0# $a Bible.
                   $p O.T.
                   $p Psalms.
```

240 Uniform title (NR)

Indicator 1: Uniform title printed or displayed
0 -- Not printed or displayed
1 -- Printed or displayed (most common)
Indicator 2: Nonfiling characters

Ford Motor Co.
Exhibit 1030
Page 065

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

Subfields used most often:

$a -- Uniform title
$l -- Language of a work
$f -- Date of a work

```
Example:    240 10 $a Ile mystérieuse.
                   $l English.
                   $f 1978
```

245 Title Statement (NR)

Indicator 1: Title added entry

(Should the title be indexed as a title added entry?)

0 -- No title added entry

(indicates a title main entry; i.e. no author is given)

1 -- Title added entry

(the proper indicator when an author given in 1XX; the most common situation)

Indicator 2: Nonfiling characters

0-9 -- Number of nonfiling characters present, including spaces; usually set at zero, except when the title begins with an article; e.g., for *The robe*, the second indicator would be set to 4. The letters *T*, *h*, *e*, and the space following them are then ignored in alphabetizing titles. The record will be automatically filed under "*r*" -- for *Robe*.

Subfields used most often:

$a -- Title proper
$h -- Medium (often used for non-book media)
$p -- Name of part/section of a work (R)
$b -- Reminder of title (subtitles, etc.)
$c -- Remainder of title page transcription/Statement of responsibility

```
Example:    245 14 $a The DNA story :
                   $b a documentary history of gene
                      cloning /
                   $c James D. Watson, John Tooze.
```

246 Varying form of title (R)

Ford Motor Co.
Exhibit 1030
Page 066

Indicator 1: Note/title added entry controller
1 -- Note, title added entry
3 -- No note, title added entry
Indicator 2: Type of title
-- No information provided
0 -- Portion of title
1 -- Parallel title
4 -- Cover title
8 -- Spine title
Subfield used most often:
$a -- Title proper

```
Example:    246 3# $a Four corners power review
```

250 Edition statement (NR)

Indicators undefined.
Subfield used most often:
$a -- Edition statement

```
Example:    250 ## $a 6th ed.
```

260 Publication, distribution, etc. (Imprint) (R)

Indicator 1: Sequence of publishing statements
-- No information provided
Indicator 2: Undefined
Subfields used most often:
$a -- Place of publication, distribution, etc. (R)
$b -- Name of publisher, distributor, etc. (R)
$c -- Date of publication, distribution, etc. (R)

```
Example:    260 ## $a New York :
                   $b Chelsea House,
                   $c 1986.
```

300 Physical description (R)

Ford Motor Co.
Exhibit 1030
Page 067

Indicators undefined.
Subfields used most often:
$a -- Extent (number of pages) (R)
$b -- Other physical details (usually illustration information)
$c -- Dimensions (cm.) (R)
$e -- Accompanying material (for example, "teacher's guide" or "manual")

```
Example:    300 ## $a 139 p. :
                   $b ill. ;
                   $c 24 cm.
```

440 Series statement / Added entry--Title

This field was made obsolete in 2008 to simplify the series statement. See 490 and 830.

490 Series statement (No added entry is traced from field) (R)

Indicator 1: Specifies whether series is traced (whether an 8XX tag is also present)
0 -- Series not traced
1 -- Series traced (8XX is in record)
Indicator 2 undefined.
Subfield used most often:
$a -- Series statement (R)
$v -- Volume number (R)

```
Example:    490 1# $a Colonial American craftsmen
```

500 General note (R)

Indicators undefined.
Subfield used most often:
$a -- General note (Used when no specialized note field has been defined for the information. Examples: Notes regarding the index; the source of the title; variations in title; descriptions of the nature, form, or scope of the item.)

```
Example:    500 ##  $a Includes index.
```

504 Bibliography, etc. note (R)

Ford Motor Co.
Exhibit 1030
Page 068

Indicators undefined.
Subfield used most often:
$a -- Bibliography, etc. note

```
Example:    504 ## $a Includes bibliographical
                      references.
```

505 Formatted contents note (R)

Indicator 1: Type of contents note
0 -- Complete contents
1 -- Incomplete contents (used with multivolume set when some volumes are not yet published)
2 -- Partial contents
Indicator 2: Level of content designation
-- Basic
Subfield used most often:
$a -- Formatted contents note

```
Example:    505 0# $a Pride and prejudice -- Emma
                      -- Northanger Abbey.
```

520 Summary, etc. note (R)

Indicator 1: Display constant controller
-- Summary
1 -- Review
2 -- Scope and content
3 -- Abstract
Indicator 2 undefined
Subfields used most often
$a -- Summary, abstract, or annotation
$b -- Expansion of summary note

```
Example:    520 ## $a This basic guide to parliamentary
                      procedure tells how to conduct
                      and participate in a meeting
                      properly.
```


Ford Motor Co.
Exhibit 1030
Page 069

600 Subject added entry -- Personal name (R)

Indicator 1: Type of personal name entry element
- 0 -- Forename
- 1 -- Surname (this is the most common form)
- 3 -- Family name

Indicator 2: Subject heading system/thesaurus (identifies the specific list or file which was used)
- 0 -- Library of Congress Subject Headings
- 1 -- LC subject headings for children's literature
- 2 -- Medical Subject Headings
- 3 -- National Agricultural Library subject authority file
- 4 -- Source not specified
- 5 -- Canadian Subject Headings
- 6 -- Répertoire de vedettes-matière
- 7 -- Source specified in subfield $2

(Note regarding Sears subject headings: The MARC 21 format does not provide an assigned indicator for Sears subject headings. Therefore, an indicator of 7 is used, and the MARC defined code "sears" is placed in subfield $2.)

Subfields used most often:
- **$a** -- Personal name (surname and forename)
- **$b** -- Numeration
- **$c** -- Titles and other words associated with a name (R)
- **$q** -- Fuller form of name
- **$d** -- Dates associated with a name (generally, year of birth)
- **$t** -- Title of a work
- **$v** -- Form subdivision (R)
- **$x** -- General subdivision (R)
- **$y** -- Chronological subdivision (R)
- **$z** -- Geographic subdivision (R)
- **$2** -- Source of heading or term (used with 2nd indicator of 7)

```
Example:    600 10 $a Shakespeare, William,
                   $d 1564-1616
                   $x Comedies
                   $x Stage history.
```

```
Example:    600 10 $a Shakespeare, William,
                   $d 1564-1616
                   $x Knowledge
```


Ford Motor Co.
Exhibit 1030
Page 070

```
                        $z Rome
                        $v Congresses.
```

Notice that subfields $v, $x, and $z in the 600 field are repeatable. Subfields $v, $x, $y, and $z do not have to be in alphabetical order. They will be in the order prescribed by the instructions given by the subject heading system.

610 Subject added entry -- Corporate name (R)

Indicator 1: Type of corporate name entry element
- 0 -- Inverted name (not used with AACR2)
- 1 -- Jurisdiction name
- 2 -- Name in direct order

Indicator 2: Subject heading system/thesaurus.
- See indicator 2 under 600

Subfields used most often:
- **$a** -- Corporate name or jurisdiction name as entry element
- **$b** -- Subordinate unit (R)
- **$v** -- Form subdivision (R)
- **$x** -- General subdivision (R)
- **$y** -- Chronological subdivision (R)
- **$z** -- Geographic subdivision (R)
- **$2** -- Source of heading or term (used with 2nd indicator of 7)

```
Example:    610 10 $a United States.
                   $b Army Air Forces
                   $v Biography.
```

650 Subject added entry -- Topical term (Most subject headings fit here.) (R)

Indicator 1: Level of subject
- # -- No information provided

Indicator 2: Subject heading system/thesaurus
(identifies the specific list or file which was used)
- 0 -- Library of Congress Subject Headings
- 1 -- LC subject headings for children's literature
- 2 -- Medical Subject Headings
- 3 -- National Agricultural Library subject authority file
- 4 -- Source not specified
- 5 -- Canadian Subject Headings
- 6 -- Répertoire de vedettes-matière
- 7 -- Source specified in subfield $2

Ford Motor Co.
Exhibit 1030
Page 071

Note regarding Sears subject headings: The MARC 21 format does not provide an assigned indicator for Sears subject headings. Therefore, an indicator of 7 is used, and the MARC defined code "sears" is placed in subfield $2.)

Subfields used most often:
$a -- Topical term
$v -- Form subdivision (R)
$x -- General subdivision (R)
$y -- Chronological subdivision (R)
$z -- Geographic subdivision (R)
$2 -- Source of heading or term used with 2nd indicator of 7)

```
Example:    650 #0 $a Theater
                   $z United States
                   $v Biography
                   $v Dictionaries.
```

Notice that subfields $v, $x, and $z in the 650 field are repeatable. Subfields $v, $x, $y, and $z do not have to be in alphabetical order. They will be in the order prescribed by the instructions given by the subject heading system.

651 Subject added entry -- Geographic name (R)

Indicator 1: undefined.
Indicator 2: Subject heading system/thesaurus.
See indicator 2 under 600
Subfields used most often:
$a -- Geographic name
$v -- Form subdivision (R)
$x -- General subdivision (R)
$y -- Chronological subdivision (R)
$z -- Geographic subdivision (R)
$2 -- Source of heading or term (used with 2nd indicator of 7)

```
Example:    651 #0 $a United States
                   $x History
                   $v Chronology.
```

Notice that subfields $v, $x, and $z in the 651 field are repeatable. Subfields $v, $x, $y, and $z do not have to be in alphabetical order. They will be in the order prescribed by the instructions given by the subject heading system.

Ford Motor Co.
Exhibit 1030
Page 072

700 Added entry -- Personal name (R)

Indicator 1: Type of personal name entry element
0 -- Forename
1 -- Surname (this is the most common form)
3 -- Family name
Indicator 2: Type of added entry
-- No information provided (most common; co-authors, editors, etc.)
2 -- Analytical entry (The values for Indicator 2 changed in 1994 with Format Integration, and older records may display additional values. An analytical entry involves an author/title of an item contained in a work.)
Subfields used most often:
$a -- Personal name
$b -- Numeration
$c -- Titles and other words associated with a name (R)
$q -- Fuller form of name
$d -- Dates associated with a name (generally, year of birth)
$e -- Relator term (such as ill.) (R)
$4 -- Relator code (R)

```
Example:    700 1# $a Baldridge, Letitia.
```

710 Added entry -- Corporate name (R)

Indicator 1: Type of corporate name entry element
0 -- Inverted name (not used with AACR2)
1 -- Jurisdiction name
2 -- Name in direct order
Indicator 2: Type of added entry.
See Indicator 2 under 700
-- No information provided
2 -- Analytical entry
Subfields used most often:
$a -- Corporate name or jurisdiction name as entry element
$b -- Subordinate unit (R)

```
Example:    710 2# $a Sunburst Communications (Firm)
```

740 Added entry -- Uncontrolled related/analytical title (R)

Indicator 1: Nonfiling characters

Ford Motor Co.
Exhibit 1030
Page 073

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

*Indicator 2:*Type of added entry. See Indicator 2 under 700

-- No information provided

2 -- Analytical entry

(This field was redefined in 1994 with Format Integration. Prior to 1994, the field was also used for variant titles, such as a different wording on a spine title. In records created since Format Integration, those variant tiles appear in a 246 field.)

Subfield used most often:

$a -- Title

```
Example:    740 02 $a Uncle Vanya.
```

800 Series added entry -- Personal name (R)

Indicator 1: Type of personal name entry element

0 -- Forename

1 -- Surname

3 -- Family name

Indicator 2 undefined.

Subfields used most often:

$a -- Personal name

$b -- Numeration

$c -- Titles and other words associated with a name (R)

$q -- Fuller form of name

$d -- Dates associated with a name (generally, year of birth)

$t -- Title of a work (the series)

$v -- Volume number

```
Example:    800 1# $a Fisher, Leonard Everett.
                   $t Colonial American craftsmen.
```

830 Series added entry -- Uniform title (R)

Indicator 1 undefined.

Indicator 2: Nonfiling characters

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

Subfield used most often:

$a -- Uniform title

Ford Motor Co.
Exhibit 1030
Page 074

$v -- Volume number

```
Example:    830 #0 $a Railroads of America (Macmillan)
```

[Back to Top of Page]

Part VIII:

A List of Other Fields Often Seen in MARC Records

001	Control number
003	Control number identifier
005	Date and time of latest transaction
006	Fixed-length data elements -- additional material characteristics
007	Physical description fixed field
008	Fixed length data elements (See Part X)
022	International Standard Serial Number (ISSN)
037	Source of acquisition
041	Language code
043	Geographic area code
050	Library of Congress call number
060	National Library of Medicine call number
082	Dewey Decimal classification number (the one recommended by the Library of Congress; locally-assigned call numbers may appear elsewhere)
110	Main entry -- Corporate name (less frequent under AACR2 rules)
256	Computer file characteristics
263	Projected publication date (indicates a CIP -- Cataloging in Publication -- record)
306	Playing time
508	Creation/production credits note
510	Citation/references note (review sources)
511	Participant or performer note
521	Target audience note (first indicator: 0 = reading grade level, 1 = interest age level, 2 = interest grade level, 3 = special audience characteristics, 4 = motivation interest level)
530	Additional physical form available note
538	System details note
586	Awards note
656	Index term -- Occupation
730	Added entry -- Uniform title
852	Location

Ford Motor Co.
Exhibit 1030
Page 075

856 Electronic location and access
9XX Reserved for local use. (They are used by vendors, systems, or individual libraries to exchange additional data)

[Back to Top of Page]

Part IX:

The Leader

There are 24 positions in the Leader, numbered from 00 to 23. For fuller explanation, see the *MARC 21 Format for Bibliographic Data*.

00-04 Record length (calculated by the computer for each record)
05 Record status
- a = increase in encoding level
- c = corrected or revised
- d = deleted
- n = new
- p = increase in encoding from prepublication (previous CIP)

06 Type of record
- a = language material
- c = printed music
- d = manuscript music
- e = cartographic material
- f = manuscript cartographic material
- g = projected medium
- i = nonmusical sound recording
- j = musical sound recording
- k = 2-dimensional nonprojectable graphic
- m = computer file
- o = kit
- p = mixed materials
- r = 3-dimensional artifact or naturally occurring object
- t = manuscript language material

07 Bibliographic level
- a = monographic component part
- b = serial component part
- c = collection
- d = subunit
- i = integrating resource
- m = monograph/item
- s = serial

08 Type of control
- # = no specified type
- a = archival

Ford Motor Co.
Exhibit 1030
Page 076

09 Character coding scheme
- # = MARC-8
- a = UCS/Unicode

10 Indicator count (always "2")
11 Subfield code count (always "2")
12-16 Base address of data (calculated by the computer for each record)
17 Encoding level
- # = full level
- 1 = full level, material not examined
- 2 = less-than-full level, material not examined
- 3 = abbreviated level
- 4 = core level
- 5 = partial (preliminary) level
- 7 = minimal level
- 8 = prepublication level (CIP)
- u = unknown
- z = not applicable

18 Descriptive cataloging form
- # = non-ISBD
- a = AACR2
- i = ISBD
- u = unknown

19 Multipart resource record level
- # = Not specified or not applicable
- a = Set
- b = Part with independent title
- c = Part with dependent title

20 Length of the length-of-field portion (always "4")
21 Length of the starting-character-position portion (always "5")
22 Length of the implementation-defined portion (always "0")
23 Undefined (always "0")

[Back to Top of Page]

Part X:

Field 008 for Books

Field 008 is used for Fixed Length Data Elements ("Fixed Field Codes"). There are 40 character positions in field 008, numbered from 00-39. Undefined positions must contain either a blank (#) or a fill character (|). Positions 00-17 and 35-39 are defined the same way for all media.

The information shown here for positions 18-34 applies only to books. For explanation of all the positions below and for positions 18-34 for other media, see the *MARC 21 Format*

Ford Motor Co.
Exhibit 1030
Page 077

for Bibliographic Data.

Note that field 008 has no indicators or subfield codes.

00-05 Date entered on file (YYMMDD),
where Y=year, M=month, and D=day

06 Type of date/publication status:
- b = no dates given; B.C. date involved
- e = detailed date
- s = single known date/probable date
- m = multiple dates
- r = reprint/reissue date (Date 1) and original date (Date 2)
- n = dates unknown
- q = questionable date
- t = publication date and copyright date
- | = no attempt to code

07-10 Date 1/beginning date of publication

11-14 Date 2/ending date of publication

Date fields contain the year(s) of publication. The type of date(s) in these elements are specified in fixed field element 06: Type of date/publication status. (For further details, see the field 008 description in the *MARC 21 Format for Bibliographic Data*.)

15-17 Place of publication, production, or execution
For example:
- pk# = Pakistan
- cau = California (US)

(For a full list of codes used in these positions, see the *MARC Code List for Countries*.)

18-21 Illustrations (up to 4 codes):
- # = no illustrations
- a = illustrations
- b = maps
- c = portraits
- d = charts
- e = plans
- f = plates
- g = music
- h = facsimiles
- i = coats of arms
- j = genealogical tables
- k = forms
- l = samples
- m = phonodisc, phonowire, etc.
- o = photographs
- p = illuminations
- | = no attempt to code

Ford Motor Co.
Exhibit 1030
Page 078

22 Target audience:
- # = unknown or not specified
- a = preschool
- b = primary
- c = pre-adolescent
- d = adolescent
- e = adult
- f = specialized
- g = general
- j = juvenile
- | = no attempt to code

23 Form of item:
- # = none of the following
- a = microfilm
- b = microfiche
- c = microopaque
- d = large print
- f = braille
- r = regular print reproduction
- s = electronic
- | = no attempt to code

24-27 Nature of contents (up to 4):
- # = no specified nature of contents
- a = abstracts/summaries
- b = bibliographies (is one or contains one)
- c = catalogs
- d = dictionaries
- e = encyclopedias
- f = handbooks
- g = legal articles
- i = indexes
- j = patent document
- k = discographies
- l = legislation
- m = theses
- n = surveys of literature
- o = reviews
- p = programmed texts
- q = filmographies
- r = directories
- s = statistics
- t = technical reports
- u = standards/specifications
- v = legal cases and notes
- w = law reports and digests
- z = treaties
- | = no attempt to code

Ford Motor Co.
Exhibit 1030
Page 079

28 Government publication:
- # = not a government publication
- i = international intergovernmental
- f = federal/national
- a = autonomous or semi-autonomous component
- s = state, provincial, territorial, dependent, etc.
- m = multistate
- c = multilocal
- l = local
- z = other type of government publication
- o = government publication -- level undetermined
- u = unknown if item is government publication
- | = no attempt to code

29 Conference publication:
- 0 = not a conference publication
- 1 = conference publication
- | = no attempt to code

30 Festschrift:
- 0 = not a festschrift
- 1 = festschrift
- | = no attempt to code

31 Index:
- 0 = no index
- 1 = index present
- | = no attempt to code

32 Undefined (since 1990) (Earlier records may contain the values 0 or 1)
- # = Undefined
- | = no attempt to code

33 Literary form:
- 0 = not fiction (not further specified)
- 1 = fiction (not further specified)
- c = comic strips
- d = dramas
- e = essays
- f = novels
- h = humor, satires, etc.
- i = letters
- j = short stories
- m = mixed forms
- p = poetry
- s = speeches
- u = unknown
- | = no attempt to code

34 Biography:
- # = no biographical material
- a = autobiography
- b = individual biography

Ford Motor Co.
Exhibit 1030
Page 080

c = collective biography
d = contains biographical information
| = no attempt to code

35-37 Language:
A three-letter code. For example: eng fre ger spa rus ita

(For a full list of codes used in these positions, see the *MARC Code List for Languages*.)

38 Modified record:
= not modified
x = missing characters (because of characters unavailable in MARC character set)
s = shortened
d = "dashed-on" information omitted
r = completely romanized/printed cards in script
o = completely romanized/printed cards romanized
| = no attempt to code

39 Cataloging source:
= national bibliographic agency
c = cooperative cataloging program
d = other sources
u = unknown
| = no attempt to code

[Back to Top of Page]

[Back to Table of Contents] -- [Continue to Part 11]



Library of Congress
Library of Congress Help Desk (10/27/2009)

Ford Motor Co.
Exhibit 1030
Page 081

Exhibit 1007

A Flexible Spread-Spectrum Multi-Carrier Multiple-Access System for Multi-Media Applications

Stefan Kaiser and Khaled Fazel
German Aerospace Research Establishment (DLR), Institute for Communications Technology
D-82234 Oberpfaffenhofen, Germany, Tel.: +49 8153 28-2805, Fax: +49 8153 28-1442
e-mail: Stefan.Kaiser@dlr.de, Khaled.Fazel@dlr.de

Abstract

This paper presents a novel multiple access scheme designed for mobile multi-media applications. The scheme is referred to as spread-spectrum multi-carrier multiple-access (SS-MC-MA) and exploits the advantages given by the combination of the spread-spectrum technique with multi-carrier modulation, known from multi-carrier code division multiple access (MC-CDMA) schemes. In contrast to MC-CDMA, the SS-MC-MA separates the users' signals according to an FDMA scheme on sub-carrier level. Thus, the novel scheme avoids multiple access interference (MAI) within a cell. SS-MC-MA can use the same radio interface for the up- and the down-link. Its bandwidth efficiency is comparable to the high bandwidth efficiency achievable with MC-CDMA in the down-link. By providing a high degree of flexibility, this system can be a good candidate for the future universal mobile telecommunications systems (UMTS).

1 Introduction

Personal communication services are of great importance for the third-generation mobile radio communications systems such as UMTS [1]. The objectives of the third-generation systems go far beyond that, what is known from the second-generation systems such as IS-54, GSM, or IS-95, especially with respect to the wide range of multi-media services (speech, audio, image, video, data) with bit rates up to 2 Mbit/s, the high quality of service requirements (speech/image quality, bit error rate (BER) performance, and user capacity), and the operation in mixed cell scenarios (macro, micro, pico) and different environments (indoor/outdoor, business/domestic, cellular/cord-less). Present second-generation systems are primarily designed as low-rate voice and data systems (usually single-environment such as indoor or outdoor) and fall short of meeting important requirements of the third-generation systems.

A CDMA scheme could be a potential candidate for a third-generation system [2, 3]. However, it has to cope with the presence of MAI. The recently proposed SS-MC-MA scheme [4] avoids MAI due to an FDMA scheme on sub-carrier level by nevertheless exploiting the diversity gain offered by the spread-spectrum technique. Furthermore, the same radio interface can be used for the up- and the down-link. SS-MC-MA is related to CDMA schemes using orthogonal multi-carrier modulation, known as MC-CDMA or orthogonal frequency division multiplexing (OFDM)-CDMA [5, 6]. Thus, SS-MC-MA achieves similar performance as MC-CDMA. Performance comparisons between direct sequence (DS)-CDMA and MC-CDMA for the down-link of a mobile communication system show that MC-CDMA can achieve a higher bandwidth efficiency than DS-CDMA [7, 8]. The application of MC-CDMA for the up-link has rarely been investigated. The main problem encountered here, is the design of an efficient channel estimation technique for the up-link.

In this paper an SS-MC-MA scheme is investigated for the up-link of a mobile radio system by taking into account channel coding with soft decision decoding and real channel estimation. The multiple access scheme is based on a combination of the spread-spectrum technique and orthogonal multi-carrier modulation. The basic difference between SS-MC-MA and MC-CDMA is that with SS-MC-MA the code division is used for simultaneous transmission of the data of a single user on the same sub-carriers, where in the case of MC-CDMA the code division is used for the transmission of the data of different users on the same sub-carriers. SS-MC-MA assigns each user exclusively its own sub-carriers according to an FDMA scheme. Since all superimposed modulated spreading codes of one sub-set of sub-carriers are affected by the same channel fading, a simple channel estimation can be realized. For joint detection of the data of a single user a low complex maximum-likelihood sequence estimation (MLSE) is applied.

The SS-MC-MA scheme provides a high degree of flexibility due to variable sub-carrier allocation for variable data rate transmission and can offer unequal error protection due to variable spreading code mapping.

2 SS-MC-MA System

The basic idea of SS-MC-MA is to exploit the diversity gain offered by the spread-spectrum technique and interleaving by nevertheless avoiding MAI within a cell. This can be achieved by using synchronous MC transmission, where the multiple access is performed in the frequency domain on sub-carrier level. The data of each active user are mapped to a set of sub-carriers, which is exclusively assigned to this user. In order to exploit the frequency diversity of the channel, a user spreads the data over its set of sub-carriers and superimpose them by code multiplex. The superimposing of the data assigned to a single

0-7803-3871-5/97/$10.00 © 1997 IEEE

Authorized licensed use limited to: Ingrid Hsieh-Yee. Downloaded on March 19,2023 at 18:43:25 UTC from IEEE Xplore. Restrictions apply.

Ford Motor Co.
Exhibit 1030
Page 083

user introduces self interference which is comparable to MAI in MC-CDMA systems. However, the self interference can easy be counteracted by use of an MLSE, since implicitly knowledge about the used spreading codes exists. In the case of MC-CDMA with MLSE, knowledge about the actually used spreading codes of the other users is required, which additionally increases the signaling overhead.

2.1 Transmitter

Fig. 1 shows the block diagram of an SS-MC-MA transmitter (same for up- and down-link) for the data of user i, $i = 1 \ldots N_u$, where N_u is the number of active users within a cell. After channel encoding, interleaving, and

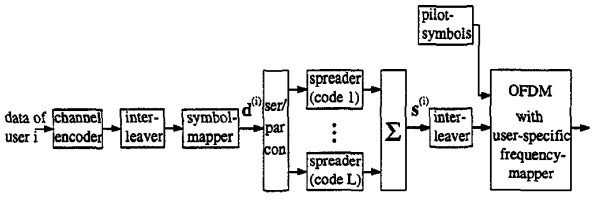


Figure 1: SS-MC-MA transmitter for the data of user i.

symbol mapping, L subsequent complex-valued data-symbols $d_l^{(i)}, l = 1 \ldots L$, of user i are serial to parallel converted. The vector $\mathbf{d}^{(i)}$ represents one block of L parallel data-symbols of user i,

$$\mathbf{d}^{(i)} = (d_1^{(i)}, d_2^{(i)} \ldots d_L^{(i)})^{\mathrm{T}}, \qquad (1)$$

where $(.)^{\mathrm{T}}$ denotes transposition. Each of the L data-symbols is multiplied with an orthogonal spreading code (e.g. Walsh-Hadamard code [9]) of length L. The $L \times L$ matrix $\mathbf{C}$ represents the L orthogonal spreading codes. The modulated spreading codes are symbol- and with that chip-synchronously added, resulting in the transmission vector

$$\mathbf{s}^{(i)} = \mathbf{C} \cdot \mathbf{d}^{(i)} = (S_1^{(i)}, S_2^{(i)} \ldots S_L^{(i)})^{\mathrm{T}}, \qquad (2)$$

consisting of L chips. To increase the robustness of the SS-MC-MA scheme e.g. against inter cell interference, less than L modulated spreading codes of length L can be added to one transmission vector $\mathbf{s}^{(i)}$.

The chips of M subsequent vectors $\mathbf{s}^{(i)}$ are frequency- and time-interleaved and fed to the OFDM unit, which also performs a user-specific frequency mapping, illustrated in Fig. 2. The frequency mapping may be equidistant or pseudo-random. After the mapping, the chips $S_l^{(i)}, l = 1 \ldots L$, are interpreted as complex frequency values of the sub-carriers. Each user of a cell transmits on a sub-set of sub-carriers, where the sub-sets of sub-carriers of the users within a cell are disjunct. To exploit the frequency diversity of the whole transmission channel, the sub-carriers assigned to a user are distributed over the whole transmission bandwidth. Finally, each OFDM-symbol of duration T_s is cyclic extended by a guard interval Δ in the OFDM unit to prevent inter symbol interference (ISI) in multipath channels. In order to perform coherent detection at the receiver and to

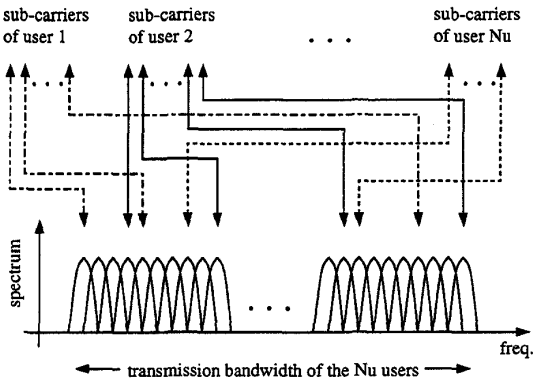


Figure 2: Pseudo-random user-specific frequency-mapping and -demapping.

guarantee robust time and frequency synchronization, pilot-symbols are multiplexed in the transmitted data.

2.2 Receiver

The block diagram of the SS-MC-MA receiver with coherent detection of the data of the i-th user is shown in Fig. 3. The received data are fed to the inverse OFDM

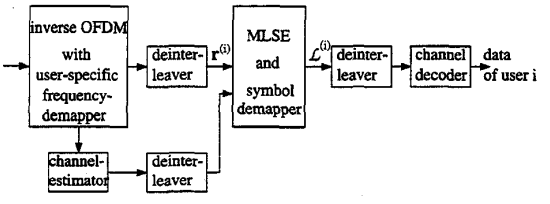


Figure 3: SS-MC-MA receiver for the data of user i.

unit which performs the user specific frequency demapping (see Fig. 2) and, furthermore, extracts the pilot-symbols from the symbols with user data. The symbols with user data are frequency and time deinterleaved, resulting in the vector

$$\mathbf{r}^{(i)} = \mathbf{H}^{(i)} \cdot \mathbf{s}^{(i)} + \mathbf{n}^{(i)} = (R_1^{(i)}, R_2^{(i)} \ldots R_L^{(i)})^{\mathrm{T}}. \qquad (3)$$

The $L \times L$ diagonal matrix $\mathbf{H}^{(i)}$ describes the complex channel fading on the L sub-carriers of user i, and the vector $\mathbf{n}^{(i)}$ represents additive white Gaussian noise. The vector $\mathbf{r}^{(i)}$ is fed to the detector.

From the received pilot-symbols, the channel estimator estimates the channel state information (CSI), represented by the $L \times L$ diagonal matrix $\hat{\mathbf{H}}^{(i)}$, which is also frequency- and time-deinterleaved and used for coherent detection.

Joint detection with MLSE is applied for the superimposed data of a user. In the SS-MC-MA system we have the advantage that the MLSE in one estimation decides on $2L$ code bits of a single user if QPSK mapping is used. Thus, compared to MC-CDMA, the complexity of the MLSE per bit reduces by a factor of $2L$ in the downlink. Conventional and other joint detection techniques

Authorized licensed use limited to: Ingrid Hsieh-Yee. Downloaded on March 19,2023 at 18:43:25 UTC from IEEE Xplore. Restrictions apply.

Ford Motor Co.
Exhibit 1030
Page 084

known for MC-CDMA receivers [10] can also be applied in SS-MC-MA receivers.

The MLSE evaluates the squared Euclidean distances $\delta^2_{m,I}$ and $\delta^2_{m,Q}$ between the received and all possible transmitted sequences $\mathbf{s}_m, m = 1 \ldots 2^{L+1}$, separately for the I- and the Q-components of $\mathbf{s}^{(i)}$ to reduce the complexity of the MLSE,

$$\delta^2_{m,I} = \| \Re\{\hat{\mathbf{H}}^{(i)*} |\hat{\mathbf{H}}^{(i)}|^{-1} \mathbf{r}^{(i)} - |\hat{\mathbf{H}}^{(i)}| \mathbf{s}_m\}\|^2 , \quad (4)$$

$$\delta^2_{m,Q} = \| \Im\{\hat{\mathbf{H}}^{(i)*} |\hat{\mathbf{H}}^{(i)}|^{-1} \mathbf{r}^{(i)} - |\hat{\mathbf{H}}^{(i)}| \mathbf{s}_m\}\|^2 , \quad (5)$$

where $\Re\{.\}$ and $\Im\{.\}$ denote the real and the imaginary part of a complex-valued vector. So, for QPSK-mapping instead of 2^{2L} only 2^{L+1} possible transmitted sequences have to be evaluated in the MLSE. The QPSK symbol demapping is inherently performed in the detection process. From the Euclidean distances the log-likelihood ratios for the received code bits are calculated as presented in [11, 4]. The log-likelihood ratios can be approximated by

$$\mathcal{L}^{(i)} \approx \frac{E_s}{N_0}(\delta^2_{m'} - \delta^2_{m''}) , \quad (6)$$

where m' is the index of the possible sequence with the smallest squared Euclidean distance δ^2_m where the estimated channel coded bit is equal to -1 and m'' is the index of the possible sequence with the smallest squared Euclidean distance δ^2_m where the estimated channel coded bit is equal to +1. Finally, after de-interleaving, the channel decoder exploits the log-likelihood ratios, yielding the detected data bits of user i.

3 Synchronization and Channel Estimation

The data transmission can be based on OFDM-frames, where each frame consists of a group of several OFDM-symbols. Each OFDM-frame may start with a null symbol (without signal energy) that is used in the receiver for coarse time-synchronization and for noise and interference estimation [12]. The fine time synchronization can be achieved by frequency domain processing of the pilot-symbols inserted into the OFDM-frame. The frequency synchronization can use a frequency offset estimation technique also based on the pilot-symbols [12].

For a synchronous transmission of the user data, the null symbol can be used to enable each user a self-synchronization. This can be achieved by searching only the null symbol in the OFDM frames transmitted by the other active users. If a larger guard-time Δ is chosen, a higher tolerance especially for the up-link on the accuracy of the synchronous transmission can be allowed.

Pilot-symbol-aided channel estimation is used for coherent detection. For the up-link one-dimensional filtering in time direction is used, where in the down-link, a more efficient two-dimensional filtering is preferable [13] since the down-link corresponds to a simple broadcasting scenario. In this paper, we focus on the up-link and thus apply one-dimensional FIR filters dimensioned as Wiener filters for the worst case channel conditions. The FIR filters are designed as Wiener filters with rectangular Doppler power spectrum, where $f_{D_{max}} = 333.3$ Hz is chosen as worst case condition. The pilot symbol spacing is chosen to $2\times$ oversample the dynamics of the fading in time direction, given $f_{D_{max}} = 333.3$ Hz. The one-dimensional FIR filter is realized with 5 taps, which is a good choice between performance and complexity [13].

4 Flexibility and Robustness

The proposed SS-MC-MA scheme inherently offers the following flexibilities and advantages:

Variable data rate transmission: It is not necessary, that the transmission vector $\mathbf{s}^{(i)}$ of length L uses L sub-carriers per OFDM-symbol for its transmission. Since, in contrast to MC-CDMA, the chips of $\mathbf{s}^{(i)}$ contains the data of only one user, each user can transmit on a different number of sub-carriers. With that, the transmission of a wide range of multi-media services with different data rates (speech, audio, image, video, data, etc.) can be easily supported by a SS-MC-MA scheme. For instance, the transmission of low data rate services will need to choose only one or few sub-carriers (see Section 5). However, for the transmission of high data rate (video, data of about 2 Mbit/s) the majority or all of the sub-carriers should be used. The required data rate is transmitted to the base station, which allocates an appropriate set of sub-carriers. The number of active users depends on the number of available sub-carriers.

Flexible TDMA component: Since a synchronous transmission is assumed, we can easily introduce a time division multiple access (TDMA) component in the OFDM-frame structure. With a TDMA component different groups of users can be supplied in subsequent OFDM-frames, resulting in an increasing number of active users.

Unequal error protection: To increase the robustness of the SS-MC-MA scheme e.g. against inter cell interference, less than L modulated spreading codes of length L can be added to one transmission vector $\mathbf{s}^{(i)}$, resulting in a higher redundancy, i.e. protection, due to spreading for a specific user.

Same radio interface for up- and down-link: Since the SS-MC-MA scheme prevents MAI within a cell, the same SS-MC-MA radio interface can be used for the up-link as well as for the down-link.

Near-far resistance: Due to the absence of MAI, the near-far resistance of SS-MC-MA is only limited by the dynamic range of the receiver. However, power control in the up- and the down-link, adapted to each user separately, would minimizes the inter cell interference. Information about the transmitted power of each user in the up-link is not necessary for detection, hence reducing the complexity of the system.

When considering inter cell interference, it is of advantage that adjacent cells operating in the same frequency band uses different frequency mapping patterns, so that the sub-carriers of a user only partly overlap with the sub-carriers used by a user of the adjacent cell. Thus, when the SS-MC-MA system is not fully loaded, the

Authorized licensed use limited to: Ingrid Hsieh-Yee. Downloaded on March 19,2023 at 18:43:25 UTC from IEEE Xplore. Restrictions apply.

Ford Motor Co.
Exhibit 1030
Page 085

distortion due to inter cell interference can be reduced. Naturally, this frequency mapping concept is also of advantage for MC-CDMA systems.

To meet the requirements of the third-generation mobile radio systems the SS-MC-MA system should in addition be combined with a flexible channel coding, where different BERs for different services have to be guaranteed by using different error-protection schemes. For instance, for speech or audio transmission a BER of about 10^{-3} is needed, where in this case one may use a simple convolutional code [14]. However, for video transmission (such as MPEG-1 or MPEG-4) one may need a much lower BER of less than 10^{-6}. In this case, a good choice would be to use a concatenated coding scheme (outer block code and inner convolutional code). Alternatively, the use of an adaptive Turbo coding scheme with different interleaver sizes can be used, where especially for larger interleaver sizes (applicable where greater delays are acceptable) significant performance improvements can be achieved [11].

Furthermore, the transmission parameters (guard-time, pilot-symbols) of an SS-MC-MA system can easily be adapted to the transmission environment by offering for instance two different modes (indoor and outdoor). The indoor channel (pico-cell) has a Doppler-shift and delay-spread much lower than the outdoor channel. This environment needs a short guard-time of about 100 ns and much less error-protection. However, the outdoor environment needs higher protection and a much longer guard time of about 20 μs. In addition, the amount of pilot-symbols used for channel estimation, depending on the Doppler-frequency and delay-spread, can be matched to the requirements of the different modes.

5 System Performance

In the following investigations, the up-link of a SS-MC-MA system is considered. The available transmission bandwidth is $B = 2$ MHz and the used carrier frequency is $f_c = 1.8$ GHz. By applying OFDM, the bandwidth B is split into $N_c = 256$ sub-channels with a sub-carrier spacing of 7.81 kHz. The resulting OFDM-symbol duration is $T_s = 128\mu$s. The OFDM and inverse OFDM is realized by an IFFT and a FFT, respectively. A guard interval of $\Delta = 20\mu$s is inserted between adjacent OFDM-symbols.

The SS-MC-MA system has a TDMA component, where, in a cycle of 4 frames (each frame consists of 31 OFDM-symbols, see Fig. 4), each frame transmits the data of another group of 32 users. Thus, $4 \times 32 = 128$ active users can transmit with a net data rate of 10.46 kbit/s. The speech encoder of the mobile station delivers a data rate of 10.46 kbit/s to the channel encoder. The binary data are convolutionally encoded with rate $R = 1/2$ and memory 6. A block of 348 code bits is randomly interleaved and QPSK-mapped into 192 data-symbols. A block of 8 serial data-symbols is serial to parallel converted and spread with orthogonal Walsh-Hadamard codes of length $L = 8$, where each of the 8 data-symbols is multiplied with another spreading

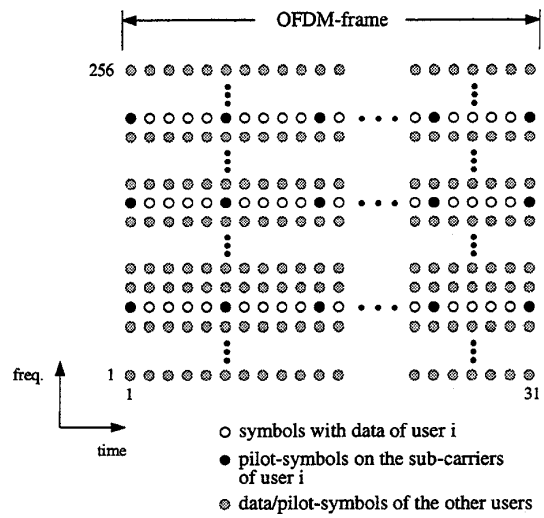


Figure 4: OFDM frame with 256 sub-carriers and 31 OFDM-symbols.

code. The 8 data modulated spreading codes are added symbol-, and with that, chip-synchronously, resulting in a transmission sequence of 8 chips. A block of 192 chips of 24 subsequent transmission sequences are randomly frequency- and time-interleaved. In the OFDM unit 8 chips are modulated on 8 sub-carriers per OFDM-symbol, where the user-specific sub-carriers are located in distances of 250 kHz, distributed over the whole bandwidth.

The receiver applies an MLSE which detects the received data-symbols of a user separately in the I- and Q-components by exploiting the estimated CSI. The CSI is estimated by FIR filtering (with 5 filter taps) of the received pilot-symbols on the user-specific sub-carriers over the time. To achieve 2× oversampling of the fading process in time direction a pilot-symbol spacing of 5 is required for the given scenarios (c.f. Fig. 4) [13].

The performance of the SS-MC-MA system is investigated for four propagation scenarios defined in COST 207 [15] by taking into account different velocities of the mobile users:

BU3: Bad urban channel with maximum delay of 10μs. The velocity of the mobile user is 3 km/h, resulting in a maximum Doppler frequency of 5 Hz.

TU50: Typical urban channel with maximum delay of 7μs. The velocity of the mobile user is 50 km/h, resulting in a maximum Doppler frequency of 83.3 Hz.

HT100: Hilly terrain with maximum delay of 20μs. The velocity of the mobile user is 100 km/h, resulting in a maximum Doppler frequency of 166.7 Hz.

RA250: Rural area with maximum delay of 0.7μs. The velocity of the mobile user is 250 km/h, resulting in a maximum Doppler frequency of 416.7 Hz. This channel should demonstrate that the proposed SS-MC-MA scheme with real channel estimation operates even in this worse scenario.

In the four channels, the Doppler spectrum is chosen according to isotropic scattering. The BER P_b versus the signal to noise ratio (SNR) E_b/N_0 of SS-MC-MA with

Authorized licensed use limited to: Ingrid Hsieh-Yee. Downloaded on March 19,2023 at 18:43:25 UTC from IEEE Xplore. Restrictions apply.

Ford Motor Co.
Exhibit 1030
Page 086

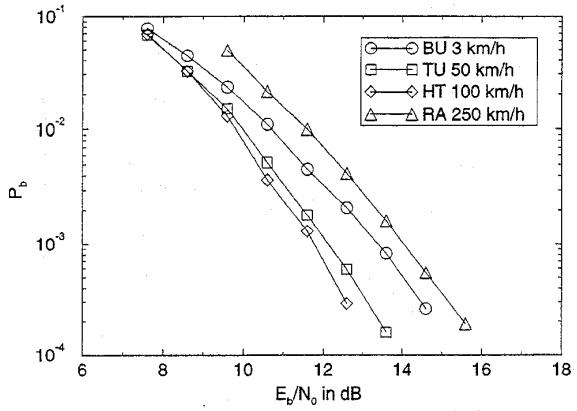


Figure 5: Performance of SS-MC-MA in different propagation scenarios with full system load.

real channel estimation is presented in Fig. 5. E_b is the average energy per received bit (data and pilot symbols are assumed to be transmitted with the same energy) and N_0 is the one-sided noise spectral density. The performance is given for a full system load, resulting in a bandwidth efficiency of 0.67 bit/s/Hz. The total SNR loss due to the guard interval Δ and the pilot symbols is 1.6 dB and is taken into account in the results of Fig. 5. It can be observed from Fig. 5, that even in a scenario with velocity of 250 km/h in a rural area, a SNR of 14 dB is sufficient to guarantee a bit error rate of $P_b < 10^{-3}$. Less critical propagation conditions requires up to 2 dB less E_b/N_0 due to higher frequency/time diversity gains and better channel estimates, respectively.

In Table 1 the SNR degradation in dB due to the proposed channel estimation scheme at a BER of 10^{-3} compared to perfect channel estimation is given for the four mobile radio channels. The loss in SNR with the

	BU3	TU50	HT100	RA250
SNR degrad.	3.1 dB	2.8 dB	2.7 dB	4.1 dB

Table 1: SNR degradation due to real channel estimation compared to perfect channel estimation.

HT100 scenario is smallest since the assigned Doppler power spectrum with $f_{D_{max}} = 166.7$ Hz best matches to the filter function of the channel estimator (c.f. Section 3) by guaranteeing 2× oversampling of the fading process.

6 Conclusions

The suitability of a novel multiple access scheme based on the spread spectrum technique and orthogonal multi-carrier modulation for mobile multi-media applications has been demonstrated. The scheme is referred to as SS-MC-MA and is based on FDMA on sub-carrier level by additionally exploiting the frequency diversity gain offered by the spread-spectrum technique. Due to the absence of MAI, SS-MC-MA is suitable for the up- and the down-link. The high flexibility of SS-MC-MA has been shown and its robustness in quite different propagation scenarios has been demonstrated, taking into account channel coding and real channel estimation. For data detection a low complex MLSE was used. The presented results show that SS-MC-MA can be applied in a typical urban environment where a mobile user has a velocity of 50 km/h @ 1.8 GHz with an SNR of $E_b/N_0 = 12$ dB to guarantee a bit error rate of $P_b = 10^{-3}$ with full system load. This yields a bandwidth efficiency of 0.67 bit/s/Hz, for the up-link as well as for the down-link.

References

[1] J. Rapeli, "UMTS: Targets, system concept, and standardization in a global framework," *IEEE Personal Communications*, pp. 20–28, February 1995.

[2] R. L. Pickholtz, L. B. Milstein, and D. L. Schilling, "Spread spectrum for mobile communications," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 313–322, May 1991.

[3] P. Jung, P. W. Baier, and A. Steil, "Advantages of CDMA and spread spectrum techniques over FDMA and TDMA in cellular mobile radio applications," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 357–364, August 1993.

[4] S. Kaiser and K. Fazel, "A spread-spectrum multi-carrier multiple-access system for mobile communications," in *Proc. IEEE Int. Workshop on Multi-Carrier Spread-Spectrum*, April 1997.

[5] K. Fazel and L. Papke, "On the performance of convolutionally-coded CDMA/OFDM for mobile communication system," in *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC'93)*, pp. 468–472, Sept. 1993.

[6] N. Yee, J.-P. Linnarz, and G. Fettweis, "Multi-carrier CDMA in indoor wireless radio networks," in *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC'93)*, pp. 109–113, Sept. 1993.

[7] S. Kaiser, "OFDM-CDMA versus DS-CDMA: Performance evaluation for fading channels," in *Proc. IEEE Int. Conf. on Commun. (ICC'95)*, pp. 1722–1726, June 1995.

[8] S. Hara and R. Prasad, "DS-CDMA, MC-CDMA and MT-CDMA for mobile multi-media communications," in *Proc. IEEE Vehic. Technol. Conf. (VTC'96)*, pp. 1106–1110, April/May 1996.

[9] J. G. Proakis, *Digital Communications*. McGraw-Hill, 1995.

[10] S. Kaiser, "Analytical performance evaluation of OFDM-CDMA mobile radio systems," in *Proc. European Personal and Mobile Commun. Conf. (EPMCC '95)*, pp. 215–220, Nov. 1995.

[11] S. Kaiser and L. Papke, "Optimal detection when combining OFDM-CDMA with convolutional and Turbo channel coding," in *Proc. IEEE Int. Conf. on Commun. (ICC'96)*, pp. 343–348, June 1996.

[12] K. Fazel, S. Kaiser, P. Robertson, and M. J. Ruf, "A concept of digital terrestrial television broadcasting," *Wireless Personal Commun.*, vol. 2, no. 1 & 2, pp. 9–27, 1995.

[13] P. Hoeher, S. Kaiser, and P. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97)*, pp. 1845–1848, April 1997.

[14] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. Commun.*, vol. 36, pp. 389–400, April 1988.

[15] Cost 207, "Digital land mobile radio communications," tech. rep., Luxembourg: Office for Official Publications of the European Communities, 1989.

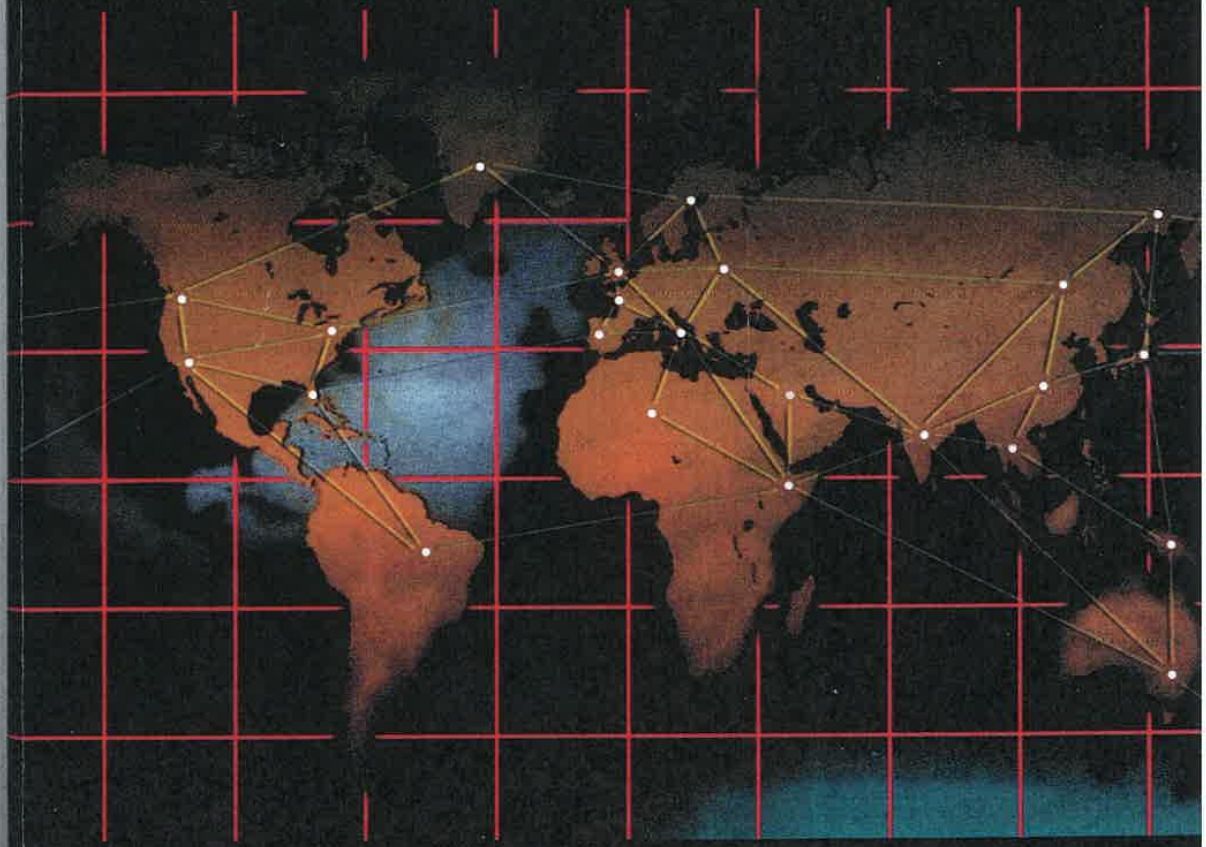
Authorized licensed use limited to: Ingrid Hsieh-Yee. Downloaded on March 19,2023 at 18:43:25 UTC from IEEE Xplore. Restrictions apply.

Ford Motor Co.
Exhibit 1030
Page 087

Exhibit 1008

THE MOBILE COMMUNICATIONS HANDBOOK

Second Edition



Editor-in-Chief

JERRY D. GIBSON

CRC PRESS

IEEE PRESS

A CRC Handbook Published in Cooperation with IEEE Press

VWGoA EX1008
U.S. Patent No. 10,833,908

ATURES

- Outstanding coverage of all aspects of mobile communications
- Expert contributions from international authors
- Authoritative account of modern technological applications
- Thorough grounding in basic principles

What's New the Second Edition?

scussions on:

- Rayleigh fading channels
- space-time processing
- radiolocation techniques
- enhancements in second generation systems
- wireless video communications
- wireless ATM
- evolving third generation standards

dates on:

- wireless data
- power control
- channel assignment
- diversity techniques
- error correction coding
- pseudonoise sequences

CRC Press Catalog Number 8597

EEE Order Number: PC5772

ISBN 0-8493-8597-0

90000

9 780849 385971

GIBSON

THE MOBILE COMMUNICATIONS HANDBOOK

525

THRIFTBOOKS

12 38501 525

CRC PRESS

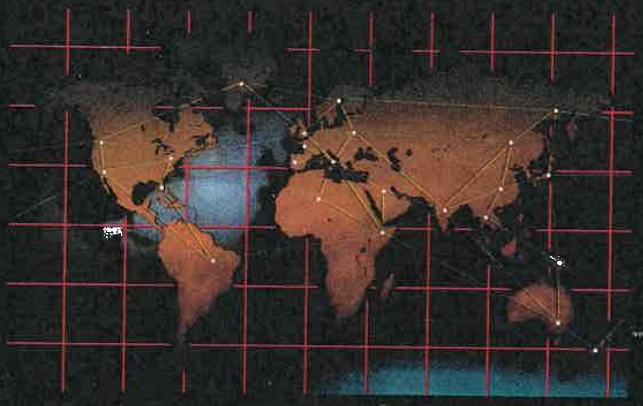
IEEE PRESS

Ford Motor Co.
Exhibit 1030
Page 090

DATA COMMUNICATIONS
NETWORKING

THE MOBILE COMMUNICATIONS HANDBOOK

Second Edition



Editor-in-Chief
JERRY D. GIBSON

The ***Mobile Communications Handbook*** covers the entire field — from principles of analog and digital communications to cordless telephones, wireless local area networks (LANs), and international technology standards. The amazing scope of the handbook ensures that it will be the primary reference for every aspect of mobile communications.

Organized into two sections, ***The Mobile Communications Handbook*** first introduces the basic principles of analog and digital communications. Consisting of tutorial articles, this section provides background information and technical details, offering a solid foundation for the spectrum of mobile communication technologies. The second part of the handbook consists of articles covering such topics as cellular mobile radio, personal communication systems, user location and addressing, wireless data, wireless LANs, and technology standards. Each article has been written as an independent contribution, but with intentional overlap between articles, providing essentially self-contained articles that can be read in whichever order desired by the reader. Nevertheless, the articles follow a logical order, thus creating continuity for the reader who wishes to proceed systematically through the handbook.

FEATURES

- Outstanding coverage of all aspects of mobile communications
- Expert contributions from international authors
- Authoritative account of modern technological applications
- Thorough grounding in basic principles

What's New in the Second Edition?

Discussions on:

- Rayleigh fading channels
- space-time processing
- radiolocation techniques
- enhancements in second generation systems
- wireless video communications
- wireless ATM
- evolving third generation standards

Updates on:

- wireless data
- power control
- channel assignment
- diversity techniques
- error correction coding
- pseudonoise sequences

CRC Press Catalog Number 8597

IEEE Order Number: PC5772

ISBN 0-8493-8597-0

9 780849 385971 90000

Ford Motor Co.
Exhibit 1030
Page 091

THE MOBILE COMMUNICATIONS HANDBOOK

Second Edition

Ford Motor Co.
Exhibit 1030
Page 092

The Electrical Engineering Handbook Series

Series Editor
Richard C. Dorf
University of California, Davis

Titles Included in the Series

The Avionics Handbook, Cary R. Spitzer
The Biomedical Engineering Handbook, Joseph D. Bronzino
The Circuits and Filters Handbook, Wai-Kai Chen
The Communications Handbook, Jerry D. Gibson
The Control Handbook, William S. Levine
The Digital Signal Processing Handbook, Vijay K. Madisetti & Douglas B. Williams
The Electrical Engineering Handbook, Richard C. Dorf
The Electric Power Engineering Handbook, L.L. Grigsby
The Electronics Handbook, Jerry C. Whitaker
The Engineering Handbook, Richard C. Dorf
The Handbook of Formulas and Tables for Signal Processing, Alexander D. Poularikas
The Industrial Electronics Handbook, J. David Irwin
The Measurement, Instrumentation, and Sensors Handbook, John G. Webster
The Mechanical Systems Design Handbook, Osita D.I. Nwokah
The Microwave Engineering Handbook, J. Michael Golio
The Mobile Communications Handbook, Jerry D. Gibson
The Ocean Engineering Handbook, Ferial El-Hawary
The Technology Management Handbook, Richard C. Dorf
The Transforms and Applications Handbook, Alexander D. Poularikas
The VLSI Handbook, Wai-Kai Chen

Ford Motor Co.
Exhibit 1030
Page 093

THE MOBILE COMMUNICATIONS HANDBOOK

Second Edition

Editor-in-Chief
JERRY D. GIBSON
Southern Methodist University
Dallas, Texas

Managing Editor
ELAINE M. GIBSON

CRC PRESS IEEE PRESS

A CRC Handbook Published in Cooperation with IEEE Press

Ford Motor Co.
Exhibit 1030
Page 094

Acquiring Editor: Nora Konopka
Project Editor: Carol Whitehead
Production Manager: Suzanne Lassandro
Marketing Manager: Barbara Glunn, Jane Lewis, Arline Massey, Jane Stark
Cover design: Jonathan Pennell
PrePress: Gary Bennett
Manufacturing: Carol Slatter

Library of Congress Cataloging-in-Publication Data

The mobile communications handbook / editor-in-chief, Jerry D. Gibson.
-- 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-8493-8597-0 (alk. paper)
1. Mobile communication systems. I. Gibson, Jerry D.
TK6570.M6M5934 1999
621.3845--dc21 98-46558
CIP

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

Neither this book nor any part may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, microfilming, and recording, or by any information storage or retrieval system, without prior permission in writing from the publisher.

All rights reserved. Authorization to photocopy items for internal or personal use, or the personal or internal use of specific clients, may be granted by CRC Press LLC, provided that $.50 per page photocopied is paid directly to Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923 USA. The fee code for users of the Transactional Reporting Service is ISBN 0-8493-8597-0/99/$0.00+$.50. The fee is subject to change without notice. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

The consent of CRC Press LLC does not extend to copying for general distribution, for promotion, for creating new works, or for resale. Specific permission must be obtained in writing from CRC Press LLC for such copying.

Direct all inquiries to CRC Press LLC, 2000 Corporate Blvd., N.W., Boca Raton, Florida 33431.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation, without intent to infringe.

© 1999 by CRC Press LLC

No claim to original U.S. Government works
International Standard Book Number 0-8493-8597-0
Library of Congress Card Number 98-46558
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

Preface

To keep pace with the rapid evolution of the mobile and wireless communications field, this second edition of *The Mobile Communications Handbook* contains 12 revised chapters and 10 entirely new chapters to complement another 14 chapters carried forward from the first edition. We use the term mobile communications to include technologies ranging from cordless telephones, digital cellular mobile radio, and evolving personal communications systems to wireless data and wireless networks. To cover this range of topics, we present 36 concise chapters by experts from industry and academia. The chapters are written to provide a succinct overview of each topic, quickly bringing the reader up to date; additionally, the chapters contain sufficient detail and references to encourage further study. The chapters are more than a "just the facts" presentation, with many of the authors using their experience to provide insights into forthcoming developments in the various fields.

The Handbook is divided into two parts. The first part covers the basic principles of analog and digital communications that pertain to mobile communications, and consists of 14 tutorial/review chapters that lay a solid groundwork for the wide-ranging aspects of mobile communications technologies. The second part of the Handbook contains 22 chapters covering such topics as cellular mobile radio, Rayleigh fading channels, space-time processing, power control, personal communications systems, user location and addressing, wireless data, wireless local area networks (LANs), wireless ATM, and third generation wireless standards. The basic principles chapters readily allow the reader to jump right to the mobile communications topic of interest, with the option of efficiently filling in any gaps in his/her background by referring back to the Basic Principles section as needed, without searching through a textbook or shuffling off to the library.

Although there is an ordering to the articles, the sequence of articles does not have to be read from the beginning to end. Each article was written to be an independent contribution but with intentional overlap between some articles to prevent the reader from having to flip back and forth to extract the desired information. Interestingly, as the reader will discover, this overlap often admits alternative views of difficult issues.

As in the first edition, the lead-off article for the wireless section is written by an acknowledged leader and innovator in wireless communications, Dr. Don Cox of Stanford University. This chapter is unique in that Dr. Cox inserts comments and provides updates directly in the text of his chapter from the first edition. Besides being informative and easy to read, the article implicitly challenges those active in wireless communications to continually question the common knowledge in the field.

It has been a great pleasure to work with the authors to publish this second edition. The authors come from all parts of the world and constitute a literal who's who of workers in digital and mobile communications. Certainly, each article is an extraordinary contribution to the communications field, and the collection, I believe, is the most comprehensive treatment of the mobile communications field available in one volume today. I sincerely acknowledge the exceptional efforts and talents of my wife, Elaine M. Gibson, who served as Managing Editor for the project, and was the principal contact point for the authors and the publisher. Her energy and organizational skills in coordinating each contribution and in organizing the Handbook were essential to its timely delivery and to its quality. Elaine

Ford Motor Co.
Exhibit 1030
Page 096

and I both appreciate the patience, support, and guidance of the staff at CRC Press during all stages of developing and producing this Handbook.

Jerry D. Gibson
Dallas, Texas

Ford Motor Co.
Exhibit 1030
Page 097

About the Editor

Jerry D. Gibson currently serves as Chairman of the Department of Electrical Engineering at Southern Methodist University in Dallas, Texas. He has held positions at General Dynamics-Fort Worth (1969–1972), the University of Notre Dame (1973–1974), and the University of Nebraska-Lincoln (1974–1976), and during the Fall of 1991, Dr. Gibson was on sabbatical with the Information Systems Laboratory and the Telecommunications Program in the Department of Electrical Engineering at Stanford University. From 1987–1997, he held the J. W. Runyon, Jr. Professorship in the Department of Electrical Engineering at Texas A&M University.

Dr. Gibson is coauthor of the book *Introduction to Nonparametric Detection with Applications* (Academic Press, 1975 and IEEE Press, 1995), the author of the textbook, *Principles of Digital and Analog Communications* (Prentice-Hall, second edition, 1993), and coauthor of the book *Digital Compression for Multimedia* (Morgan Kaufmann, 1998). He was Associate Editor for Speech Processing for the *IEEE Transactions on Communications* from 1981 to 1985 and Associate Editor for Communications for the *IEEE Transactions on Information Theory* from 1988–1991. He has served as a member of the Speech Technical Committee of the IEEE Signal Processing Society (1992–1995), on the Editorial Board for the *Proceedings of the IEEE* (1991–1997), and as a member of the IEEE Information Theory Society Board of Governors (1990-1998). He was President of the IEEE Information Theory Society in 1996. Dr. Gibson is Editor-in-Chief of *The Mobile Communications Handbook* (CRC Press, 1996) and Editor-in-Chief of *The Communications Handbook* (CRC Press, 1997).

In 1990, Dr. Gibson received The Fredrick Emmons Terman Award from The American Society for Engineering Education, and in 1992, was elected Fellow of the IEEE "for contributions to the theory and practice of adaptive prediction and speech waveform coding." He was corecipient of the 1993 IEEE Signal Processing Society Senior Paper Award for the Speech Processing area. His research interests include data, speech, image, and video compression, multimedia over networks, wireless communications, information theory, and digital signal processing.

Ford Motor Co.
Exhibit 1030
Page 098

Contributors

Melbourne Barton
Bellcore
Red Bank, New Jersey

V. K. Bhargava
University of Victoria
Department of Electrical and Computer Engineering
Victoria, Canada

Madhukar Budagavi
Texas Instruments
DSP Solutions R and D Center
Dallas, Texas

James J. Caffery, Jr.
Georgia Institute of Technology
School of Electrical and Computer Engineering
Atlanta, Georgia

Wai-Yip Chan
Electrical and Computer Engineering Department
Illinois Institute of Technology
Chicago, Illinois

Li Fung Chang
Bellcore
Morristown, New Jersey

Matthew Cheng
Bellcore
Morristown, New Jersey

Giovanni Cherubini
IBM Zurich Research Laboratory
Ruschlikon, Switzerland

Leon W. Couch, II
University of Florida
Electrical and Computer Engineering
Gainesville, Florida

Donald C. Cox
Stanford University
Stanford, California

Marc Delprat
Alcatel Telcom/MCD
Colombes, France

Spiros Dimolitsas
Lawrence Livermore National Laboratory
Livermore, California

I. J. Fair
Department of Electrical and Computer Engineering
Technical University of Nova Scotia
Halifax, Nova Scotia
Canada

Ira Gerson
Motorola Semiconductor Products
Schaumburg, Illinois

Steven D. Gray
Nokia Research Center
Irving, Texas

Lajos Hanzo
Department of Electrical and Computer Science
University of Southampton
Highfield, Southampton
United Kingdom

Tor Helleseth
Department of Informatics
University of Bergen
Bergen, Norway

Michael L. Honig
Northwestern University
Department of EECS
Evanston, Illinois

Hwei P. Hsu
Fairleigh Dickinson University
Teaneck, New Jersey

Bijan Jabbari
George Mason University
Department of Electrical and Computer Engineering
Fairfax, Virginia

Ravi Jain
Bell Communications Research
Red Bank, New Jersey

P. Vijay Kumar
University of Southern California
Los Angeles, California

Vinod Kumar
Alcatel Telecom
Mobile Communication Division
Colombes, France

Allen H. Levesque
GTE Laboratories
Waltham, Massachusetts

Yi-Bing Lin
National Chaio Tung University
Hsinchu, Taiwan
Republic of China

Ford Motor Co.
Exhibit 1030
Page 099

Joseph L. LoCicero
Illinois Institute of Technology
Armour College of Engineering
Department of Electrical and Computer Engineering
Chicago, Illinois

Paul Mermelstein
INRS—Telecommunications
Ile des Soeurs
Verdun, Quebec
Canada

Toshio Miki
Mobile Communication Network, Inc.
Yokosuka, Kanagawa
Japan

Laurence B. Milstein
University of California–San Diego
Department of Electrical and Computer Engineering
La Jolla, California

Seshadri Mohan
Bell Communications Research
Morristown, New Jersey

Tero Ojanperä
Nokia Research Center
Irving, Texas

Michael Onufry
COMSAT Laboratories
Clarksburg, Maryland

Geoffrey C. Orsak
Southern Methodist University
Electrical Engineering Department
Dallas, Texas

Kaveh Pahlavan
Worcester Polytechnic Institute
Worcester, Massachusetts

Bernd-Peter Paris
George Mason University
Department of Electrical and Computer Engineering
Fairfax, Virginia

Bhasker P. Patel
Illinois Institute of Technology
Chicago, Illinois

Arogyaswami J. Paulraj
Stanford University
Department of Electrical Engineering
Stanford, California

Roman Pichna
University of Oulu
Center for Wireless Communication
Oulu, Finland

John G. Proakis
Northeastern University
Electrical Engineering
Boston, Massachusetts

Bala Rajagopalan
NEC C&C Research Laboratories
Princeton, New Jersey

Daniel Reininger
NEC USA, Inc.
Princeton, New Jersey

Marvin K. Simon
Jet Propulsion Laboratory
Pasadena, California

Suresh Singh
Oregon State University
Electrical and Computer Engineering Department
Corvallis, Oregon

Bernard Sklar
Communications Engineering Services
Tarzana, California

Raymond Steele
MAC, Ltd.
Southampton
United Kingdom

Gordon L. Stüber
Georgia Institute of Technology
School of Electrical and Computer Engineering
Atlanta, Georgia

Raj Talluri
Texas Instruments
DSP Solutions R and D Center
Dallas, Texas

Qiang Wang
University of Victoria
Victoria, British Columbia
Canada

Michel Daoud Yacoub
The University of Campinas
San Paulo, Brazil

Ford Motor Co.
Exhibit 1030
Page 100

Contents

SECTION I Basic Principles

1 Complex Envelope Representations for Modulated Signals *Leon W. Couch, II* 1-1

2 Sampling *Hwei P. Hsu* 2-1

3 Pulse Code Modulation *Leon W. Couch, II* 3-1

4 Baseband Signalling and Pulse Shaping *Michael L. Honig and Melbourne Barton* 4-1

5 Channel Equalization *John G. Proakis* 5-1

6 Line Coding *Joseph L. LoCicero and Bhasker P. Patel* 6-1

7 Echo Cancellation *Giovanni Cherubini* 7-1

8 Pseudonoise Sequences *Tor Helleseth and P. Vijay Kumar* 8-1

9 Optimum Receivers *Geoffrey C. Orsak* 9-1

10 Forward Error Correction Coding *V.K. Bhargava and I.J. Fair* 10-1

11 Spread Spectrum Communications *Laurence B. Milstein and Marvin K. Simon* 11-1

12 Diversity *Arogyaswami J. Paulraj* 12-1

13 Digital Communication System Performance *Bernard Sklar* 13-1

14 Telecommunications Standardization *Spiros Dimolitsas and Michael Onufry* 14-1

SECTION II Wireless

15 Wireless Personal Communications: A Perspective *Donald C. Cox* 15-1

16 Modulation Methods *Gordon L. Stüber* 16-1

17 Access Methods *Bernd-Peter Paris* 17-1

18 Rayleigh Fading Channels *Bernard Sklar* 18-1

Ford Motor Co.
Exhibit 1030
Page 101

19 Space-Time Processing *Arogyaswami J. Paulraj* 19-1

20 Location Strategies for Personal Communications Services *Ravi Jain, Yi-Bing Lin, and Seshadri Mohan* 20-1

21 Cell Design Principles *Michel Daoud Yacoub* 21-1

22 Microcellular Radio Communications *Raymond Steele* 22-1

23 Fixed and Dynamic Channel Assignment *Bijan Jabbari* 23-1

24 Radiolocation Techniques *Gordon L. Stüber and James J. Caffery, Jr.* 24-1

25 Power Control *Roman Pichna and Qiang Wang* 25-1

26 Enhancements in Second Generation Systems *Marc Delprat and Vinod Kumar* 26-1

27 The Pan-European Cellular System *Lajos Hanzo* 27-1

28 Speech and Channel Coding for North American TDMA Cellular Systems *Paul Mermelstein* 28-1

29 The British Cordless Telephone Standard: CT-2 *Lajos Hanzo* 29-1

30 Half-Rate Standards *Wai-Yip Chan, Ira Gerson, and Toshio Miki* 30-1

31 Wireless Video Communications *Madhukar Budagavi and Raj Talluri* . . . 31-1

32 Wireless LANs *Suresh Singh* 32-1

33 Wireless Data *Allen H. Levesque and Kaveh Pahlavan* 33-1

34 Wireless ATM: Interworking Aspects *Melbourne Barton, Matthew Cheng, and Li Fung Chang* 34-1

35 Wireless ATM: QoS and Mobility Management *Bala Rajagopalan and Daniel Reininger* 35-1

36 An Overview of cdma2000, WCDMA, and EDGE *Tero Ojanperä and Steven D. Gray* 36-1

Index I-1

Ford Motor Co.
Exhibit 1030
Page 102

I

Basic Principles

1 **Complex Envelope Representations for Modulated Signals** *Leon W. Couch, II* . 1-1
Introduction • Complex Envelope Representation • Representation of Modulated Signals • Generalized Transmitters and Receivers • Spectrum and Power of Bandpass Signals • Amplitude Modulation • Phase and Frequency Modulation • QPSK Signalling

2 **Sampling** *Hwei P. Hsu* 2-1
Introduction • Instantaneous Sampling • Sampling Theorem • Sampling of Sinusoidal Signals • Sampling of Bandpass Signals • Practical Sampling • Sampling Theorem in the Frequency Domain • Summary and Discussion

3 **Pulse Code Modulation** *Leon W. Couch, II* 3-1
Introduction • Generation of PCM • Percent Quantizing Noise • Practical PCM Circuits • Bandwidth of PCM • Effects of Noise • Nonuniform Quantizing: μ-Law and A-Law Companding • Example: Design of a PCM System

4 **Baseband Signalling and Pulse Shaping** *Michael L. Honig and Melbourne Barton* 4-1
Communications System Model • Intersymbol Interference and the Nyquist Criterion • Nyquist Criterion with Matched Filtering • Eye Diagrams • Partial-Response Signalling • Additional Considerations • Examples

5 **Channel Equalization** *John G. Proakis* 5-1
Characterization of Channel Distortion • Characterization of Intersymbol Interference • Linear Equalizers • Decision-Feedback Equalizer • Maximum-Likelihood Sequence Detection • Conclusions

6 **Line Coding** *Joseph L. LoCicero and Bhasker P. Patel* 6-1
Introduction • Common Line Coding Formats • Alternate Line Codes • Multilevel Signalling, Partial Response Signalling, and Duobinary Coding • Bandwidth Comparison • Concluding Remarks

7 **Echo Cancellation** *Giovanni Cherubini* 7-1
Introduction • Echo Cancellation for PAM Systems • Echo Cancellation for QAM Systems • Echo Cancellation for OFDM Systems • Summary and Conclusions

8 **Pseudonoise Sequences** *Tor Helleseth and P. Vijay Kumar* 8-1
Introduction • m Sequences • The q-ary Sequences with Low Autocorrelation • Families of Sequences with Low Crosscorrelation • Aperiodic Correlation • Other Correlation Measures

Ford Motor Co.
Exhibit 1030
Page 103

9 Optimum Receivers *Geoffrey C. Orsak* 9-1
Introduction • Preliminaries • Karhunen–Loève Expansion • Detection Theory • Performance • Signal Space • Standard Binary Signalling Schemes • *M*-ary Optimal Receivers • More Realistic Channels • Dispersive Channels

10 Forward Error Correction Coding *V.K. Bhargava and I.J. Fair* 10-1
Introduction • Fundamentals of Block Coding • Structure and Decoding of Block Codes • Important Classes of Block Codes • Principles of Convolutional Coding • Decoding of Convolutional Codes • Trellis-Coded Modulation • Additional Measures • Turbo Codes • Applications

11 Spread Spectrum Communications *Laurence B. Milstein and Marvin K. Simon* 11-1
A Brief History • Why Spread Spectrum? • Basic Concepts and Terminology • Spread Spectrum Techniques • Applications of Spread Spectrum

12 Diversity *Arogyaswami J. Paulraj* 12-1
Introduction • Diversity Schemes • Diversity Combining Techniques • Effect of Diversity Combining on Bit Error Rate • Concluding Remarks

13 Digital Communication System Performance *Bernard Sklar* 13-1
Introduction • Bandwidth and Power Considerations • Example 1: Bandwidth-Limited Uncoded System • Example 2: Power-Limited Uncoded System • Example 3: Bandwidth-Limited Power-Limited Coded System • Example 4: Direct-Sequence Spread-Spectrum Coded System • Conclusion

14 Telecommunications Standardization *Spiros Dimolitsas and Michael Onufry* .. 14-1
Introduction • Global Standardization • Regional Standardization • National Standardization • Intellectual Property • Standards Coordination • Scientific • Standards Development Cycle

Ford Motor Co.
Exhibit 1030
Page 104

1

Complex Envelope Representations for Modulated Signals[1]

1.1 Introduction 1-1
1.2 Complex Envelope Representation 1-2
1.3 Representation of Modulated Signals 1-4
1.4 Generalized Transmitters and Receivers 1-4
1.5 Spectrum and Power of Bandpass Signals 1-7
1.6 Amplitude Modulation 1-8
1.7 Phase and Frequency Modulation 1-10
1.8 QPSK Signalling 1-11
Defining Terms 1-14
References 1-15
Further Information 1-15

Leon W. Couch, II
University of Florida

1.1 Introduction

What is a general representation for bandpass digital and analog signals? How do we represent a **modulated signal**? How do we evaluate the spectrum and the power of these signals? These are some of the questions that are answered in this chapter.

A *baseband* waveform has a spectral magnitude that is nonzero for frequencies in the vicinity of the origin (i.e., $f = 0$) and negligible elsewhere. A *bandpass* waveform has a spectral magnitude that is nonzero for frequencies in some band concentrated about a frequency $f = \pm f_c$ (where $f_c \gg 0$), and the spectral magnitude is negligible elsewhere. f_c is called the *carrier frequency*. The value of f_c may be arbitrarily assigned for mathematical convenience in some problems. In others, namely, **modulation** problems, f_c is the frequency of an oscillatory signal in the transmitter circuit and is the assigned frequency of the transmitter, such as 850 kHz for an AM broadcasting station.

In communication problems, the information source signal is usually a baseband signal—for example, a transistor-transistor logic (TTL) waveform from a digital circuit or an audio (analog) signal from a microphone. The communication engineer has the job of building a system that will transfer the information from this source signal to the desired destination.

[1] *Source:* Couch, Leon W., II. 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ.

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 105

As shown in Fig. 1.1, this usually requires the use of a bandpass signal, $s(t)$, which has a bandpass spectrum that is concentrated at $\pm f_c$ where f_c is selected so that $s(t)$ will propagate across the communication channel (either a wire or a wireless channel).

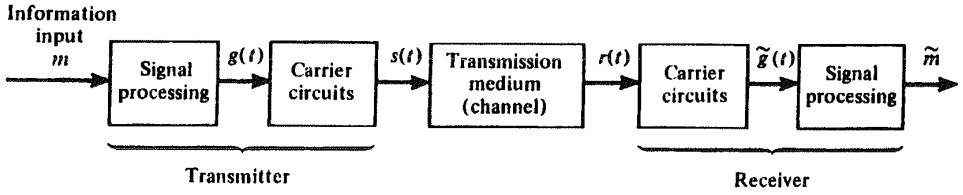


FIGURE 1.1 Bandpass communication system. *Source:* Couch, L.W., II. 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 227. With permission.

Modulation is the process of imparting the source information onto a bandpass signal with a carrier frequency f_c by the introduction of amplitude and/or phase perturbations. This bandpass signal is called the *modulated* signal $s(t)$, and the baseband source signal is called the *modulating* signal $m(t)$. Examples of exactly how modulation is accomplished are given later in this chapter. This definition indicates that modulation may be visualized as a mapping operation that maps the source information onto the bandpass signal $s(t)$ that will be transmitted over the channel.

As the modulated signal passes through the channel, noise corrupts it. The result is a bandpass signal-plus-noise waveform that is available at the receiver input, $r(t)$, as illustrated in Fig. 1.1. The receiver has the job of trying to recover the information that was sent from the source; $\tilde{m}$ denotes the corrupted version of m.

1.2 Complex Envelope Representation

All bandpass waveforms, whether they arise from a modulated signal, interfering signals, or noise, may be represented in a convenient form given by the following theorem. $v(t)$ will be used to denote the **bandpass waveform** canonically. That is, $v(t)$ can represent the signal when $s(t) \equiv v(t)$, the noise when $n(t) \equiv v(t)$, the filtered signal plus noise at the channel output when $r(t) \equiv v(t)$, or any other type of bandpass waveform[2].

THEOREM 1.1 *Any physical bandpass waveform can be represented by*

$$v(t) = Re\left\{g(t)e^{j\omega_c t}\right\} \tag{1.1a}$$

$Re\{\cdot\}$ denotes the real part of $\{\cdot\}$. $g(t)$ is called the complex envelope of $v(t)$, and f_c is the associated carrier frequency (hertz) where $\omega_c = 2\pi f_c$. Furthermore, two other equivalent representations are

[2]The symbol $\equiv$ denotes an equivalence and the symbol $\triangleq$ denotes a definition.

Ford Motor Co.
Exhibit 1030
Page 106

$$v(t) = R(t)\cos[\omega_c t + \theta(t)] \tag{1.1b}$$

and

$$v(t) = x(t)\cos\omega_c t - y(t)\sin\omega_c t \tag{1.1c}$$

where

$$g(t) = x(t) + jy(t) = |g(t)|e^{j\angle g(t)} \equiv R(t)e^{j\theta(t)} \tag{1.2}$$

$$x(t) = Re\{g(t)\} \equiv R(t)\cos\theta(t) \tag{1.3a}$$
$$y(x) = Im\{g(t)\} \equiv R(t)\sin\theta(t) \tag{1.3b}$$

$$R(t) \triangleq |g(t)| \equiv \sqrt{x^2(t) + y^2(t)} \tag{1.4a}$$
$$\theta(t) \triangleq \angle g(t) = \tan^{-1}\left(\frac{y(t)}{x(t)}\right) \tag{1.4b}$$

The waveforms $g(t), x(t), y(t), R(t)$, and $\theta(t)$ are all **baseband waveforms**, and, except for $g(t)$, they are all real waveforms. $R(t)$ is a nonnegative real waveform. Equation (1.1) is a low-pass-to-bandpass transformation. The $e^{j\omega_c t}$ factor in (1.1a) shifts (i.e., translates) the spectrum of the baseband signal $g(t)$ from baseband up to the carrier frequency f_c. In communications terminology the frequencies in the baseband signal $g(t)$ are said to be *heterodyned* up to f_c. The **complex envelope**, $g(t)$, is usually a complex function of time and it is the generalization of the phasor concept. That is, if $g(t)$ happens to be a complex constant, then $v(t)$ is a pure sine wave of frequency f_c and this complex constant is the phasor representing the sine wave. If $g(t)$ is not a constant, then $v(t)$ is not a pure sine wave because the amplitude and phase of $v(t)$ varies with time, caused by the variations of $g(t)$.

Representing the complex envelope in terms of two real functions in Cartesian coordinates, we have

$$g(x) \equiv x(t) + jy(t) \tag{1.5}$$

where $x(t) = \text{Re}\{g(t)\}$ and $y(t) = \text{Im}\{g(t)\}$. $x(t)$ is said to be the *in-phase modulation* associated with $v(t)$, and $y(t)$ is said to be the *quadrature modulation* associated with $v(t)$. Alternatively, the polar form of $g(t)$, represented by $R(t)$ and $\theta(t)$, is given by (1.2), where the identities between Cartesian and polar coordinates are given by (1.3) and (1.4). $R(t)$ and $\theta(t)$ are real waveforms and, in addition, $R(t)$ is always nonnegative. $R(t)$ is said to be the *amplitude modulation* (AM) on $v(t)$, and $\theta(t)$ is said to be the *phase modulation* (PM) on $v(t)$.

The usefulness of the complex envelope representation for bandpass waveforms cannot be overemphasized. In modern communication systems, the bandpass signal is often partitioned into two channels, one for $x(t)$ called the I (in-phase) channel and one for $y(t)$ called the Q (quadrature-phase) channel. In digital computer simulations of bandpass signals, the sampling rate used in the simulation can be minimized by working with the complex envelope, $g(t)$, instead of with the bandpass signal, $v(t)$, because $g(t)$ is the baseband equivalent of the bandpass signal [1].

Ford Motor Co.
Exhibit 1030
Page 107

1.3 Representation of Modulated Signals

Modulation is the process of encoding the source information $m(t)$ (modulating signal) into a bandpass signal $s(t)$ (modulated signal). Consequently, the modulated signal is just a special application of the bandpass representation. The *modulated signal* is given by

$$s(t) = \mathrm{Re}\left\{g(t)e^{j\omega_c t}\right\} \tag{1.6}$$

where $\omega_c = 2\pi f_c$. f_c is the carrier frequency. The complex envelope $g(t)$ is a function of the modulating signal $m(t)$. That is,

$$g(t) = g[m(t)] \tag{1.7}$$

Thus $g[\cdot]$ performs a mapping operation on $m(t)$. This was shown in Fig. 1.1.

Table 1.1 gives an overview of the *big picture* for the modulation problem. Examples of the mapping function $g[m]$ are given for amplitude modulation (AM), double-sideband suppressed carrier (DSB-SC), phase modulation (PM), frequency modulation (FM), single-sideband AM suppressed carrier (SSB-AM-SC), single-sideband PM (SSB-PM), single-sideband FM (SSB-FM), single-sideband envelope detectable (SSB-EV), single-sideband square-law detectable (SSB-SQ), and quadrature modulation (QM). For each $g[m]$, Table 1.1 also shows the corresponding $x(t)$ and $y(t)$ quadrature modulation components, and the corresponding $R(t)$ and $\theta(t)$ amplitude and phase modulation components. Digitally modulated bandpass signals are obtained when $m(t)$ is a digital baseband signal—for example, the output of a transistor transistor logic (TTL) circuit.

Obviously, it is possible to use other $g[m]$ functions that are not listed in Table 1.1. The question is: Are they useful? $g[m]$ functions are desired that are easy to implement and that will give desirable spectral properties. Furthermore, in the receiver the inverse function $m[g]$ is required. The inverse should be single valued over the range used and should be easily implemented. The inverse mapping should suppress as much noise as possible so that $m(t)$ can be recovered with little corruption.

1.4 Generalized Transmitters and Receivers

A more detailed description of transmitters and receivers as first shown in Fig. 1.1 will now be illustrated.

There are two canonical forms for the generalized transmitter, as indicated by (1.1b) and (1.1c). Equation (1.1b) describes an AM-PM type circuit as shown in Fig. 1.2. The baseband signal processing circuit generates $R(t)$ and $\theta(t)$ from $m(t)$. The R and θ are functions of the modulating signal $m(t)$, as given in Table 1.1, for the particular modulation type desired. The signal processing may be implemented either by using nonlinear analog circuits or a digital computer that incorporates the R and θ algorithms under software program control. In the implementation using a digital computer, one analog-to-digital converter (ADC) will be needed at the input of the baseband signal processor and two digital-to-analog converters (DACs) will be needed at the output. The remainder of the AM-PM canonical form requires radio frequency (RF) circuits, as indicated in the figure.

Figure 1.3 illustrates the second canonical form for the generalized transmitter. This uses in-phase and quadrature-phase (IQ) processing. Similarly, the formulas relating $x(t)$ and $y(t)$ to $m(t)$ are shown in Table 1.1, and the baseband signal processing may be implemented by using either analog hardware or digital hardware with software. The remainder of the canonical form uses RF circuits as indicated.

Ford Motor Co.
Exhibit 1030
Page 108

TABLE 1.1 Complex Envelope Functions for Various Types of Modulation[a]

Type of Modulation	Mapping Functions $g(m)$	Corresponding Quadrature Modulation	
		$x(t)$	$y(t)$
AM	$A_c[1 + m(t)]$	$A_c[1 + m(t)]$	0
DSB-SC	$A_c m(t)$	$A_c m(t)$	0
PM	$A_c e^{jD_p m(t)}$	$A_c \cos[D_p m(t)]$	$A_c \sin[D_p m(t)]$
FM	$A_c e^{jD_f \int_{-\infty}^{t} m(\sigma)d\sigma}$	$A_c \cos\left[D_f \int_{-\infty}^{t} m(\sigma)d\sigma\right]$	$A_c \sin\left[D_f \int_{-\infty}^{t} m(\sigma)d\sigma\right]$
SSB-AM-SC[b]	$A_c[m(t) \pm j\hat{m}(t)]$	$A_c m(t)$	$\pm A_c \hat{m}(t)$
SSB-PM[b]	$A_c e^{jD_p[m(t) \pm j\hat{m}(t)]}$	$A_c e^{\mp D_p \hat{m}(t)} \cos[D_p m(t)]$	$A_c e^{\mp D_p \hat{m}(t)} \sin[D_p m(t)]$
SSB-FM[b]	$A_c e^{jD_f \int_{-\infty}^{t} [m(\sigma) \pm j\hat{m}(\sigma)]d\sigma}$	$A_c e^{\mp D_f \int_{-\infty}^{t} \hat{m}(\sigma)d\sigma} \cos\left[D_f \int_{-\infty}^{t} m(\sigma)d\sigma\right]$	$A_c e^{\mp D_f \int_{-\infty}^{t} \hat{m}(\sigma)d\sigma} \sin\left[D_f \int_{-\infty}^{t} m(\sigma)d\sigma\right]$
SSB-EV[b]	$A_c e^{\{\ln[1+m(t)] \pm j\hat{\ln}[1+m(t)]\}}$	$A_c[1 + m(t)] \cos\{\hat{\ln}[1 + m(t)]\}$	$\pm A_c[1 + m(t)] \sin\{\hat{\ln}[1 + m(t)]\}$
SSB-SQ[b]	$A_c e^{(1/2)\{\ln[1+m(t)] \pm j\hat{\ln}[1+m(t)]\}}$	$A_c\sqrt{1 + m(t)} \cos\{\frac{1}{2}\hat{\ln}[1 + m(t)]\}$	$\pm A_c\sqrt{1 + m(t)} \sin\{\frac{1}{2}\hat{\ln}[1 + m(t)]\}$
QM	$A_c[m_1(t) + jm_2(t)]$	$A_c m_1(t)$	$A_c m_2(t)$

Ford Motor Co.
Exhibit 1030
Page 109

TABLE 1.1 Complex Envelope Functions for Various Types of Modulation[a] *(Continued)*

Type of Modulation	Corresponding Amplitude and Phase Modulation		Linearity	Remarks
	$R(t)$	$\theta(t)$		
AM	$A_c\|1+m(t)\|$	$\begin{cases} 0, & m(t)>-1 \\ 180°, & m(t)<-1 \end{cases}$	L[c]	$m(t)>-1$ required for envelope detection
DSB-SC	$A_c\|m(t)\|$	$\begin{cases} 0, & m(t)>0 \\ 180°, & m(t)<0 \end{cases}$	L	Coherent detection required
PM	A_c	$D_p m(t)$	NL	D_p is the phase deviation constant (rad/volt)
FM	A_c	$D_f\int_{-\infty}^{t} m(\sigma)d\sigma$	NL	D_f is the frequency deviation constant (rad/volt-sec)
SSB-AM-SC[b]	$A_c\sqrt{[m(t)]^2+[\hat{m}(t)]^2}$	$\tan^{-1}[\pm\hat{m}(t)/m(t)]$	L	Coherent detection required
SSB-PM[b]	$A_c e^{\pm D_p\hat{m}(t)}$	$D_p m(t)$	NL	
SSB-FM[b]	$A_c e^{\pm D_f\int_{-\infty}^{t}\hat{m}(\sigma)d\sigma}$	$D_f\int_{-\infty}^{t} m(\sigma)d\sigma$	NL	
SSB-EV[b]	$A_c\|1+m(t)\|$	$\pm\hat{\ln}[1+m(t)]$	NL	$m(t)>-1$ is required so that the $\ln(\cdot)$ will have a real value
SSB-SQ[b]	$A_c\sqrt{1+m(t)}$	$\pm\frac{1}{2}\hat{\ln}[1+m(t)]$	NL	$m(t)>-1$ is required so that the $\ln(\cdot)$ will have a real value
QM	$A_c\sqrt{m_1^2(t)+m_2^2(t)}$	$\tan^{-1}[m_2(t)/m_1(t)]$	L	Used in NTSC color television; requires coherent detection

Source: Couch, L.W., II, 1997, *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, pp. 231-232. With permission.

[a] $A_c>0$ is a constant that sets the power level of the signal as evaluated by use of (1.11); L, linear; NL, nonlinear; and $\hat{[\cdot]}$ is the Hilbert transform (a $-90°$ phase-shifted version of $[\cdot]$). For example, $\hat{m}(t)=m(t)*\frac{1}{\pi t}=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{m(\lambda)}{t-\lambda}d\lambda$.

[b] Use upper signs for upper sideband signals and lower signals for lower sideband signals.

[c] In the strict sense, AM signals are not linear because the carrier term does not satisfy the linearity (superposition) condition.

Analogous to the transmitter realizations, there are two canonical forms of receiver. Each one consists of RF carrier circuits followed by baseband signal processing as illustrated in Fig. 1.1. Typically the carrier circuits are of the superheterodyne-receiver type which consist of an RF amplifier, a down converter (mixer plus local oscillator) to some intermediate frequency (IF), an IF amplifier and then detector circuits [1]. In the first canonical form of the receiver, the carrier circuits have amplitude and phase detectors that output $\tilde{R}(t)$ and $\tilde{\theta}(t)$, respectively. This pair, $\tilde{R}(t)$ and $\tilde{\theta}(t)$, describe the polar form of the received complex envelope, $\tilde{g}(t)$. $\tilde{R}(t)$ and $\tilde{\theta}(t)$ are then fed into the signal processor which uses the inverse functions of Table 1.1 to generate the recovered modulation, $\tilde{m}(t)$. The second canonical form of the receiver uses quadrature product detectors in the carrier circuits to produce the Cartesian form of the received complex envelope, $\tilde{x}(t)$ and $\tilde{y}(t)$. $\tilde{x}(t)$ and $\tilde{y}(t)$ are then inputted to the signal processor which generates $\tilde{m}(t)$ at its output.

Once again, it is stressed that any type of signal modulation (see Table 1.1) may be generated (transmitted) or detected (received) by using either of these two canonical forms. Both of these forms conveniently separate baseband processing from RF processing. Digital

Ford Motor Co.
Exhibit 1030
Page 110

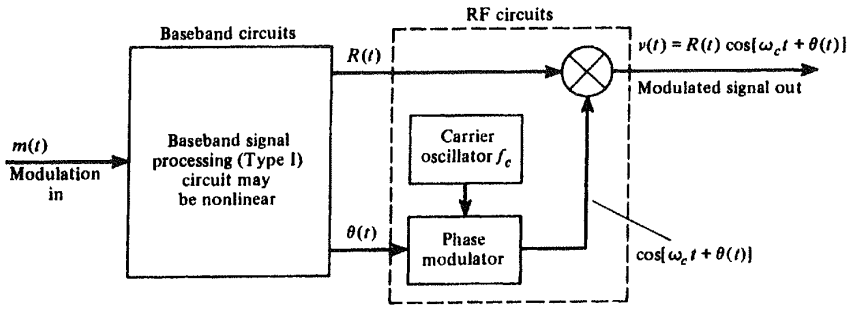


FIGURE 1.2 Generalized transmitter using the AM-PM generation technique. *Source:* Couch, L.W., II. 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 278. With permission.

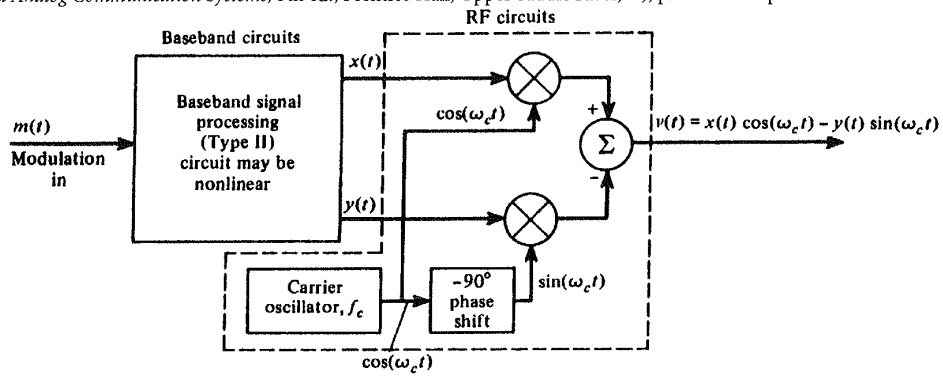


FIGURE 1.3 Generalized transmitter using the quadrature generation technique. *Source:* Couch, L.W., II. 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 278. With permission.

techniques are especially useful to realize the baseband processing portion. Furthermore, if digital computing circuits are used, any desired modulation type can be realized by selecting the appropriate software algorithm.

1.5 Spectrum and Power of Bandpass Signals

The spectrum of the bandpass signal is the translation of the spectrum of its complex envelope. Taking the **Fourier transform** of (1.1a), the spectrum of the bandpass waveform is [1]

$$V(f) = \frac{1}{2}\left[G\left(f - f_c\right) + G^*\left(-f - f_c\right)\right] \tag{1.8}$$

where $G(f)$ is the Fourier transform of $g(t)$,

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-j2\pi ft}dt \ ,$$

and the asterisk superscript denotes the complex conjugate operation. The *power spectra density* (PSD) of the bandpass waveform is [1]

$$\mathcal{P}_v(f) = \frac{1}{4}\left[\mathcal{P}_g\left(f - f_c\right) + \mathcal{P}_g\left(-f - f_c\right)\right] \tag{1.9}$$

where $\mathcal{P}_g(f)$ is the PSD of $g(t)$.

The average power dissipated in a resistive load is V_{rms}^2/R_L or $I_{\text{rms}}^2 R_L$ where V_{rms} is the rms value of the voltage waveform across the load and I_{rms} is the rms value of the current through the load. For bandpass waveforms, Eq. (1.1) may represent either the voltage or the current. Furthermore, the rms values of $v(t)$ and $g(t)$ are related by [1]

$$v_{\text{rms}}^2 = \langle v^2(t) \rangle = \frac{1}{2} \langle |g(t)|^2 \rangle = \frac{1}{2} g_{\text{rms}}^2 \tag{1.10}$$

where $\langle \cdot \rangle$ denotes the time average and is given by

$$\left\langle \left[\quad \right] \right\rangle = \lim_{t \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} \left[\quad \right] dt$$

Thus, if $v(t)$ of (1.1) represents the bandpass voltage waveform across a resistive load, the average power dissipated in the load is

$$P_L = \frac{v_{\text{rms}}^2}{R_L} = \frac{\langle v^2(t) \rangle}{R_L} = \frac{\langle |g(t)|^2 \rangle}{2R_L} = \frac{g_{\text{rms}}^2}{2R_L} \tag{1.11}$$

where g_{rms} is the rms value of the complex envelope and R_L is the resistance of the load.

1.6 Amplitude Modulation

Amplitude modulation (AM) will now be examined in more detail. From Table 1.1 the complex envelope of an AM signal is

$$g(t) = A_c[1 + m(t)] \tag{1.12}$$

so that the spectrum of the complex envelope is

$$G(f) = A_c \delta(f) + A_c M(f) \tag{1.13}$$

Using (1.6), we obtain the AM signal waveform

$$s(t) = A_c[1 + m(t)] \cos \omega_c t \tag{1.14}$$

and, using (1.8), the AM spectrum

$$S(f) = \frac{1}{2} A_c \left[\delta(f - f_c) + M(f - f_c) + \delta(f + f_c) + M(f + f_c) \right] \tag{1.15}$$

where $\delta(f) = \delta(-f)$ and, because $m(t)$ is real, $M^*(f) = M(-f)$. Suppose that the magnitude spectrum of the modulation happens to be a triangular function, as shown in Fig. 1.4(a). This spectrum might arise from an analog audio source where the bass frequencies are emphasized. The resulting AM spectrum, using (1.15), is shown in Fig. 1.4(b). Note that because $G(f - f_c)$ and $G^*(-f - f_c)$ do not overlap, the magnitude spectrum is

$$|S(f)| = \begin{cases} \frac{1}{2} A_c \delta(f - f_c) + \frac{1}{2} A_c |M(f - f_c)|, & f > 0 \\ \frac{1}{2} A_c \delta(f + f)_c + \frac{1}{2} A_c |M(-f - f)_c|, & f < 0 \end{cases} \tag{1.16}$$

The 1 in

$$g(t) = A_c[1 + m(t)]$$

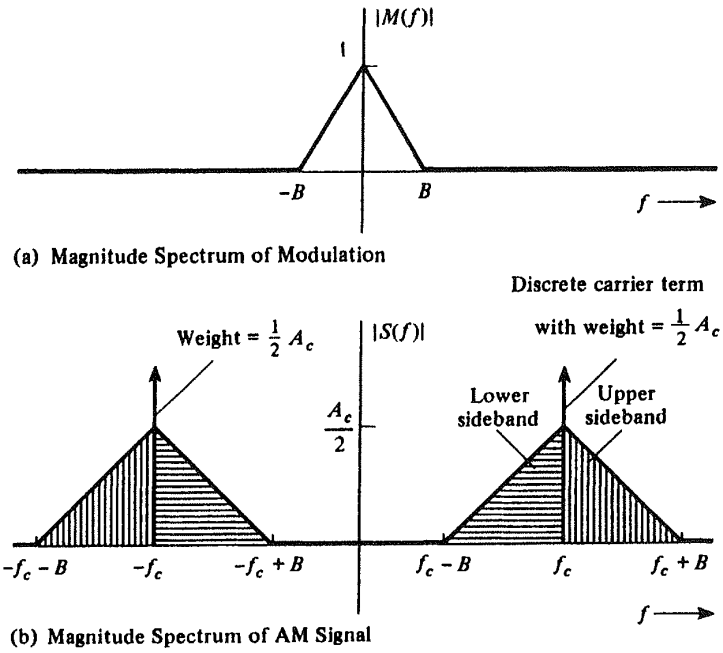


FIGURE 1.4 Spectrum of an AM signal. *Source:* Couch, L.W., II. 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 235. With permission.

causes delta functions to occur in the spectrum at $f = \pm f_c$, where f_c is the assigned carrier frequency. Also, from Fig. 1.4 and (1.16), it is realized that the bandwidth of the AM signal is 2B. That is, the bandwidth of the AM signal is twice the bandwidth of the baseband modulating signal.

The average power dissipated into a resistive load is found by using (1.11).

$$P_L = \frac{A_c^2}{2R_L}\left\langle |1 + m(t)|^2 \right\rangle = \frac{A_c^2}{2R_L}\left[1 + 2\langle m(t)\rangle + \langle m^2(t)\rangle\right]$$

If we assume that the *dc* value of the modulation is zero, $\langle m(t)\rangle = 0$, then the average power dissipated into the load is

$$P_L = \frac{A_c^2}{2R_L}\left\langle 1 + m_{\text{rms}}^2 \right\rangle \tag{1.17}$$

where m_{rms} is the rms value of the modulation, $m(t)$. Thus, the average power of an AM signal changes if the rms value of the modulating signal changes. For example, if $m(t)$ is a sine wave test tone with a peak value of 1.0 for 100% modulation,

$$m_{\text{rms}} = 1/\sqrt{2}\ .$$

Assume that $A_c = 1000$ volts and $R_L = 50$ ohms, which are typical values used in AM broadcasting. Then the averagepower dissipated into the 50 Ω load for this AM signal is

$$P_L = \frac{(1000)^2}{2(50)}\left[1 + \frac{1}{2}\right] = 15{,}000 \text{ watts} \tag{1.18}$$

The Federal Communications Commission (FCC) rated carrier power is obtained when $m(t) = 0$. In this case, (1.17) becomes $P_L = (1000)^2/100 = 10{,}000$ watts and the FCC would rate this as a 10,000 watt AM station. The sideband power for 100% sine wave modulation is 5,000 watts.

Now let the modulation on the AM signal be a binary digital signal such that $m(t) = \pm 1$ where $+1$ is used for a binary one and -1 is used for a binary 0. Referring to (1.14), this AM signal becomes an *on-off keyed* (OOK) digital signal where the signal is on when a binary one is transmitted and off when a binary zero is transmitted. For $A_c = 1000$ and $R_L = 50\ \Omega$, the average power dissipated would be 20,000 watts since $m_{\text{rms}} = 1$ for $m(t) = \pm 1$.

1.7 Phase and Frequency Modulation

Phase modulation (PM) and *frequency modulation* (FM) are special cases of angle-modulated signalling. In angle-modulated signalling the complex envelope is

$$g(t) = A_c e^{j\theta(t)} \tag{1.19}$$

Using (1.6), the resulting *angle-modulated* signal is

$$s(t) = A_c \cos\left[\omega_c + \theta(t)\right] \tag{1.20}$$

For PM the phase is directly proportional to the modulating signal:

$$\theta(t) = D_p m(t) \tag{1.21}$$

where the proportionality constant D_p is the phase sensitivity of the phase modulator, having units of radians per volt [assuming that $m(t)$ is a voltage waveform]. For FM the phase is proportional to the integral of $m(t)$:

$$\theta(t) = D_f \int_{-\infty}^{t} m(\sigma) d\sigma \tag{1.22}$$

where the frequency deviation constant D_f has units of radians/volt-second. These concepts are summarized by the PM and FM entries in Table 1.1.

By comparing the last two equations, it is seen that if we have a PM signal modulated by $m_p(t)$, there is *also* FM on the signal corresponding to a *different* modulating waveshape that is given by

$$m_f(t) = \frac{D_p}{D_f}\left[\frac{dm_p(t)}{dt}\right] \tag{1.23}$$

where the subscripts f and p denote frequency and phase, respectively. Similarly, if we have an FM signal modulated by $m_f(t)$, the corresponding phase modulation on this signal is

$$m_p(t) = \frac{D_f}{D_p} \int_{-\infty}^{t} m_f(\sigma) d\sigma \tag{1.24}$$

By using (1.24), a PM circuit may be used to synthesize an FM circuit by inserting an integrator in cascade with the phase modulator input.

Other properties of PM and FM are that the **real envelope**, $R(t) = |g(t)| = A_c$, is a constant, as seen from (1.19). Also, $g(t)$ is a *nonlinear* function of the modulation. However,

Ford Motor Co.
Exhibit 1030
Page 114

from (1.21) and (1.22), $\theta(t)$ is a linear function of the modulation, $m(t)$. Using (1.11), the average power dissipated by a PM or FM signal is the constant

$$P_L = \frac{A_c^2}{2R_L} \tag{1.25}$$

That is, the average power of a PM or FM signal does not depend on the modulating waveform, $m(t)$.

The *instantaneous frequency deviation* for an FM signal from its carrier frequency is given by the derivative of its phase $\theta(t)$. Taking the derivative of (1.22), the *peak frequency deviation* is

$$\Delta F = \frac{1}{2\pi} D_f M_p \text{ Hz} \tag{1.26}$$

where $M_p = \max[m(t)]$ is the peak value of the modulation waveform and the derivative has been divided by 2π to convert from radians/sec to Hz units.

For FM and PM signals, Carson's rule estimates the transmission bandwidth containing approximately 98% of the total power. This FM or PM signal bandwidth is

$$B_T = 2(\beta + 1)B \tag{1.27}$$

where B is bandwidth (highest frequency) of the modulation. The modulation index β, is $\beta = \Delta F/B$ for FM and $\beta = \max[D_p m(t)] = D_p M_p$ for PM.

The AMPS (Advanced Mobile Phone System) analog cellular phones use FM signalling. A peak deviation of 12 kHz is specified with a modulation bandwidth of 3 kHz. From (1.27), this gives a bandwidth of 30 kHz for the AMPS signal and allows a channel spacing of 30 kHz to be used. To accommodate more users, narrow-band AMPS (NAMPS) with a 5 kHz peak deviation is used in some areas. This allows 10 kHz channel spacing if the carrier frequencies are carefully selected to minimize interference to used adjacent channels. A maximum FM signal power of 3 watts is allowed for the AMPS phones. However, hand-held AMPS phones usually produce no more than 600 mW which is equivalent to 5.5 volts rms across the 50 Ω antenna terminals.

The GSM (Group Special Mobile) digital cellular phones use FM with *minimum frequency-shift-keying* (MSK) where the peak frequency deviation is selected to produce orthogonal waveforms for binary one and binary zero data. (Digital phones use a speech codec to convert the analog voice source to a digital data source for transmission over the system.) Orthogonality occurs when $\Delta F = 1/4R$ where R is the bit rate (bits/sec) [1]. Actually, GSM uses Gaussian shaped MSK (GMSK). That is, the digital data waveform (with rectangular binary one and binary zero pulses) is first filtered by a low-pass filter having a Gaussian shaped frequency response (to attenuate the higher frequencies). This Gaussian filtered data waveform is then fed into the frequency modulator to generate the GMSK signal. This produces a digitally modulated FM signal with a relatively small bandwidth.

Other digital cellular standards use QPSK signalling as discussed in the next section.

1.8 QPSK Signalling

Quadrature phase-shift-keying (QPSK) is a special case of quadrature modulation as shown in Table 1.1 where $m_1(t) = \pm 1$ and $m_2(t) = \pm 1$ are two binary bit streams. The complex envelope for QPSK is

$$g(t) = x(t) + jy(t) = A_c\left[m_1(t) + jm_2(t)\right]$$

where $x(t) = \pm A_c$ and $y(t) = \pm A_c$. The permitted values for the complex envelope are illustrated by the QPSK **signal constellation** shown in Fig. 1.5a. The *signal constellation* is a plot of the permitted values for the complex envelope, $g(t)$. QPSK may be generated by using the quadrature generation technique of Fig. 1.3 where the baseband signal processor is a serial-to-parallel converter that reads in two bits of data at a time from the serial binary input stream, $m(t)$ and outputs the first of the two bits to $x(t)$ and the second bit to $y(t)$. If the two input bits are both binary ones, (11), then $m_1(t) = +A_c$ and $m_2(t) = +A_c$. This is represented by the top right-hand dot for $g(t)$ in the signal constellation for QPSK signalling in Fig. 1.5a. Likewise, the three other possible two-bit words, (10), (01), and (00), are also shown. The QPSK signal is also equivalent to a four-phase phase-shift-keyed signal (4PSK) since all the points in the signal constellation fall on a circle where the permitted phases are $\theta(t) = 45°$, 135°, 225°, and 315°. There is no amplitude modulation on the QPSK signal since the distances from the origin to all the signal points on the signal constellation are equal.

For QPSK, the spectrum of $g(t)$ is of the $\sin x/x$ type since $x(t)$ and $y(t)$ consists of rectangular data pulses of value $\pm A_c$. Moreover, it can be shown that for equally likely independent binary one and binary zero data, the power spectral density of $g(t)$ for digitally modulated signals with M point signal constellations is [1]

$$\mathcal{P}_g(f) = K\left(\frac{\sin \pi f \ell T_b}{\pi f \ell T_b}\right)^2 \tag{1.28}$$

where K is a constant, $R = 1/T_b$ is the data rate (bits/sec) of $m(t)$ and $M = 2^\ell$. M is the number of points in the signal constellation. For QPSK, $M = 4$ and $\ell = 2$. This PSD for the complex envelope, $\mathcal{P}_g(f)$, is plotted in Fig. 1.6. The PSD for the QPSK signal ($\ell = 2$) is given by translating $\mathcal{P}_g(f)$ up to the carrier frequency as indicated by (1.9).

Referring to Fig. 1.6 or using (1.28), the first-null bandwidth of $g(t)$ is R/ℓ Hz. Consequently, the null-to-null bandwidth of the modulated RF signal is

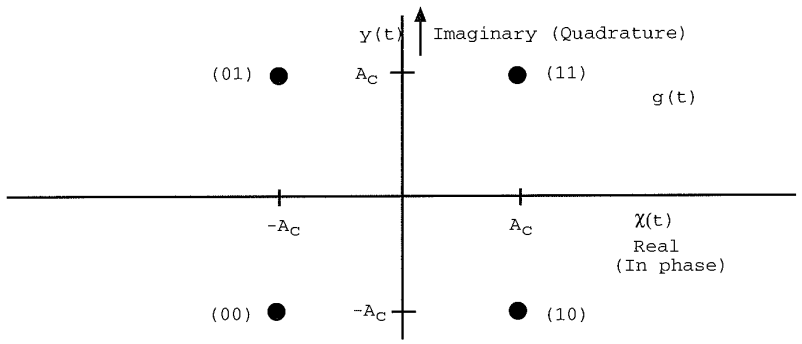
$$B_{\text{null}} = \frac{2R}{\ell} \text{ Hz} \tag{1.29}$$

For example, if the data rate of the baseband information source is 9600 bits/sec, then the null-to-null bandwidth of the QPSK signal would be 9.6 Hz since $\ell = 2$.

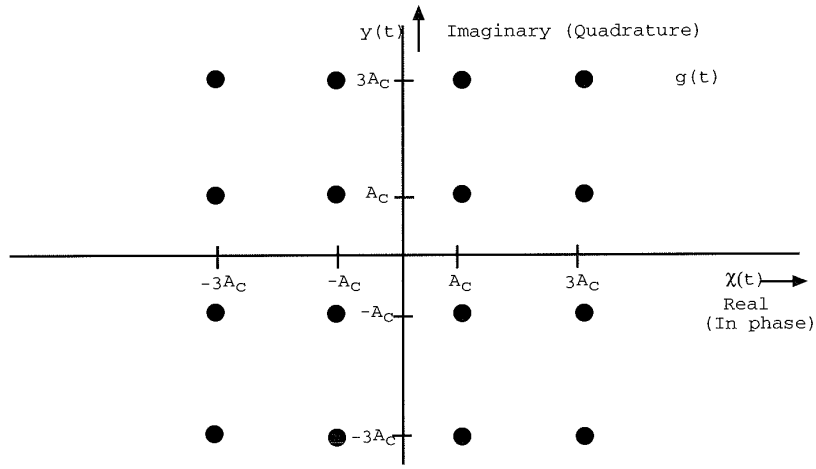
Referring to Fig. 1.6, it is seen that the sidelobes of the spectrum are relatively large so, in practice, the sidelobes of the spectrum are filtered off to prevent interference to the adjacent channels. This filtering rounds off the edges of the rectangular data pulses and this causes some amplitude modulation on the QPSK signal. That is, the points in the signal constellation for the filtered QPSK signal would be fuzzy since the transition from one constellation point to another point is not instantaneous because the filtered data pulses are not rectangular. QPSK is the modulation used for digital cellular phones with the IS-95 Code Division Multiple Access (CDMA) standard.

Equation (1.28) and Fig. 1.6 also represent the spectrum for *quadrature modulation amplitude modulation* (QAM) signalling. QAM signalling allows more than two values for $x(t)$ and $y(t)$. For example QAM where $M = 16$ has 16 points in the signal constellation with 4 values for $x(t)$ and 4 values for $y(t)$ such as, for example, $x(t) = +A_c$, $-A_c$, $+3A_c$, $-3A_c$ and $y(t) = +A_c$, $-A_c$, $+3A_c$, $-3A_c$. This is shown in Fig. 1.5b. Each point in the $M = 16$ QAM signal constellation would represent a unique four-bit data word, as compared with the $M = 4$ QPSK signal constellation shown in Fig. 1.5a where each point represents a unique two-bit data word. For a $R = 9600$ bits/sec information source data rate, a $M = 16$ QAM signal would have a null-to-null bandwidth of 4.8 kHz since $\ell = 4$.

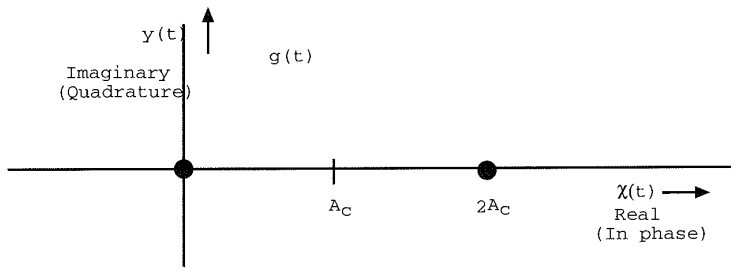
Ford Motor Co.
Exhibit 1030
Page 116



(a) QPSK Signal Constellation



(b) 16 QAM Signal Constellation



(c) OOK Signal Constellation

FIGURE 1.5 Signal constellations (permitted values of the complex envelope).

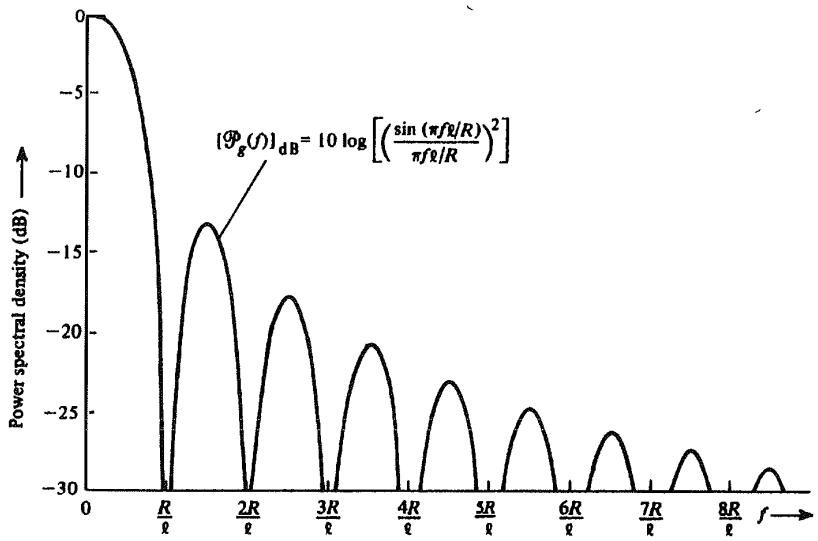


FIGURE 1.6 PSD for the complex envelope of MPSK and QAM where $M = 2^\ell$ and R is bit rate (positive frequencies shown). *Source:* Couch, L.W., II. 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 350. With permission.

For OOK signalling as described at the end of Section 1.6, the signal constellation would consist of $M = 2$ points along the x axis where $x = 0, 2A_c$ and $y = 0$. This is illustrated in Fig. 1.5c. For a $R = 9600$ bit/sec information source data rate, an OOK signal would have a null-to-null bandwidth of 19.2 kHz since $\ell = 1$.

Defining Terms

Bandpass waveform: The spectrum of the waveform is nonzero for frequencies in some band concentrated about a frequency $f_c \gg 0$; f_c is called the carrier frequency.

Baseband waveform: The spectrum of the waveform is nonzero for frequencies near $f = 0$.

Complex envelope: The function $g(t)$ of a bandpass waveform $v(t)$ where the bandpass waveform is described by

$$v(t) = Re\left\{g(t)e^{j\omega_c t}\right\}$$

Fourier transform: If $w(t)$ is a waveform, then the Fourier transform of $w(t)$ is

$$W(f) = \Im[w(t)] = \int_{-\infty}^{\infty} w(t)e^{-j2\pi ft}dt$$

where f has units of hertz.

Modulated signal: The bandpass signal

$$s(t) = Re\left\{g(t)e^{j\omega_c t}\right\}$$

Ford Motor Co.
Exhibit 1030
Page 118

where fluctuations of $g(t)$ are caused by the information source such as audio, video, or data.

Modulation: The information source, $m(t)$, that causes fluctuations in a bandpass signal.

Real envelope: The function $R(t) = |g(t)|$ of a bandpass waveform $v(t)$ where the bandpass waveform is described by

$$v(t) = Re\left\{g(t)e^{j\omega_c t}\right\}$$

Signal constellation: The permitted values of the complex envelope for a digital modulating source.

References

[1] Couch, L.W., II, *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, 1997.

Further Information

[1] Bedrosian, E., The analytic signal representation of modulated waveforms. *Proc. IRE,* vol. 50, October, 2071–2076, 1962.

[2] Couch, L.W., II, *Modern Communication Systems: Principles and Applications,* Macmillan Publishing, New York, (now Prentice Hall, Upper Saddle River, NJ), 1995.

[3] Dugundji, J., Envelopes and pre-envelopes of real waveforms. *IRE Trans. Information Theory,* vol. IT-4, March, 53–57, 1958.

[4] Voelcker, H.B., Toward the unified theory of modulation—Part I: Phase-envelope relationships. *Proc. IRE,* vol.54, March, 340–353, 1966.

[5] Voelcker, H.B., Toward the unified theory of modulation—Part II: Zero manipulation. *Proc. IRE,* vol. 54, May, 735–755, 1966.

[6] Ziemer, R.E. and Tranter, W.H., *Principles of Communications,* 4th ed., John Wiley and Sons, New York, 1995.

Ford Motor Co.
Exhibit 1030
Page 119

2

Sampling

2.1 Introduction 2-1
2.2 Instantaneous Sampling 2-1
Ideal Sampled Signal • Band-Limited Signals
2.3 Sampling Theorem 2-3
2.4 Sampling of Sinusoidal Signals 2-4
2.5 Sampling of Bandpass Signals 2-4
2.6 Practical Sampling 2-6
Natural Sampling • Flat-Top Sampling
2.7 Sampling Theorem in the Frequency Domain 2-9
2.8 Summary and Discussion 2-9
Defining Terms 2-10
References 2-10
Further Information 2-10

Hwei P. Hsu
Fairleigh Dickinson University

2.1 Introduction

To transmit analog message signals, such as speech signals or video signals, by digital means, the signal has to be converted into digital form. This process is known as analog-to-digital conversion. The sampling process is the first process performed in this conversion, and it converts a continuous-time signal into a discrete-time signal or a sequence of numbers. Digital transmission of analog signals is possible by virtue of the sampling theorem, and the sampling operation is performed in accordance with the sampling theorem.

In this chapter, using the Fourier transform technique, we present this remarkable sampling theorem and discuss the operation of sampling and practical aspects of sampling.

2.2 Instantaneous Sampling

Suppose we sample an arbitrary analog signal $m(t)$ shown in Fig. 2.1(a) instantaneously at a uniform rate, once every T_s seconds. As a result of this sampling process, we obtain an infinite sequence of samples $\{m(nT_s)\}$, where n takes on all possible integers. This form of sampling is called *instantaneous sampling.* We refer to T_s as the **sampling interval,** and its reciprocal $1/T_s = f_s$ as the **sampling rate.** Sampling rate (samples per second) is often cited in terms of sampling frequency expressed in hertz.

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 120

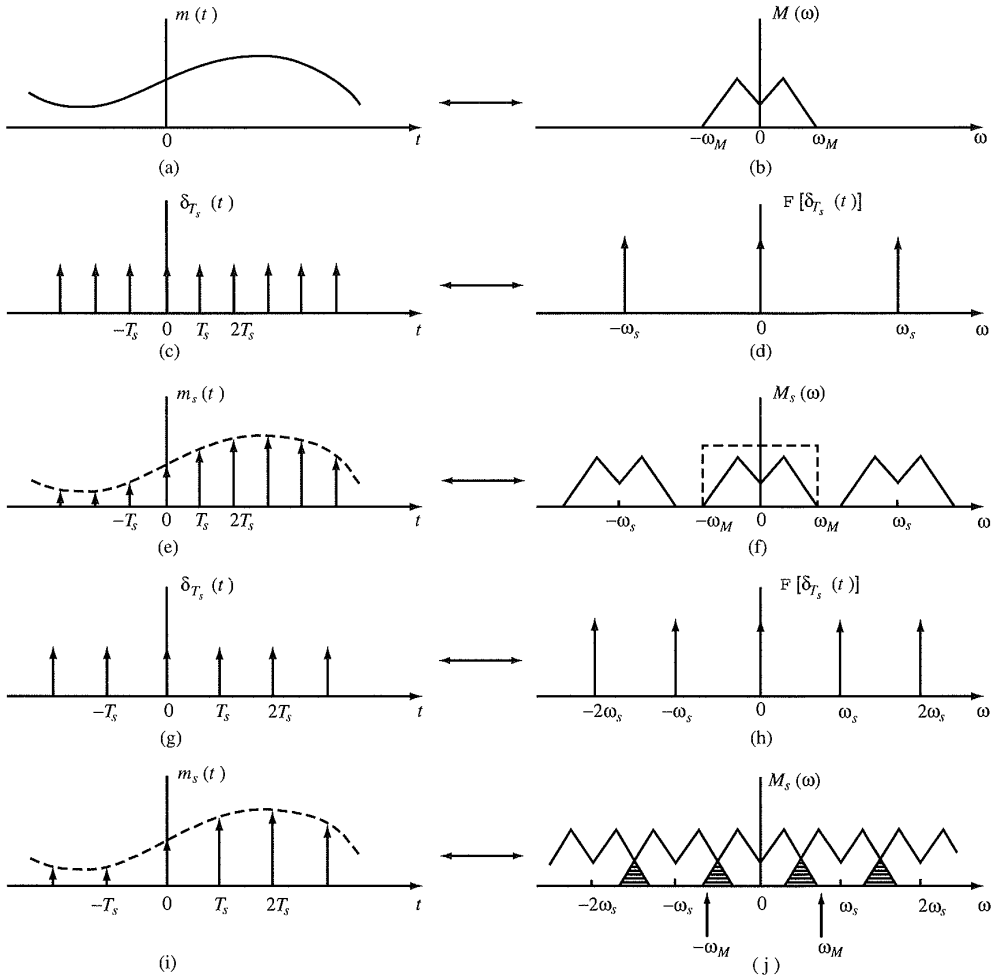


FIGURE 2.1 Illustration of instantaneous sampling and sampling theorem.

2.2.1 Ideal Sampled Signal

Let $m_s(t)$ be obtained by multiplication of $m(t)$ by the unit impulse train $\delta_T(t)$ with period T_s [Fig. 2.1(c)], that is,

$$\begin{aligned} m_s(t) &= m(t)\delta_{T_s}(t) = m(t)\sum_{n=-\infty}^{\infty}\delta(t-nT_s) \\ &= \sum_{n=-\infty}^{\infty} m(t)\delta(t-nT_s) = \sum_{n=-\infty}^{\infty} m(nT_s)\,\delta(t-nT_s) \end{aligned} \tag{2.1}$$

where we used the property of the δ function, $m(t)\delta(t-t_0) = m(t_0)\delta(t-t_0)$. The signal $m_s(t)$ [Fig. 2.1(e)] is referred to as the **ideal sampled signal.**

Ford Motor Co.
Exhibit 1030
Page 121

2.2.2 Band-Limited Signals

A real-valued signal $m(t)$ is called a **band-limited signal** if its Fourier transform $M(\omega)$ satisfies the condition

$$M(\omega) = 0 \qquad \text{for } |\omega| > \omega_M \tag{2.2}$$

where $\omega_M = 2\pi f_M$ [Fig. 2.1(b)]. A band-limited signal specified by Eq. (2.2) is often referred to as a *low-pass signal.*

2.3 Sampling Theorem

The sampling theorem states that a band-limited signal $m(t)$ specified by Eq. (2.2) can be uniquely determined from its values $m(nT_s)$ sampled at uniform interval T_s if $T_s \leq \pi/\omega_M = 1/(2f_M)$. In fact, when $T_s = \pi/\omega_M, m(t)$ is given by

$$m(t) = \sum_{n=-\infty}^{\infty} m(nT_s) \frac{\sin \omega_M (t - nT_s)}{\omega_M (t - nT_s)} \tag{2.3}$$

which is known as the **Nyquist–Shannon interpolation formula** and it is also sometimes called the *cardinal series.* The sampling interval $T_s = 1/(2f_M)$is called the *Nyquist interval* and the minimum rate $f_s = 1/T_s = 2f_M$ is known as the **Nyquist rate.**

Illustration of the instantaneous sampling process and the sampling theorem is shown in Fig. 2.1. The Fourier transform of the unit impulse train is given by [Fig. 2.1(d)]

$$\mathcal{F}\{\delta_{T_s}(t)\} = \omega_s \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_s) \qquad \omega_s = 2\pi/T_s \tag{2.4}$$

Then, by the convolution property of the Fourier transform, the Fourier transform $M_s(\omega)$ of the ideal sampled signal $m_s(t)$ is given by

$$\begin{aligned} M_s(\omega) &= \frac{1}{2\pi}\left[M(\omega) * \omega_s \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_s)\right] \\ &= \frac{1}{T_s} \sum_{n=-\infty}^{\infty} M(\omega - n\omega_s) \end{aligned} \tag{2.5}$$

where $*$ denotes convolution and we used the convolution property of the δ-function $M(\omega) * \delta(\omega - \omega_0) = M(\omega - \omega_0)$. Thus, the sampling has produced images of $M(\omega)$ along the frequency axis. Note that $M_s(\omega)$ will repeat periodically without overlap as long as $\omega_s \geq 2\omega_M$ or $f_s \geq 2f_M$ [Fig. 2.1(f)]. It is clear from Fig. 2.1(f) that we can recover $M(\omega)$ and, hence, $m(t)$ by passing the sampled signal $m_s(t)$ through an ideal low-pass filter having frequency response

$$H(\omega) = \begin{cases} T_s, & |\omega| \leq \omega_M \\ 0, & \text{otherwise} \end{cases} \tag{2.6}$$

where $\omega_M = \pi/T_s$. Then

$$M(\omega) = M_s(\omega)H(\omega) \tag{2.7}$$

Taking the inverse Fourier transform of Eq. (2.6), we obtain the impulse response $h(t)$ of the ideal low-pass filter as

$$h(t) = \frac{\sin \omega_M t}{\omega_M t} \tag{2.8}$$

Ford Motor Co.
Exhibit 1030
Page 122

Taking the inverse Fourier transform of Eq. (2.7), we obtain

$$\begin{aligned} m(t) &= m_s(t) * h(t) \\ &= \sum_{n=-\infty}^{\infty} m(nT_s)\,\delta(t-nT_s) * \frac{\sin \omega_M t}{\omega_M t} \\ &= \sum_{n=-\infty}^{\infty} m(nT_s)\,\frac{\sin \omega_M (t-nT_s)}{\omega_M (t-nT_s)} \end{aligned} \tag{2.9}$$

which is Eq. (2.3).

The situation shown in Fig. 2.1(j) corresponds to the case where $f_s < 2f_M$. In this case there is an overlap between $M(\omega)$ and $M(\omega - \omega_M)$. This overlap of the spectra is known as *aliasing* or *foldover*. When this aliasing occurs, the signal is distorted and it is impossible to recover the original signal $m(t)$ from the sampled signal. To avoid aliasing, in practice, the signal is sampled at a rate slightly higher than the Nyquist rate. If $f_s > 2f_M$, then as shown in Fig. 2.1(f), there is a gap between the upper limit ω_M of $M(\omega)$ and the lower limit $\omega_s - \omega_M$ of $M(\omega - \omega_s)$. This range from ω_M to $\omega_s - \omega_M$ is called a *guard band.* As an example, speech transmitted via telephone is generally limited to $f_M = 3.3$ kHz (by passing the sampled signal through a low-pass filter). The Nyquist rate is, thus, 6.6 kHz. For digital transmission, the speech is normally sampled at the rate $f_s = 8$ kHz. The guard band is then $f_s - 2f_M = 1.4$ kHz. The use of a sampling rate higher than the Nyquist rate also has the desirable effect of making it somewhat easier to design the low-pass reconstruction filter so as to recover the original signal from the sampled signal.

2.4 Sampling of Sinusoidal Signals

A special case is the sampling of a sinusoidal signal having the frequency f_M. In this case we require that $f_s > 2f_M$ rather that $f_s \geq 2f_M$. To see that this condition is necessary, let $f_s = 2f_M$. Now, if an initial sample is taken at the instant the sinusoidal signal is zero, then all successive samples will also be zero. This situation is avoided by requiring $f_s > 2f_M$.

2.5 Sampling of Bandpass Signals

A real-valued signal $m(t)$ is called a **bandpass signal** if its Fourier transform $M(\omega)$ satisfies the condition

$$M(\omega) = 0 \quad \text{except for} \quad \begin{cases} \omega_1 < \omega < \omega_2 \\ -\omega_2 < \omega < -\omega_1 \end{cases} \tag{2.10}$$

where $\omega_1 = 2\pi f_1$ and $\omega_2 = 2\pi f_2$ [Fig. 2.2(a)].

The sampling theorem for a band-limited signal has shown that a sampling rate of $2f_2$ or greater is adequate for a low-pass signal having the highest frequency f_2. Therefore, treating $m(t)$ specified by Eq. (2.10) as a special case of such a low-pass signal, we conclude that a sampling rate of $2f_2$ is adequate for the sampling of the bandpass signal $m(t)$. But it is not necessary to sample this fast. The minimum allowable sampling rate depends on f_1, f_2, and the bandwidth $f_B = f_2 - f_1$.

Let us consider the direct sampling of the bandpass signal specified by Eq. (2.10). The spectrum of the sampled signal is periodic with the period $\omega_s = 2\pi f_s$, where f_s is the sampling frequency, as in Eq. (2.4). Shown in Fig. 2.2(b) are the two right shifted spectra of the negative side spectrum $M_-(\omega)$. If the recovering of the bandpass signal is achieved by

Ford Motor Co.
Exhibit 1030
Page 123

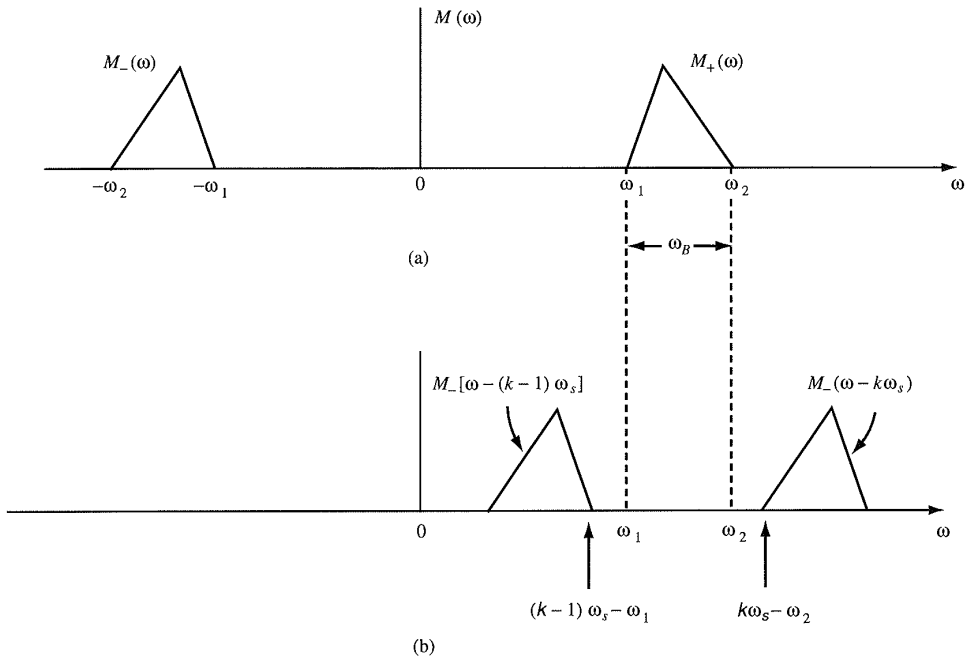


FIGURE 2.2 (a) Spectrum of a bandpass signal; (b) Shifted spectra of $M_{-}(\omega)$.

passing the sampled signal through an ideal bandpass filter covering the frequency bands $(-\omega_2, -\omega_1)$ and (ω_1, ω_2), it is necessary that there be no aliasing problem. From Fig. 2.2(b), it is clear that to avoid overlap it is necessary that

$$\omega_s \geq 2(\omega_2 - \omega_1) \tag{2.11}$$

$$(k-1)\omega_s - \omega_1 \leq \omega_1 \tag{2.12}$$

and

$$k\omega_s - \omega_2 \geq \omega_2 \tag{2.13}$$

where $\omega_1 = 2\pi f_1$, $\omega_2 = 2\pi f_2$, and k is an integer ($k = 1, 2, \ldots$). Since $f_1 = f_2 - f_B$, these constraints can be expressed as

$$1 \leq k \leq \frac{f_2}{f_B} \leq \frac{k}{2}\frac{f_s}{f_B} \tag{2.14}$$

and

$$\frac{k-1}{2}\frac{f_s}{f_B} \leq \frac{f_2}{f_B} - 1 \tag{2.15}$$

A graphical description of Eqs. (2.14) and (2.15) is illustrated in Fig. 2.3. The unshaded regions represent where the constraints are satisfied, whereas the shaded regions represent the regions where the constraints are not satisfied and overlap will occur. The solid line in Fig. 2.3 shows the locus of the minimum sampling rate. The minimum sampling rate is given by

$$\min\{f_s\} = \frac{2f_2}{m} \tag{2.16}$$

Ford Motor Co.
Exhibit 1030
Page 124

where m is the largest integer not exceeding f_2/f_B. Note that if the ratio f_2/f_B is an integer, then the minimum sampling rate is $2f_B$. As an example, consider a bandpass signal with $f_1 = 1.5$ kHz and $f_2 = 2.5$ kHz. Here $f_B = f_2 - f_1 = 1$ kHz, and $f_2/f_B = 2.5$. Then from Eq. (2.16) and Fig. 2.3 we see that the minimum sampling rate is $2f_2/2 = f_2 = 2.5$ kHz, and allowable ranges of sampling rate are 2.5 kHz $\leq f_s \leq 3$ kHz and $f_s \geq 5$ kHz $(= 2f_2)$.

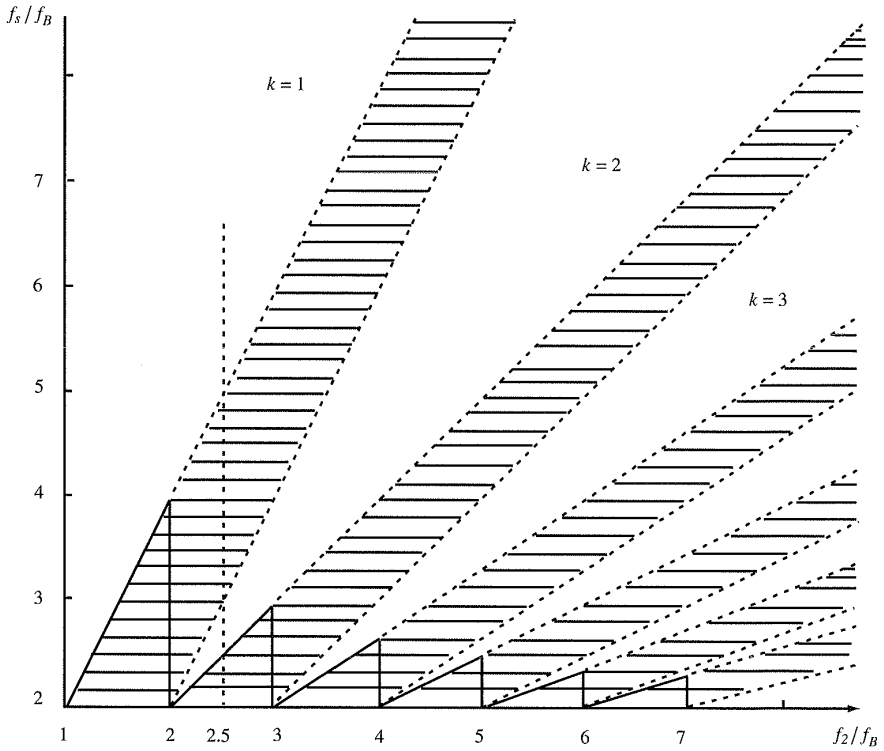


FIGURE 2.3 Minimum and permissible sampling rates for a bandpass signal.

2.6 Practical Sampling

In practice, the sampling of an analog signal is performed by means of high-speed switching circuits, and the sampling process takes the form of *natural sampling* or **flat-top sampling.**

2.6.1 Natural Sampling

Natural sampling of a band-limited signal $m(t)$ is shown in Fig. 2.4. The sampled signal $m_{\text{ns}}(t)$ can be expressed as

$$m_{\text{ns}}(t) = m(t)x_p(t) \tag{2.17}$$

Ford Motor Co.
Exhibit 1030
Page 125

where $x_p(t)$ is the periodic train of rectangular pulses with fundamental period T_s, and each rectangular pulse in $x_p(t)$ has duration d and unit amplitude [Fig. 2.4(b)]. Observe that the sampled signal $m_{\text{ns}}(t)$ consists of a sequence of pulses of varying amplitude whose tops follow the waveform of the signal $m(t)$ [Fig. 2.4(c)].

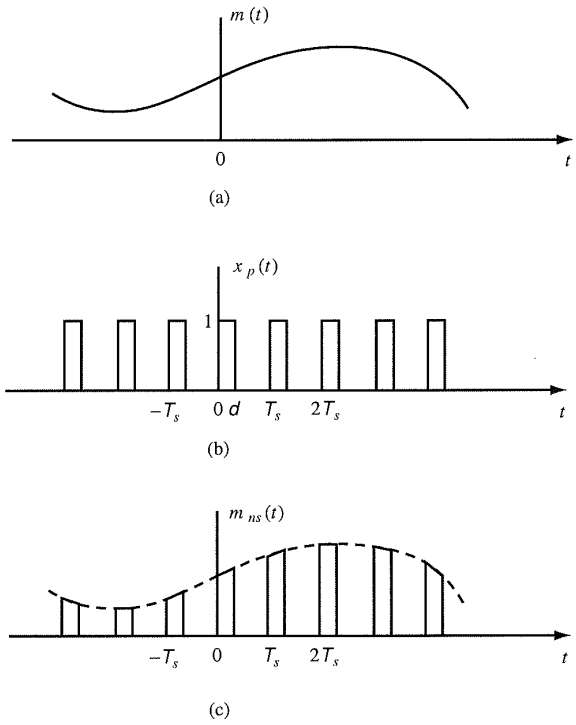


FIGURE 2.4 Natural sampling.

The Fourier transform of $x_p(t)$ is

$$X_p(\omega) = \sum_{n=-\infty}^{\infty} c_n \delta(\omega - n\omega_s) \qquad \omega_s = 2\pi/T_s \tag{2.18}$$

where

$$c_n = \frac{d}{T_s} \frac{\sin(n\omega_s d/2)}{n\omega_s d/2} e^{-jn\omega_s d/2} \tag{2.19}$$

Then the Fourier transform of $m_{\text{ns}}(t)$ is given by

$$M_{\text{ns}}(\omega) = M(\omega) * X_p(\omega) = \sum_{n=-\infty}^{\infty} c_n M(\omega - n\omega_s) \tag{2.20}$$

from which we see that the effect of the natural sampling is to multiply the nth shifted spectrum $M(\omega - n\omega_s)$ by a constant c_n. Thus, the original signal $m(t)$ can be reconstructed from $m_{\text{ns}}(t)$ with no distortion by passing $m_{\text{ns}}(t)$ through an ideal low-pass filter if the sampling rate f_s is equal to or greater than the Nyquist rate $2f_M$.

Ford Motor Co.
Exhibit 1030
Page 126

2.6.2 Flat-Top Sampling

The sampled waveform, produced by practical sampling devices that are the sample and hold types, has the form [Fig. 2.5(c)]

$$m_{\text{fs}}(t) = \sum_{n=-\infty}^{\infty} m(nT_s)\, p(t - nT_s) \tag{2.21}$$

where $p(t)$ is a rectangular pulse of duration d with unit amplitude [Fig. 2.5(a)]. This type of sampling is known as **flat-top sampling.** Using the ideal sampled signal $m_s(t)$ of Eq. (2.1), $m_{\text{fs}}(t)$ can be expressed as

$$m_{\text{fs}}(t) = p(t) * \left[\sum_{n=-\infty}^{\infty} m(nT_s)\, \delta(t - nT_s) \right] = p(t) * m_s(t) \tag{2.22}$$

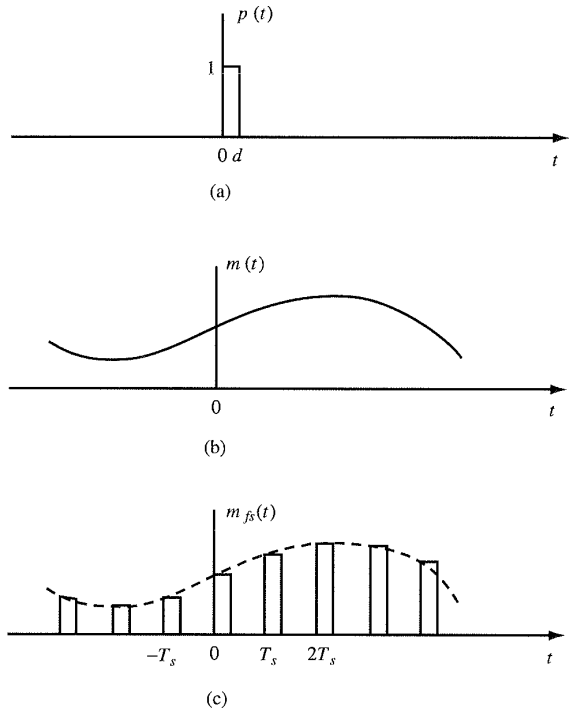


FIGURE 2.5 Flat-top sampling.

Using the convolution property of the Fourier transform and Eq. (2.4), the Fourier transform of $m_{\text{fs}}(t)$ is given by

$$M_{\text{fs}}(\omega) = P(\omega)M_s(\omega) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} P(\omega)M(\omega - n\omega_s) \tag{2.23}$$

Ford Motor Co.
Exhibit 1030
Page 127

where

$$P(\omega) = d\frac{\sin(\omega d/2)}{\omega d/2}e^{-j\omega d/2} \tag{2.24}$$

From Eq.(2.23) we see that by using flat-top sampling we have introduced amplitude distortion and time delay, and the primary effect is an attenuation of high-frequency components. This effect is known as the *aperture effect.* The aperture effect can be compensated by an equalizing filter with a frequency response $H_{eq}(\omega) = 1/P(\omega)$. If the pulse duration d is chosen such that $d \ll T_s$, however, then $P(\omega)$ is essentially constant over the baseband and no equalization may be needed.

2.7 Sampling Theorem in the Frequency Domain

The sampling theorem expressed in Eq. (2.4) is the time-domain sampling theorem. There is a dual to this time-domain sampling theorem, i.e., the sampling theorem in the frequency domain.

Time-limited signals: A continuous-time signal $m(t)$ is called **time limited** if

$$m(t) = 0 \qquad \text{for} \quad |t| > |T_0| \tag{2.25}$$

Frequency-domain sampling theorem: The frequency-domain sampling theorem states that the Fourier transform $M(\omega)$ of a time-limited signal $m(t)$ specified by Eq. (2.25) can be uniquely determined from its values $M(n\omega_s)$ sampled at a uniform rate ω_s if $\omega_s \leq \pi/T_0$. In fact, when $\omega_s = \pi/T_0$, then $M(\omega)$ is given by

$$M(\omega) = \sum_{n=-\infty}^{\infty} M(n\omega_s)\frac{\sin T_0(\omega - n\omega_s)}{T_0(\omega - n\omega_s)} \tag{2.26}$$

2.8 Summary and Discussion

The sampling theorem is the fundamental principle of digital communications. We state the sampling theorem in two parts.

THEOREM 2.1 *If the signal contains no frequency higher than f_M* Hz, *it is completely described by specifying its samples taken at instants of time spaced $1/2f_M$* s.

THEOREM 2.2 *The signal can be completely recovered from its samples taken at the rate of $2f_M$ samples per second or higher.*

The preceding sampling theorem assumes that the signal is strictly band limited. It is known that if a signal is band limited it cannot be time limited and vice versa. In many practical applications, the signal to be sampled is time limited and, consequently, it cannot be strictly band limited. Nevertheless, we know that the frequency components of physically occurring signals attenuate rapidly beyond some defined bandwidth, and for practical purposes we consider these signals are band limited. This approximation of real signals by band limited ones introduces no significant error in the application of the sampling theorem. When such a signal is sampled, we band limit the signal by filtering before sampling and sample at a rate slightly higher than the nominal Nyquist rate.

Ford Motor Co.
Exhibit 1030
Page 128

Defining Terms

Band-limited signal: A signal whose frequency content (Fourier transform) is equal to zero above some specified frequency.

Bandpass signal: A signal whose frequency content (Fourier transform) is nonzero only in a band of frequencies not including the origin.

Flat-top sampling: Sampling with finite width pulses that maintain a constant value for a time period less than or equal to the sampling interval. The constant value is the amplitude of the signal at the desired sampling instant.

Ideal sampled signal: A signal sampled using an ideal impulse train.

Nyquist rate: The minimum allowable sampling rate of $2f_M$ samples per second, to reconstruct a signal band limited to f_M hertz.

Nyquist-Shannon interpolation formula: The infinite series representing a time domain waveform in terms of its ideal samples taken at uniform intervals.

Sampling interval: The time between samples in uniform sampling.

Sampling rate: The number of samples taken per second (expressed in Hertz and equal to the reciprocal of the sampling interval).

Time-limited: A signal that is zero outside of some specified time interval.

References

[1] Brown, J.L. Jr., First order sampling of bandpass signals—A new approach. *IEEE Trans. Information Theory,* IT-26(5), 613–615, 1980.

[2] Byrne, C.L. and Fitzgerald, R.M., Time-limited sampling theorem for band-limited signals, *IEEE Trans. Information Theory,* IT-28(5), 807–809, 1982.

[3] Hsu, H.P., *Applied Fourier Analysis,* Harcourt Brace Jovanovich, San Diego, CA, 1984.

[4] Hsu, H.P., *Analog and Digital Communications,* McGraw-Hill, New York, 1993.

[5] Hulthén, R., Restoring causal signals by analytical continuation: A generalized sampling theorem for causal signals. *IEEE Trans. Acoustics, Speech, and Signal Processing,* ASSP-31(5), 1294–1298, 1983.

[6] Jerri, A.J., The Shannon sampling theorem—Its various extensions and applications: A tutorial review, *Proc. IEEE.* 65(11), 1565–1596, 1977.

Further Information

For a tutorial review of the sampling theorem, historical notes, and earlier references see Jerri [6].

Ford Motor Co.
Exhibit 1030
Page 129

3

Pulse Code Modulation[1]

3.1 Introduction 3-1
3.2 Generation of PCM 3-2
3.3 Percent Quantizing Noise 3-3
3.4 Practical PCM Circuits 3-4
3.5 Bandwidth of PCM 3-6
3.6 Effects of Noise 3-6
3.7 Nonuniform Quantizing: μ-Law and A-Law Companding 3-9
3.8 Example: Design of a PCM System 3-11
Defining Terms 3-12
References 3-12
Further Information 3-13

Leon W. Couch, II
University of Florida

3.1 Introduction

Pulse code modulation (PCM) is analog-to-digital conversion of a special type where the information contained in the instantaneous samples of an analog signal is represented by digital words in a serial bit stream.

If we assume that each of the digital words has n binary digits, there are $M = 2^n$ unique code words that are possible, each code word corresponding to a certain amplitude level. Each sample value from the analog signal, however, can be any one of an infinite number of levels, so that the digital word that represents the amplitude closest to the actual sampled value is used. This is called **quantizing**. That is, instead of using the exact sample value of the analog waveform, the sample is replaced by the closest allowed value, where there are M allowed values, and each allowed value corresponds to one of the code words.

PCM is very popular because of the many advantages it offers. Some of these advantages are as follows.

- Relatively inexpensive digital circuitry may be used extensively in the system.
- PCM signals derived from all types of analog sources (audio, video, etc.) may be time-division multiplexed with data signals (e.g., from digital computers) and transmitted over a common high-speed digital communication system.

[1] *Source:* Leon W. Couch, II. 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ. With permission.

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 130

- In long-distance digital telephone systems requiring repeaters, a *clean* PCM waveform can be regenerated at the output of each repeater, where the input consists of a noisy PCM waveform. The noise at the input, however, may cause bit errors in the regenerated PCM output signal.
- The noise performance of a digital system can be superior to that of an analog system. In addition, the probability of error for the system output can be reduced even further by the use of appropriate coding techniques.

These advantages usually outweigh the main disadvantage of PCM: a much wider bandwidth than that of the corresponding analog signal.

3.2 Generation of PCM

The PCM signal is generated by carrying out three basic operations: sampling, quantizing, and encoding (see Fig. 3.1). The sampling operation generates an instantaneously-sampled flat-top **pulse-amplitude modulated** (PAM) signal.

The quantizing operation is illustrated in Fig. 3.2 for the $M = 8$ level case. This quantizer is said to be *uniform* since all of the steps are of equal size. Since we are approximating the analog sample values by using a finite number of levels ($M = 8$ in this illustration), *error* is introduced into the recovered output analog signal because of the quantizing effect. The error waveform is illustrated in Fig. 3.2c. The quantizing error consists of the difference between the analog signal at the sampler input and the output of the quantizer. Note that the peak value of the error (± 1) is one-half of the quantizer step size (2). If we sample at the Nyquist rate ($2B$, where B is the absolute bandwidth, in hertz, of the input analog signal) or faster and there is negligible channel noise, there will still be noise, called *quantizing noise,* on the recovered analog waveform due to this error. The quantizing noise can also be thought of as a round-off error. The quantizer output is a *quantized* (i.e., only M possible amplitude values) PAM signal.

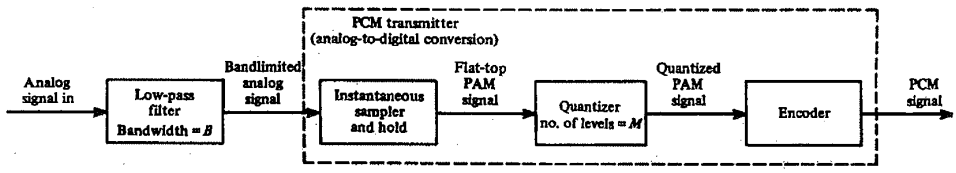


FIGURE 3.1 A PCM transmitter. *Source:* Couch, L.W. II 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 138. With permission.

The PCM signal is obtained from the quantized PAM signal by encoding each quantized sample value into a digital word. It is up to the system designer to specify the exact code word that will represent a particular quantized level. If a Gray code of Table 3.1 is used, the resulting PCM signal is shown in Fig. 3.2d where the PCM word for each quantized sample is strobed out of the encoder by the next clock pulse. The Gray code was chosen because it has only 1-b change for each step change in the quantized level. Consequently, single errors in the received PCM code word will cause minimum errors in the recovered analog level, provided that the sign bit is not in error.

Here we have described PCM systems that represent the quantized analog sample values by *binary* code words. Of course, it is possible to represent the quantized analog samples by

Ford Motor Co.
Exhibit 1030
Page 131

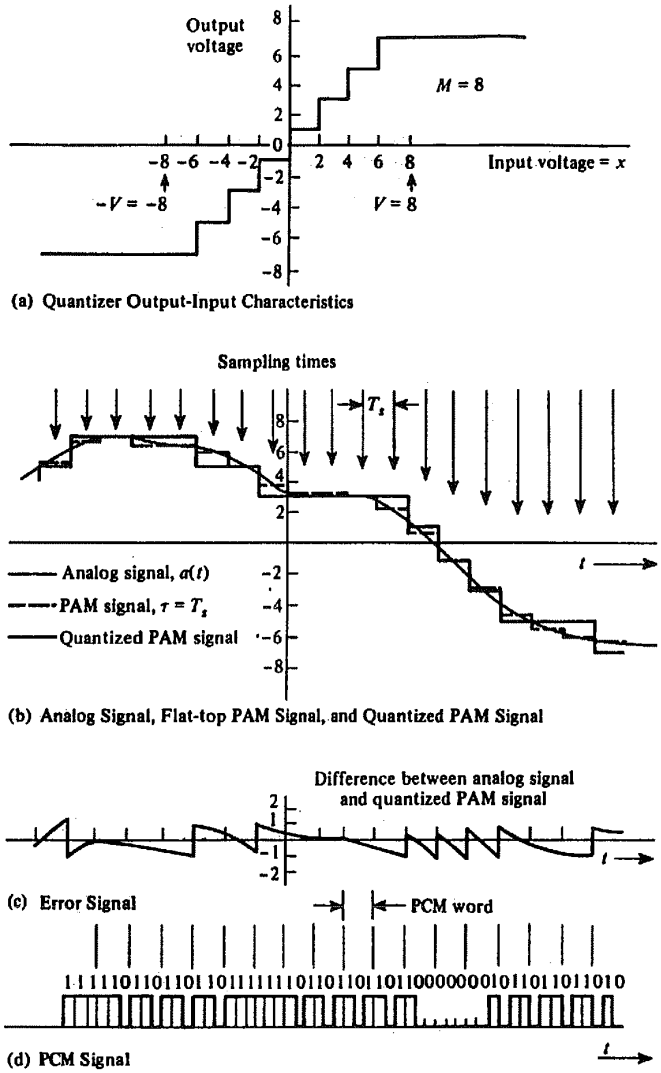


FIGURE 3.2 Illustration of waveforms in a PCM system. *Source:* Couch, L.W. II 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 139. With permission.

digital words using other than base 2. That is, for base q, the number of quantized levels allowed is $M = q^n$, where n is the number of q base digits in the code word. We will not pursue this topic since binary ($q = 2$) digital circuits are most commonly used.

3.3 Percent Quantizing Noise

The quantizer at the PCM encoder produces an error signal at the PCM decoder output as illustrated in Fig. 3.2c. The peak value of this error signal may be expressed as a percentage

TABLE 3.1 3-b Gray Code for $M = 8$ Levels

Quantized Sample Voltage	Gray Code Word (PCM Output)	
+7	110	
+5	111	
+3	101	
+1	100	
		Mirror image except for sign bit
−1	000	
−3	001	
−5	011	
−7	010	

Source: Couch, L.W., II. 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 140. With permission.

of the maximum possible analog signal amplitude. Referring to Fig. 3.2c, a peak error of 1 V occurs for a maximum analog signal amplitude of $M = 8$ V as shown Fig. 3.1c. Thus, in general,

$$\frac{2P}{100} = \frac{1}{M} = \frac{1}{2^n}$$

or

$$2^n = \frac{50}{P} \tag{3.1}$$

where P is the peak percentage error for a PCM system that uses n bit code words. The design value of n needed in order to have less than P percent error is obtained by taking the base 2 logarithm of both sides of Eq. (3.1), where it is realized that $\log_2(x) = [\log_{10}(x)]/\log_{10}(2) = 3.32\log_{10}(x)$. That is,

$$n \geq 3.32\log_{10}\left(\frac{50}{P}\right) \tag{3.2}$$

where n is the number of bits needed in the PCM word in order to obtain less than P percent error in the recovered analog signal (i.e., decoded PCM signal).

3.4 Practical PCM Circuits

Three techniques are used to implement the analog-to-digital converter (ADC) encoding operation. These are the *counting* or *ramp, serial* or *successive approximation*, and *parallel* or *flash* encoders.

In the counting encoder, at the same time that the sample is taken, a ramp generator is energized and a binary counter is started. The output of the ramp generator is continuously compared to the sample value; when the value of the ramp becomes equal to the sample value, the binary value of the counter is read. This count is taken to be the PCM word. The binary counter and the ramp generator are then reset to zero and are ready to be reenergized at the next sampling time. This technique requires only a few components, but the speed of this type of ADC is usually limited by the speed of the counter. The Maxim ICL7126 CMOS ADC integrated circuit uses this technique.

The serial encoder compares the value of the sample with trial quantized values. Successive trials depend on whether the past comparator outputs are positive or negative. The

Ford Motor Co.
Exhibit 1030
Page 133

trial values are chosen first in large steps and then in small steps so that the process will converge rapidly. The trial voltages are generated by a series of voltage dividers that are configured by (on-off) switches. These switches are controlled by digital logic. After the process converges, the value of the switch settings is read out as the PCM word. This technique requires more precision components (for the voltage dividers) than the ramp technique. The speed of the feedback ADC technique is determined by the speed of the switches. The National Semiconductor ADC0804 8-b ADC uses this technique.

The parallel encoder uses a set of parallel comparators with reference levels that are the permitted quantized values. The sample value is fed into all of the parallel comparators simultaneously. The high or low level of the comparator outputs determines the binary PCM word with the aid of some digital logic. This is a fast ADC technique but requires more hardware than the other two methods. The Harris CA3318 8-b ADC integrated circuit is an example of the technique.

All of the integrated circuits listed as examples have parallel digital outputs that correspond to the digital word that represents the analog sample value. For generation of PCM, the parallel output (digital word) needs to be converted to serial form for transmission over a two-wire channel. This is accomplished by using a parallel-to-serial converter integrated circuit, which is also known as a **serial-input-output** (SIO) chip. The SIO chip includes a shift register that is set to contain the parallel data (usually, from 8 or 16 input lines). Then the data are shifted out of the last stage of the shift register bit by bit onto a single output line to produce the serial format. Furthermore, the SIO chips are usually full duplex; that is, they have two sets of shift registers, one that functions for data flowing in each direction. One shift register converts parallel input data to serial output data for transmission over the channel, and, simultaneously, the other shift register converts received serial data from another input to parallel data that are available at another output. Three types of SIO chips are available: the *universal asynchronous receiver/transmitter* (UART), the *universal synchronous receiver/transmitter* (USRT), and the *universal synchronous/asynchronous receiver transmitter* (USART). The UART transmits and receives asynchronous serial data, the USRT transmits and receives synchronous serial data, and the USART combines both a UART and a USRT on one chip.

At the receiving end the PCM signal is decoded back into an analog signal by using a digital-to-analog converter (DAC) chip. If the DAC chip has a parallel data input, the received serial PCM data are first converted to a parallel form using a SIO chip as described in the preceding paragraph. The parallel data are then converted to an approximation of the analog sample value by the DAC chip. This conversion is usually accomplished by using the parallel digital word to set the configuration of electronic switches on a resistive current (or voltage) divider network so that the analog output is produced. This is called a *multiplying* DAC since the analog output voltage is directly proportional to the divider reference voltage multiplied by the value of the digital word. The Motorola MC1408 and the National Semiconductor DAC0808 8-b DAC chips are examples of this technique. The DAC chip outputs samples of the quantized analog signal that approximates the analog sample values. This may be smoothed by a low-pass reconstruction filter to produce the analog output.

The Communications Handbook [6, pp 107–117] and The Electrical Engineering Handbook [5, pp. 771–782] give more details on ADC, DAC, and PCM circuits.

3.5 Bandwidth of PCM

A good question to ask is: What is the spectrum of a PCM signal? For the case of PAM signalling, the spectrum of the PAM signal could be obtained as a function of the spectrum of the input analog signal because the PAM signal is a linear function of the analog signal. This is not the case for PCM. As shown in Figs. 3.1 and 3.2, the PCM signal is a nonlinear function of the input signal. Consequently, the spectrum of the PCM signal is not directly related to the spectrum of the input analog signal. It can be shown that the spectrum of the PCM signal depends on the bit rate, the correlation of the PCM data, and on the PCM waveform pulse shape (usually rectangular) used to describe the bits [2, 3]. From Fig. 3.2, the bit rate is

$$R = nf_s \tag{3.3}$$

where n is the number of bits in the PCM word ($M = 2^n$) and f_s is the sampling rate. For no aliasing we require $f_s \geq 2B$ where B is the bandwidth of the analog signal (that is to be converted to the PCM signal). The dimensionality theorem [2, 3] shows that the bandwidth of the PCM waveform is bounded by

$$B_{\text{PCM}} \geq \frac{1}{2}R = \frac{1}{2}nf_s \tag{3.4}$$

where equality is obtained if a $(\sin x)/x$ type of pulse shape is used to generate the PCM waveform. The exact spectrum for the PCM waveform will depend on the pulse shape that is used as well as on the type of line encoding. For example, if one uses a rectangular pulse shape with polar nonreturn to zero (NRZ) line coding, the first null bandwidth is simply

$$B_{\text{PCM}} = R = nf_s \text{ Hz} \tag{3.5}$$

Table 3.2 presents a tabulation of this result for the case of the minimum sampling rate, $f_s = 2B$. Note that Eq. (3.4) demonstrates that the bandwidth of the PCM signal has a lower bound given by

$$B_{\text{PCM}} \geq nB \tag{3.6}$$

where $f_s > 2B$ and B is the bandwidth of the corresponding analog signal. Thus, for reasonable values of n, the bandwidth of the PCM signal will be significantly larger than the bandwidth of the corresponding analog signal that it represents. For the example shown in Fig. 3.2 where $n = 3$, the PCM signal bandwidth will be at least three times wider than that of the corresponding analog signal. Furthermore, if the bandwidth of the PCM signal is reduced by improper filtering or by passing the PCM signal through a system that has a poor frequency response, the filtered pulses will be elongated (stretched in width) so that pulses corresponding to any one bit will smear into adjacent bit slots. If this condition becomes too serious, it will cause errors in the detected bits. This pulse smearing effect is called **intersymbol interference** (ISI).

3.6 Effects of Noise

The analog signal that is recovered at the PCM system output is corrupted by noise. Two main effects produce this noise or distortion: 1) quantizing noise that is caused by the M-step quantizer at the PCM transmitter and 2) bit errors in the recovered PCM signal. The bit errors are caused by *channel noise* as well as improper channel filtering, which causes ISI. In addition, if the input analog signal is not strictly band limited, there will be some

Ford Motor Co.
Exhibit 1030
Page 135

TABLE 3.2 Performance of a PCM System with Uniform Quantizing and No Channel Noise

Number of Quantizer Levels Used, M	Length of the PCM Word, n (bits)	Bandwidth of PCM Signal (First Null Bandwidth)[a]	Recovered Analog Signal Power-to-Quantizing Noise Power Ratios (dB) $(S/N)_{out}$
2	1	2B	6.0
4	2	4B	12.0
8	3	6B	18.1
16	4	8B	24.1
32	5	10B	30.1
64	6	12B	36.1
128	7	14B	42.1
256	8	16B	48.2
512	9	18B	54.2
1,024	10	20B	60.2
2,048	11	22B	66.2
4,096	12	24B	72.2
8,192	13	26B	78.3
16,384	14	28B	84.3
32,768	15	30B	90.3
65,536	16	32B	96.3

[a] B is the absolute bandwidth of the input analog signal. *Source:* Couch, L.W. II 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 142. With permission.

aliasing noise on the recovered analog signal [12]. Under certain assumptions, it can be shown that the recovered analog *average* signal power to the average noise power [2] is

$$\left(\frac{S}{N}\right)_{\text{out}} = \frac{M^2}{1 + 4\left(M^2 - 1\right) P_e} \tag{3.7}$$

where M is the number of uniformly spaced quantizer levels used in the PCM transmitter and P_e is the probability of bit error in the recovered binary PCM signal at the receiver DAC before it is converted back into an analog signal. Most practical systems are designed so that P_e is negligible. Consequently, if we assume that there are no bit errors due to channel noise (i.e., $P_e = 0$), the S/N due only to quantizing errors is

$$\left(\frac{S}{N}\right)_{\text{out}} = M^2 \tag{3.8}$$

Numerical values for these S/N ratios are given in Table 3.2.

To realize these S/N ratios, one critical assumption is that the peak-to-peak level of the analog waveform at the input to the PCM encoder is set to the design level of the quantizer. For example, referring to Fig. 3.2, this corresponds to the input traversing the range $-V$ to $+V$ volts where $V = 8$ V is the design level of the quantizer. Equation (3.7) was derived for waveforms with equally likely values, such as a triangle waveshape, that have a peak-to-peak value of $2V$ and an rms value of $V/\sqrt{3}$, where V is the design peak level of the quantizer.

From a practical viewpoint, the quantizing noise at the output of the PCM decoder can be categorized into four types depending on the operating conditions. The four types are overload noise, random noise, granular noise, and hunting noise. As discussed earlier, the level of the analog waveform at the input of the PCM encoder needs to be set so that its peak level does not exceed the design peak of V volts. If the peak input does exceed V, the recovered analog waveform at the output of the PCM system will have flat tops near the peak values. This produces *overload noise.* The flat tops are easily seen on an oscilloscope, and the recovered analog waveform sounds distorted since the flat topping

produces unwanted harmonic components. For example, this type of distortion can be heard on PCM telephone systems when there are high levels such as dial tones, busy signals, or off-hook warning signals.

The second type of noise, *random noise,* is produced by the random quantization errors in the PCM system under normal operating conditions when the input level is properly set. This type of condition is assumed in Eq. (3.8). Random noise has a white hissing sound. If the input level is not sufficiently large, the S/N will deteriorate from that given by Eq. (3.8); the quantizing noise will still remain more or less random.

If the input level is reduced further to a relatively small value with respect to the design level, the error values are not equally likely from sample to sample, and the noise has a harsh sound resembling gravel being poured into a barrel. This is called *granular noise.* This type of noise can be randomized (noise power decreased) by increasing the number of quantization levels and, consequently, increasing the PCM bit rate. Alternatively, granular noise can be reduced by using a nonuniform quantizer, such as the μ-law or A-law quantizers that are described in Section 3.7.

The fourth type of quantizing noise that may occur at the output of a PCM system is *hunting noise.* It can occur when the input analog waveform is nearly constant, including when there is no signal (i.e., zero level). For these conditions the sample values at the quantizer output (see Fig. 3.2) can oscillate between two adjacent quantization levels, causing an undesired sinusoidal type tone of frequency $1/2f_s$ at the output of the PCM system. Hunting noise can be reduced by filtering out the tone or by designing the quantizer so that there is no vertical step at the constant value of the inputs, such as at 0-V input for the no signal case. For the no signal case, the hunting noise is also called *idle channel noise.* Idle channel noise can be reduced by using a horizontal step at the origin of the quantizer output–input characteristic instead of a vertical step as shown in Fig. 3.2.

Recalling that $M = 2^n$, we may express Eq. (3.8) in decibels by taking $10 \log_{10}(\cdot)$ of both sides of the equation,

$$\left(\frac{S}{N}\right)_{\text{dB}} = 6.02n + \alpha \tag{3.9}$$

where n is the number of bits in the PCM word and $\alpha = 0$. This equation—called the 6-dB rule—points out the significant performance characteristic for PCM: an additional 6-dB improvement in S/N is obtained for each bit added to the PCM word. This is illustrated in Table 3.2. Equation (3.9) is valid for a wide variety of assumptions (such as various types of input waveshapes and quantification characteristics), although the value of α will depend on these assumptions [7]. Of course, it is assumed that there are no bit errors and that the input signal level is large enough to range over a significant number of quantizing levels.

One may use Table 3.2 to examine the design requirements in a proposed PCM system. For example, high fidelity enthusiasts are turning to digital audio recording techniques. Here PCM signals are recorded instead of the analog audio signal to produce superb sound reproduction. For a dynamic range of 90 dB, it is seen that at least 15-b PCM words would be required. Furthermore, if the analog signal had a bandwidth of 20 kHz, the first null bandwidth for rectangular bit-shape PCM would be 2×20 kHz $\times 15 = 600$ kHz. Consequently, video-type tape recorders are needed to record and reproduce high-quality digital audio signals. Although this type of recording technique might seem ridiculous at first, it is realized that expensive high-quality analog recording devices are hard pressed to reproduce a dynamic range of 70 dB. Thus, digital audio is one way to achieve improved performance. This is being proven in the marketplace with the popularity of the digital compact disk (CD). The CD uses a 16-b PCM word and a sampling rate of 44.1 kHz on

Ford Motor Co.
Exhibit 1030
Page 137

each stereo channel [9, 10]. Reed–Solomon coding with interleaving is used to correct burst errors that occur as a result of scratches and fingerprints on the compact disk.

3.7 Nonuniform Quantizing: μ-Law and A-Law Companding

Voice analog signals are more likely to have amplitude values near zero than at the extreme peak values allowed. For example, when digitizing voice signals, if the peak value allowed is 1 V, weak passages may have voltage levels on the order of 0.1 V (20 dB down). For signals such as these with nonuniform amplitude distribution, the granular quantizing noise will be a serious problem if the step size is not reduced for amplitude values near zero and increased for extremely large values. This is called nonuniform quantizing since a variable step size is used. An example of a nonuniform quantizing characteristic is shown in Fig. 3.3.

The effect of nonuniform quantizing can be obtained by first passing the analog signal through a compression (nonlinear) amplifier and then into the PCM circuit that uses a uniform quantizer. In the U.S., a μ-law type of compression characteristic is used. It is defined [11] by

$$|w_2(t)| = \frac{\ln(1 + \mu |w_1(t)|)}{\ln(1 + \mu)} \tag{3.10}$$

where the allowed peak values of $w_1(t)$ are ± 1 (i.e., $|w_1(t)| \leq 1$), μ is a positive constant that is a parameter. This compression characteristic is shown in Fig. 3.3(b) for several values of μ, and it is noted that $\mu \to 0$ corresponds to linear amplification (uniform quantization overall). In the United States, Canada, and Japan, the telephone companies use a $\mu = 255$ compression characteristic in their PCM systems [4].

Another compression law, used mainly in Europe, is the A-law characteristic. It is defined [1] by

$$|w_2(t)| = \begin{cases} \dfrac{A|w_1(t)|}{1 + \ln A}, & 0 \leq |w_1(t)| \leq \dfrac{1}{A} \\ \dfrac{1 + \ln(A|w_1(t)|)}{1 + \ln A}, & \dfrac{1}{A} \leq |w_1(t)| \leq 1 \end{cases} \tag{3.11}$$

where $|w_1(t)| < 1$ and A is a positive constant. The A-law compression characteristic is shown in Fig. 3.3(c). The typical value for A is 87.6.

When compression is used at the transmitter, *expansion* (i.e., decompression) must be used at the receiver output to restore signal levels to their correct relative values. The *expandor* characteristic is the inverse of the compression characteristic, and the combination of a compressor and an expandor is called a *compandor*.

Once again, it can be shown that the output S/N follows the 6-dB law [2]

$$\left(\frac{S}{N}\right)_{\text{dB}} = 6.02 + \alpha \tag{3.12}$$

where for uniform quantizing

$$\alpha = 4.77 - 20\log(V/x_{\text{rms}}) \tag{3.13}$$

Ford Motor Co.
Exhibit 1030
Page 138

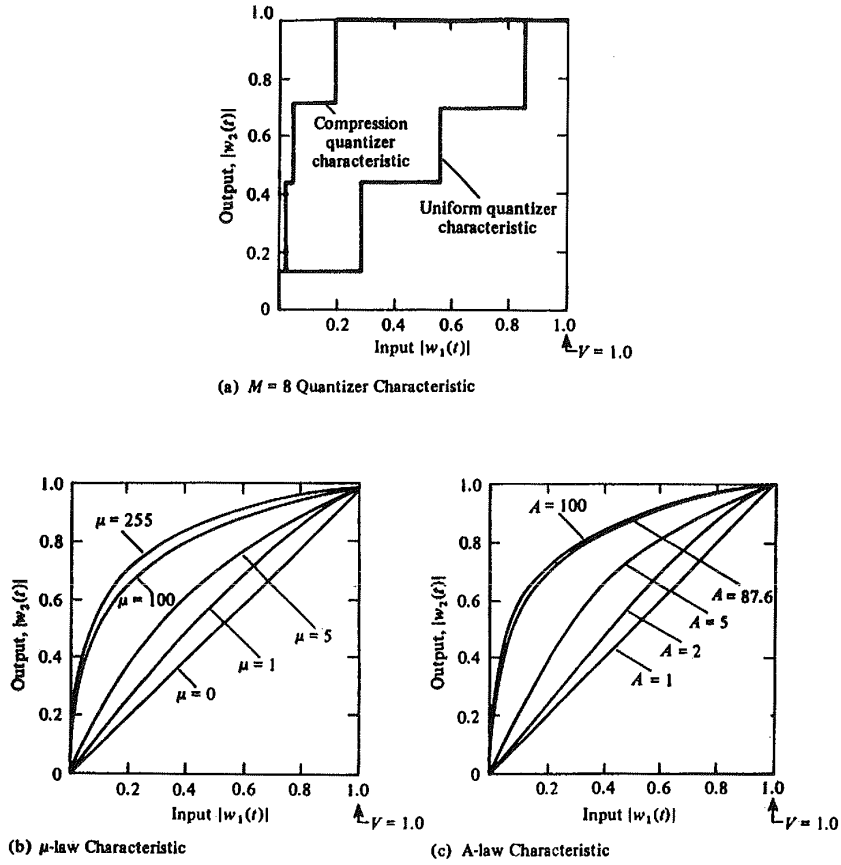


FIGURE 3.3 Compression characteristics (first quadrant shown). *Source:* Couch, L.W. II 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 147. With permission.

and for sufficiently large input levels[2] for μ-law companding

$$\alpha \approx 4.77 - 20\log[\ln(1+\mu)] \tag{3.14}$$

and for A-law companding [7]

$$\alpha \approx 4.77 - 20\log[1+\ln A] \tag{3.15}$$

n is the number of bits used in the PCM word, V is the peak design level of the quantizer, and x_{rms} is the rms value of the input analog signal. Notice that the output S/N is a function of the input level for the uniform quantizing (no companding) case but is relatively insensitive to input level for μ-law and A-law companding, as shown in Fig. 3.4. The ratio V/x_{rms} is called the *loading factor.* The input level is often set for a loading factor of 4 (12 dB) to

[2]See Lathi, 1998 for a more complicated expression that is valid for any input level.

Ford Motor Co.
Exhibit 1030
Page 139

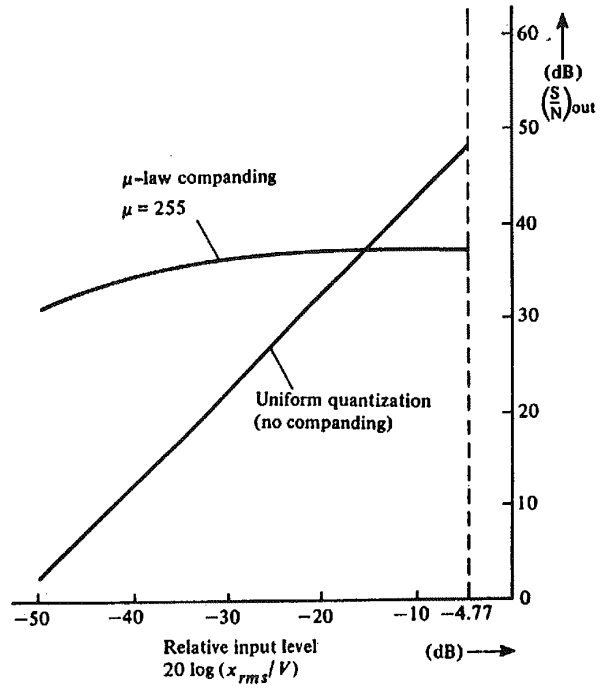


FIGURE 3.4 Output S/N of 8-b PCM systems with and without companding. *Source:* Couch, L.W. II 1997. *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, p. 149. With permission.

ensure that the overload quantizing noise will be negligible. In practice this gives $\alpha = -7.3$ for the case of uniform encoding as compared to $\alpha = 0$, which was obtained for the ideal conditions associated with Eq. (3.8).

3.8 Example: Design of a PCM System

Assume that an analog voice-frequency signal, which occupies a band from 300 to 3400 Hz, is to be transmitted over a binary PCM system. The minimum sampling frequency would be $2 \times 3.4 = 6.8$ kHz. In practice the signal is oversampled, and in the U.S. a sampling frequency of 8 kHz is the standard used for voice-frequency signals in telephone communication systems. Assume that each sample value is represented by 8 b; then the bit rate of the PCM signal is

$$\begin{aligned} R &= (f_s \text{ samples}/s)\,(n \text{ b}/s) \\ &= (8\ k \text{ samples}/s)(8 \text{ b}/s) = 64 \text{ kb}/s \end{aligned} \tag{3.16}$$

Referring to the dimensionality theorem [Eq. (3.4)], we realize that the theoretically minimum absolute bandwidth of the PCM signal is

$$B_{\min} = \frac{1}{2}D = 32 \text{ kHz} \tag{3.17}$$

and this is realized if the PCM waveform consists of $(\sin x)/x$ pulse shapes. If rectangular pulse shaping is used, the absolute bandwidth is infinity, and the first null bandwidth

[Eq. (3.5)] is

$$B_{\text{null}} = R = \frac{1}{T_b} = 64 \text{ kHz} \tag{3.18}$$

That is, we require a bandwidth of 64 kHz to transmit this digital voice PCM signal where the bandwidth of the original analog voice signal was, at most, 4 kHz. Using $n = 8$ in Eq. (3.1), the error on the recovered analog signal is $\pm 0.2\%$. Using Eqs. (3.12) and (3.13) for the case of uniform quantizing with a loading factor, V/x_{rms}, of 10 (20 dB), we get for uniform quantizing

$$\left(\frac{S}{N}\right)_{\text{dB}} = 32.9 \text{ dB} \tag{3.19}$$

Using Eqs. (3.12) and (3.14) for the case of $\mu = 255$ companding, we get

$$\left(\frac{S}{N}\right) = 38.05 \text{ dB} \tag{3.20}$$

These results are illustrated in Fig. 3.4.

Defining Terms

Intersymbol interference: Filtering of a digital waveform so that a pulse corresponding to 1 b will smear (stretch in width) into adjacent bit slots.

Pulse amplitude modulation: An analog signal is represented by a train of pulses where the pulse amplitudes are proportional to the analog signal amplitude.

Pulse code modulation: A serial bit stream that consists of binary words which represent quantized sample values of an analog signal.

Quantizing: Replacing a sample value with the closest allowed value.

References

[1] Cattermole, K.W., *Principles of Pulse-code Modulation,* American Elsevier, New York, NY, 1969.

[2] Couch, L.W., *Digital and Analog Communication Systems,* 5th ed., Prentice Hall, Upper Saddle River, NJ, 1997.

[3] Couch, L.W., *Modern Communication Systems: Principles and Applications,* Macmillan Publishing, New York, NY, 1995.

[4] Dammann, C.L., McDaniel, L.D., and Maddox, C.L., D2 Channel Bank—Multiplexing and Coding. *B. S. T. J.,* 12(10), 1675–1700, 1972.

[5] Dorf, R.C., *The Electrical Engineering Handbook,* CRC Press, Inc., Boca Raton, FL, 1993.

[6] Gibson, J.D., *The Communications Handbook,* CRC Press, Inc., Boca Raton, FL, 1997.

[7] Jayant, N.S. and Noll, P., *Digital Coding of Waveforms,* Prentice Hall, Englewood Cliffs, NJ, 1984.

[8] Lathi, B.P., *Modern Digital and Analog Communication Systems,* 3rd ed., Oxford University Press, New York, NY, 1998.

[9] Miyaoka, S., Digital Audio is Compact and Rugged. *IEEE Spectrum,* 21(3), 35–39, 1984.

[10] Peek, J.B.H., Communication Aspects of the Compact Disk Digital Audio System. *IEEE Comm. Mag.,* 23(2), 7–15, 1985.

Ford Motor Co.
Exhibit 1030
Page 141

[11] Smith, B., Instantaneous Companding of Quantized Signals. *B. S. T. J.*, 36(5), 653–709, 1957.

[12] Spilker, J.J., *Digital Communications by Satellite,* Prentice Hall, Englewood Cliffs, NJ, 1977.

Further Information

Many practical design situations and applications of PCM transmission via twisted-pair T-1 telephone lines, fiber optic cable, microwave relay, and satellite systems are given in [2] and [3].

Ford Motor Co.
Exhibit 1030
Page 142

4

Baseband Signalling and Pulse Shaping

4.1 Communications System Model 4-1
4.2 Intersymbol Interference and the Nyquist Criterion ... 4-4
Raised Cosine Pulse
4.3 Nyquist Criterion with Matched Filtering 4-8
4.4 Eye Diagrams 4-10
Vertical Eye Opening • Horizontal Eye Opening • Slope of the Inner Eye
4.5 Partial-Response Signalling 4-12
Precoding
4.6 Additional Considerations 4-18
Average Transmitted Power and Spectral Constraints • Peak-to-Average Power • Channel and Receiver Characteristics • Complexity • Tolerance to Interference • Probability of Intercept and Detection
4.7 Examples 4-20
Global System for Mobile Communications (GSM) • U.S. Digital Cellular (IS-136) • Interim Standard-95 • Personal Access Communications System (PACS)
Defining Terms 4-21
References 4-22
Further Information 4-23

Michael L. Honig
Northwestern University

Melbourne Barton
Bellcore

Many physical communications channels, such as radio channels, accept a continuous-time waveform as input. Consequently, a sequence of source bits, representing data or a digitized analog signal, must be converted to a continuous-time waveform at the transmitter. In general, each successive group of bits taken from this sequence is mapped to a particular continuous-time pulse. In this chapter we discuss the basic principles involved in selecting such a pulse for channels that can be characterized as linear and time invariant with finite bandwidth.

4.1 Communications System Model

Figure 4.1a shows a simple block diagram of a communications system. The sequence of source bits $\{b_i\}$ are grouped into sequential blocks (vectors) of m bits $\{\boldsymbol{b}_i\}$, and each binary vector $\boldsymbol{b}_i$ is mapped to one of 2^m pulses, $p(\boldsymbol{b}_i; t)$, which is transmitted over the channel.

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 143

The transmitted signal as a function of time can be written as

$$s(t) = \sum_i p(\boldsymbol{b}_i; t - iT) \tag{4.1}$$

where $1/T$ is the rate at which each group of m bits, or pulses, is introduced to the channel. The information (bit) rate is therefore m/T.

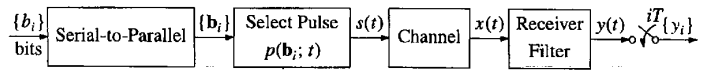


Figure 4.1a Communication system model. The source bits are grouped into binary vectors, which are mapped to a sequence of pulse shapes.

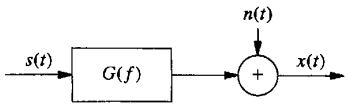


Figure 4.1b Channel model consisting of a linear, time-invariant system (transfer function) followed by additive noise.

The channel in Fig. 4.1a can be a radio link, which may distort the input signal $s(t)$ in a variety of ways. For example, it may introduce pulse dispersion (due to finite bandwidth) and multipath, as well as additive background noise. The output of the channel is denoted as $x(t)$, which is processed by the receiver to determine estimates of the source bits. The receiver can be quite complicated; however, for the purpose of this discussion, it is sufficient to assume only that it contains a front-end filter and a sampler, as shown in Fig. 4.1a. This assumption is valid for a wide variety of detection strategies. The purpose of the receiver filter is to remove noise outside of the transmitted frequency band and to compensate for the channel frequency response.

A commonly used channel model is shown in Fig. 4.1b and consists of a linear, time-invariant filter, denoted as $G(f)$, followed by additive noise $n(t)$. The channel output is, therefore,

$$x(t) = [g(t) * s(t)] + n(t) \tag{4.2}$$

where $g(t)$ is the channel impulse response associated with $G(f)$, and the asterisk denotes convolution,

$$g(t) * s(t) = \int_{-\infty}^{\infty} g(t - \tau)s(\tau)\,\mathrm{d}\tau$$

This channel model accounts for all linear, time-invariant channel impairments, such as finite bandwidth and time-invariant multipath. It does not account for time-varying impairments, such as rapid fading due to time-varying multipath. Nevertheless, this model can be considered valid over short time periods during which the multipath parameters remain constant.

In Figs. 4.1a, and 4.1b, it is assumed that all signals are **baseband signals**, which means that the frequency content is centered around $f = 0$ (DC). The channel passband, therefore, partially coincides with the transmitted spectrum. In general, this condition requires that the transmitted signal be modulated by an appropriate carrier frequency and demodulated at the receiver. In that case, the model in Figs. 4.1a, and 4.1b still applies; however, *baseband-equivalent* signals must be derived from their modulated (passband) counterparts.

Ford Motor Co.
Exhibit 1030
Page 144

Baseband signalling and *pulse shaping* refers to the way in which a group of source bits is mapped to a baseband transmitted pulse.

As a simple example of baseband signalling, we can take $m = 1$ (map each source bit to a pulse), assign a 0 bit to a pulse $p(t)$, and a 1 bit to the pulse $-p(t)$. Perhaps the simplest example of a baseband pulse is the *rectangular* pulse given by $p(t) = 1$, $0 < t \leq T$, and $p(t) = 0$ elsewhere. In this case, we can write the transmitted signal as

$$s(t) = \sum_i A_i p(t - iT) \tag{4.3}$$

where each symbol A_i takes on a value of $+1$ or -1, depending on the value of the ith bit, and $1/T$ is the *symbol rate*, namely, the rate at which the symbols A_i are introduced to the channel.

The preceding example is called *binary* **pulse amplitude modulation (PAM)**, since the data symbols A_i are binary valued, and they amplitude modulate the transmitted pulse $p(t)$. The information rate (bits per second) in this case is the same as the symbol rate $1/T$. As a simple extension of this signalling technique, we can increase m and choose A_i from one of $M = 2^m$ values to transmit at bit rate m/T. This is known as M-ary PAM. For example, letting $m = 2$, each pair of bits can be mapped to a pulse in the set $\{p(t), -p(t), 3p(t), -3p(t)\}$.

In general, the transmitted symbols $\{A_i\}$, the baseband pulse $p(t)$, and channel impulse response $g(t)$ can be *complex valued.* For example, each successive pair of bits might select a symbol from the set $\{1, -1, j, -j\}$, where $j = \sqrt{-1}$. This is a consequence of considering the baseband equivalent of passband modulation. (That is, generating a transmitted spectrum which is centered around a carrier frequency f_c.) Here we are not concerned with the relation between the passband and baseband equivalent models and simply point out that the discussion and results in this chapter apply to complex-valued symbols and pulse shapes.

As an example of a signalling technique which is not PAM, let $m = 1$ and

$$\begin{aligned} p(0;t) &= \begin{cases} \sqrt{2}\sin(2\pi f_1 t) & 0 < t < T \\ 0 & \text{elsewhere} \end{cases} \\ p(1;t) &= \begin{cases} \sqrt{2}\sin(2\pi f_2 t) & 0 < t < T \\ 0 & \text{elsewhere} \end{cases} \end{aligned} \tag{4.4}$$

where f_1 and $f_2 \neq f_1$ are fixed frequencies selected so that $f_1 T$ and $f_2 T$ (number of cycles for each bit) are multiples of $1/2$. These pulses are *orthogonal*, namely,

$$\int_0^T p(1;t)p(0;t)\,\mathrm{d}t = 0$$

This choice of pulse shapes is called binary **frequency-shift keying (FSK)**.

Another example of a set of orthogonal pulse shapes for $m = 2$ bits/T is shown in Fig. 4.2. Because these pulses may have as many as three transitions within a symbol period, the transmitted spectrum occupies roughly four times the transmitted spectrum of binary PAM with a rectangular pulse shape. The spectrum is, therefore, spread across a much larger band than the smallest required for reliable transmission, assuming a data rate of $2/T$. This type of signalling is referred to as **spread-spectrum**. Spread-spectrum signals are more

Ford Motor Co.
Exhibit 1030
Page 145

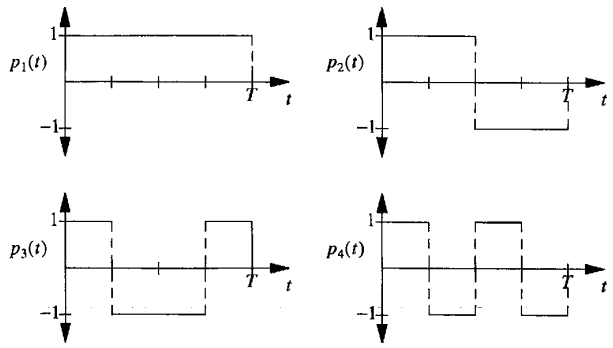


FIGURE 4.2 Four orthogonal spread-spectrum pulse shapes.

robust with respect to interference from other transmitted signals than are narrowband signals.[1]

4.2 Intersymbol Interference and the Nyquist Criterion

Consider the transmission of a PAM signal illustrated in Fig. 4.3. The source bits $\{b_i\}$ are mapped to a sequence of levels $\{A_i\}$, which modulate the transmitter pulse $p(t)$. The channel input is, therefore, given by Eq. (4.3) where $p(t)$ is the impulse response of the transmitter *pulse-shaping filter* $P(f)$ shown in Fig. 4.3. The input to the transmitter filter $P(f)$ is the modulated sequence of delta functions $\sum_i A_i\delta(t-iT)$. The channel is represented by the transfer function $G(f)$ (plus noise), which has impulse response $g(t)$, and the receiver filter has transfer function $R(f)$ with associated impulse response $r(t)$.

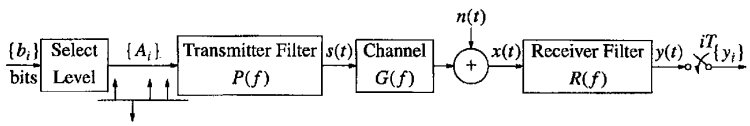


FIGURE 4.3 Baseband model of a pulse amplitude modulation system.

Let $h(t)$ be the overall impulse response of the combined transmitter, channel, and receiver, which has transfer function $H(f) = P(f)G(f)R(f)$. We can write $h(t) = p(t) * g(t) * r(t)$. The output of the receiver filter is then

$$y(t) = \sum_i A_i h(t - iT) + \tilde{n}(t) \tag{4.5}$$

where $\tilde{n}(t) = r(t) * n(t)$ is the output of the filter $R(f)$ with input $n(t)$. Assuming that samples are collected at the output of the filter $R(f)$ at the symbol rate $1/T$, we can write

[1] This example can also be viewed as coded binary PAM. Namely, each pair of two source bits are mapped to 4 coded bits, which are transmitted via binary PAM with a rectangular pulse. The current IS-95 air interface uses an extension of this signalling method in which groups of 6 b are mapped to 64 orthogonal pulse shapes with as many as 63 transitions during a symbol.

Ford Motor Co.
Exhibit 1030
Page 146

the kth sample of $y(t)$ as

$$\begin{aligned} y(kT) &= \sum_i A_i h(kT - iT) + \tilde{n}(kT) \\ &= A_k h(0) + \sum_{i \neq k} A_i h(kT - iT) + \tilde{n}(kT) \end{aligned} \tag{4.6}$$

The first term on the right-hand side of Eq. (4.6) is the kth transmitted symbol scaled by the system impulse response at $t = 0$. If this were the only term on the right side of Eq. (4.6), we could obtain the source bits without error by scaling the received samples by $1/h(0)$. The second term on the right-hand side of Eq. (4.6) is called **intersymbol interference**, which reflects the view that neighboring symbols interfere with the detection of each desired symbol.

One possible criterion for choosing the transmitter and receiver filters is to minimize intersymbol interference. Specifically, if we choose $p(t)$ and $r(t)$ so that

$$h(kT) = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases} \tag{4.7}$$

then the kth received sample is

$$y(kT) = A_k + \tilde{n}(kT) \tag{4.8}$$

In this case, the intersymbol interference has been eliminated. This choice of $p(t)$ and $r(t)$ is called a **zero-forcing** solution, since it forces the intersymbol interference to zero. Depending on the type of detection scheme used, a zero-forcing solution may not be desirable. This is because the probability of error also depends on the noise intensity, which generally increases when intersymbol interference is suppressed. It is instructive, however, to examine the properties of the zero-forcing solution.

We now view Eq. (4.7) in the frequency domain. Since $h(t)$ has Fourier transform

$$H(f) = P(f)G(f)R(f) \tag{4.9}$$

where $P(f)$ is the Fourier transform of $p(t)$, the bandwidth of $H(f)$ is limited by the bandwidth of the channel $G(f)$. We will assume that $G(f) = 0$, $|f| > W$. The sampled impulse response $h(kT)$ can, therefore, be written as the inverse Fourier transform

$$h(kT) = \int_{-W}^{W} H(f) e^{j2\pi f kT} \, \mathrm{d}f$$

Through a series of manipulations, this integral can be rewritten as an inverse discrete Fourier transform,

$$h(kT) = T \int_{-1/(2T)}^{1/(2T)} H_{\text{eq}}\left(e^{j2\pi fT}\right) e^{j2\pi f kT} \, \mathrm{d}f \tag{4.10a}$$

where

$$\begin{aligned} H_{\text{eq}}(e^{j2\pi fT}) &= \frac{1}{T} \sum_k H\left(f + \frac{k}{T}\right) \\ &= \frac{1}{T} \sum_k P\left(f + \frac{k}{T}\right) G\left(f + \frac{k}{T}\right) R\left(f + \frac{k}{T}\right) \end{aligned} \tag{4.10b}$$

This relation states that $H_{\text{eq}}(z), z = e^{j2\pi fT}$, is the discrete Fourier transform of the sequence $\{h_k\}$, where $h_k = h(kT)$. Sampling the impulse response $h(t)$ therefore changes the transfer function $H(f)$ to the *aliased* frequency response $H_{\text{eq}}(e^{j2\pi fT})$. From Eqs. (4.10), and (4.6) we conclude that $H_{\text{eq}}(z)$ is the transfer function that relates the sequence of input data symbols $\{A_i\}$ to the sequence of received samples $\{y_i\}$, where $y_i = y(iT)$, in the absence of noise. This is illustrated in Fig. 4.4. For this reason, $H_{\text{eq}}(z)$ is called the **equivalent discrete-time transfer function** for the overall system transfer function $H(f)$.

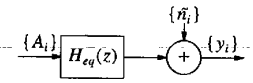


FIGURE 4.4 Equivalent discrete-time channel for the PAM system shown in Fig. 4.3 $[y_i = y(iT), \tilde{n}_i = \tilde{n}(iT)]$

Since $H_{\text{eq}}(e^{j2\pi fT})$ is the discrete Fourier transform of the sequence $\{h_k\}$, the time-domain, or sequence condition (4.7) is equivalent to the frequency-domain condition

$$H_{\text{eq}}\left(e^{j2\pi fT}\right) = 1 \tag{4.11}$$

This relation is called the **Nyquist criterion**. From Eqs. (4.10b) and (4.11) we make the following observations.

1. To satisfy the Nyquist criterion, the channel bandwidth W must be at least $1/(2T)$. Otherwise, $G(f + n/T) = 0$ for f in some interval of positive length for all n, which implies that $H_{\text{eq}}(e^{j2\pi fT}) = 0$ for f in the same interval.
2. For the minimum bandwidth $W = 1/(2T)$, Eqs. (4.10b) and (4.11) imply that $H(f) = T$ for $|f| < 1/(2T)$ and $H(f) = 0$ elsewhere. This implies that the system impulse response is given by

$$h(t) = \frac{\sin(\pi t/T)}{\pi t/T} \tag{4.12}$$

(Since $\int_{-\infty}^{\infty} h^2(t)\,dt = T$, the transmitted signal $s(t) = \sum_i A_i h(t - iT)$ has power equal to the symbol variance $E[|A_i|^2]$.) The impulse response in Eq. (4.12) is called a *minimum bandwidth* or Nyquist pulse. The frequency band $[-1/(2T), 1/(2T)]$ [i.e., the passband of $H(f)$] is called the **Nyquist band**.
3. Suppose that the channel is bandlimited to twice the Nyquist bandwidth. That is, $G(f) = 0$ for $|f| > 1/T$. The condition (4.11) then becomes

$$H(f) + H\left(f - \frac{1}{T}\right) + H\left(f + \frac{1}{T}\right) = T \tag{4.13}$$

Assume for the moment that $H(f)$ and $h(t)$ are both real valued, so that $H(f)$ is an even function of $f[H(f) = H(-f)]$. This is the case when the receiver filter is the matched filter (see Section 4.3). We can then rewrite Eq. (4.13) as

$$H(f) + H\left(\frac{1}{T} - f\right) = T, \quad 0 < f < \frac{1}{2T} \tag{4.14}$$

which states that $H(f)$ must have odd symmetry about $f = 1/(2T)$. This is illustrated in Fig. 4.5, which shows two different transfer functions $H(f)$ that

Ford Motor Co.
Exhibit 1030
Page 148

satisfy the Nyquist criterion.

4. The pulse shape $p(t)$ enters into Eq. (4.11) only through the product $P(f)R(f)$. Consequently, either $P(f)$ or $R(f)$ can be fixed, and the other filter can be adjusted or adapted to the particular channel. Typically, the pulse shape $p(t)$ is fixed, and the receiver filter is adapted to the (possibly time-varying) channel.

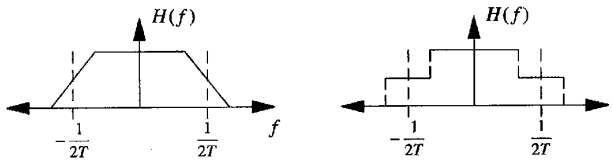


FIGURE 4.5 Two examples of frequency responses that satisfy the Nyquist criterion.

4.2.1 Raised Cosine Pulse

Suppose that the channel is ideal with transfer function

$$G(f) = \begin{cases} 1, & |f| < W \\ 0, & |f| > W \end{cases} \tag{4.15}$$

To maximize bandwidth efficiency, Nyquist pulses given by Eq. (4.12) should be used where $W = 1/(2T)$. This type of signalling, however, has two major drawbacks. First, Nyquist pulses are noncausal and of infinite duration. They can be approximated in practice by introducing an appropriate delay, and truncating the pulse. The pulse, however, decays very slowly, namely, as $1/t$, so that the truncation window must be wide. This is equivalent to observing that the ideal bandlimited frequency response given by Eq. (4.15) is difficult to approximate closely. The second drawback, which is more important, is the fact that this type of signalling is not robust with respect to sampling jitter. Namely, a small sampling offset ε produces the output sample

$$y(kT + \varepsilon) = \sum_i A_i \frac{\sin[\pi(k - i + \varepsilon/T)]}{\pi(k - i + \varepsilon/T)} \tag{4.16}$$

Since the Nyquist pulse decays as $1/t$, this sum is not guaranteed to converge. A particular choice of symbols $\{A_i\}$ can, therefore, lead to very large intersymbol interference, no matter how small the offset. Minimum bandwidth signalling is therefore impractical.

The preceding problem is generally solved in one of two ways in practice:

1. The pulse bandwidth is increased to provide a faster pulse decay than $1/t$.
2. A *controlled* amount of intersymbol interference is introduced at the transmitter, which can be subtracted out at the receiver.

The former approach sacrifices bandwidth efficiency, whereas the latter approach sacrifices power efficiency. We will examine the latter approach in Section 4.5. The most common example of a pulse, which illustrates the first technique, is the **raised cosine pulse**, given by

$$h(t) = \left[\frac{\sin(\pi t/T)}{\pi t/T}\right]\left[\frac{\cos(\alpha\pi t/T)}{1 - (2\alpha t/T)^2}\right] \tag{4.17}$$

Ford Motor Co.
Exhibit 1030
Page 149

which has Fourier transform

$$H(f) = \begin{cases} T & 0 \le |f| \le \dfrac{1-\alpha}{2T} \\ \dfrac{T}{2}\left\{1 + \cos\left[\dfrac{\pi T}{\alpha}\left(|f| - \dfrac{1-\alpha}{2T}\right)\right]\right\} & \dfrac{1-\alpha}{2T} \le |f| \le \dfrac{1+\alpha}{2T} \\ 0 & |f| > \dfrac{1+\alpha}{2T} \end{cases} \tag{4.18}$$

where $0 \le \alpha \le 1$.

Plots of $p(t)$ and $P(f)$ are shown in Figs. 4.6a, and 4.6b for different values of α. It is easily verified that $h(t)$ satisfies the Nyquist criterion (4.7) and, consequently, $H(f)$ satisfies Eq. (4.11). When $\alpha = 0$, $H(f)$ is the Nyquist pulse with minimum bandwidth $1/(2T)$, and when $\alpha > 0, H(f)$ has bandwidth $(1+\alpha)/(2T)$ with a raised cosine rolloff. The parameter α, therefore, represents the additional, or **excess bandwidth** as a fraction of the minimum bandwidth $1/(2T)$. For example, when $\alpha = 1$, we say that the pulse is a raised cosine pulse with 100% excess bandwidth. This is because the pulse bandwidth $1/T$ is twice the minimum bandwidth. Because the raised cosine pulse decays as $1/t^3$, performance is robust with respect to sampling offsets.

The raised cosine frequency response (4.18) applies to the combination of transmitter, channel, and receiver. If the transmitted pulse shape $p(t)$ is a raised cosine pulse, then $h(t)$ is a raised cosine pulse only if the combined receiver and channel frequency response is constant. Even with an ideal (transparent) channel, however, the optimum (matched) receiver filter response is generally not constant in the presence of additive Gaussian noise. An alternative is to transmit the *square-root raised cosine* pulse shape, which has frequency response $P(f)$ given by the square-root of the raised cosine frequency response in Eq. (4.18). Assuming an ideal channel, setting the receiver frequency response $R(f) = P(f)$ then results in an overall raised cosine system response $H(f)$.

4.3 Nyquist Criterion with Matched Filtering

Consider the transmission of an isolated pulse $A_0\delta(t)$. In this case the input to the receiver in Fig. 4.3 is

$$x(t) = A_0\tilde{g}(t) + n(t) \tag{4.19}$$

where $\tilde{g}(t)$ is the inverse Fourier transform of the combined transmitter-channel transfer function $\tilde{G}(f) = P(f)G(f)$. We will assume that the noise $n(t)$ is white with spectrum $N_0/2$. The output of the receiver filter is then

$$y(t) = r(t) * x(t) = A_0[r(t) * \tilde{g}(t)] + [r(t) * n(t)] \tag{4.20}$$

The first term on the right-hand side is the desired signal, and the second term is noise. Assuming that $y(t)$ is sampled at $t = 0$, the ratio of signal energy to noise energy, or signal-to-noise ratio (SNR) at the sampling instant, is

$$SNR = \frac{E\left[|A_0|^2\right]\left|\displaystyle\int_{-\infty}^{\infty} r(-t)\tilde{g}(t)\,\mathrm{d}t\right|^2}{\dfrac{N_0}{2}\displaystyle\int_{-\infty}^{\infty}|r(t)|^2\,\mathrm{d}t} \tag{4.21}$$

Ford Motor Co.
Exhibit 1030
Page 150

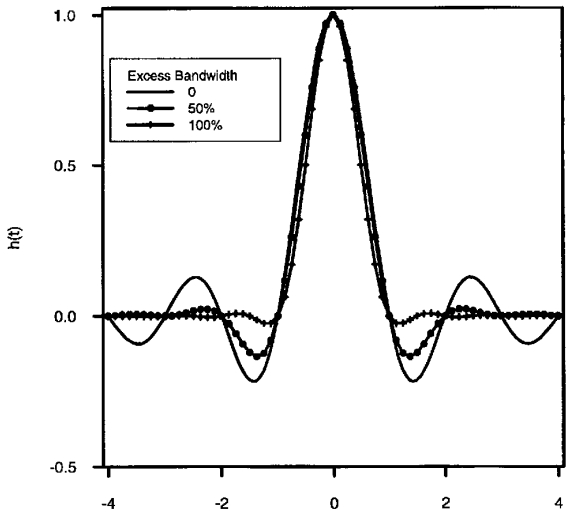


Figure 4.6a Raised cosine pulse.

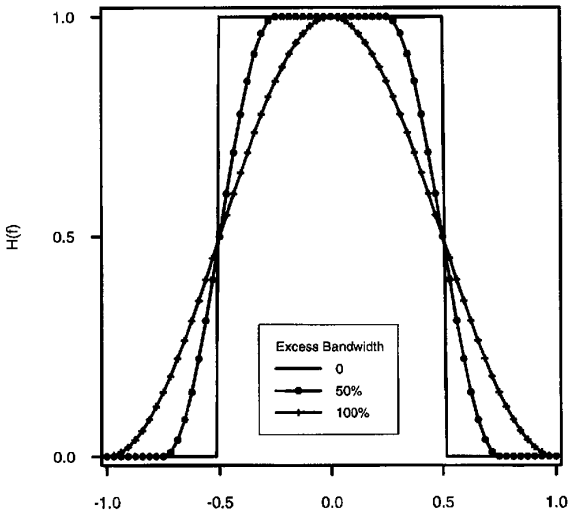


Figure 4.6b Raised cosine spectrum.

The receiver impulse response that maximizes this expression is $r(t) = \tilde{g}^*(-t)$ [complex conjugate of $\tilde{g}(-t)$], which is known as the **matched filter** impulse response. The associated transfer function is $R(f) = \tilde{G}^*(f)$.

Choosing the receiver filter to be the matched filter is optimal in more general situations, such as when detecting a sequence of channel symbols with intersymbol interference (assuming the additive noise is Gaussian). We, therefore, reconsider the Nyquist criterion when the receiver filter is the matched filter. In this case, the baseband model is shown in Fig. 4.7, and the output of the receiver filter is given by

$$y(t) = \sum_i A_i h(t - iT) + \tilde{n}(t) \tag{4.22}$$

Ford Motor Co.
Exhibit 1030
Page 151

where the baseband pulse $h(t)$ is now the impulse response of the filter with transfer function $|\tilde{G}(f)|^2 = |P(f)G(f)|^2$. This impulse response is the *autocorrelation* of the impulse response of the combined transmitter-channel filter $\tilde{G}(f)$,

$$h(t) = \int_{-\infty}^{\infty} \tilde{g}^*(s)\tilde{g}(s+t)\,\mathrm{d}s \tag{4.23}$$

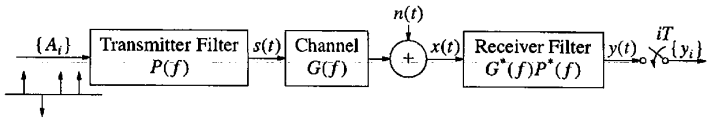


FIGURE 4.7 Baseband PAM model with a matched filter at the receiver.

With a matched filter at the receiver, the equivalent discrete-time transfer function is

$$\begin{aligned} H_{\mathrm{eq}}(e^{j2\pi fT}) &= \frac{1}{T}\sum_k \left|\tilde{G}\left(f-\frac{k}{T}\right)\right|^2 \\ &= \frac{1}{T}\sum_k \left|P\left(f-\frac{k}{T}\right)G\left(f-\frac{k}{T}\right)\right|^2 \end{aligned} \tag{4.24}$$

which relates the sequence of transmitted symbols $\{A_k\}$ to the sequence of received samples $\{y_k\}$ in the absence of noise. Note that $H_{\mathrm{eq}}(e^{j2\pi fT})$ is positive, real valued, and an even function of f. If the channel is bandlimited to twice the Nyquist bandwidth, then $H(f) = 0$ for $|f| > 1/T$, and the Nyquist condition is given by Eq. (4.14) where $H(f) = |G(f)P(f)|^2$. The aliasing sum in Eq. (4.10b) can therefore be described as a folding operation in which the channel response $|H(f)|^2$ is folded around the Nyquist frequency $1/(2T)$. For this reason, $H_{\mathrm{eq}}(e^{j2\pi fT})$ with a matched receiver filter is often referred to as the folded channel spectrum.

4.4 Eye Diagrams

One way to assess the severity of distortion due to intersymbol interference in a digital communications system is to examine the **eye diagram**. The eye diagram is illustrated in Figs. 4.8a and 4.8b, for a raised cosine pulse shape with 25% excess bandwidth and an ideal bandlimited channel. Figure 4.8a shows the data signal at the receiver

$$y(t) = \sum_i A_i h(t-iT) + \tilde{n}(t) \tag{4.25}$$

where $h(t)$ is given by Eq. (4.17), $\alpha = 1/4$, each symbol A_i is independently chosen from the set $\{\pm 1, \pm 3\}$, where each symbol is equally likely, and $\tilde{n}(t)$ is bandlimited white Gaussian noise. (The received SNR is 30 dB.) The eye diagram is constructed from the time-domain data signal $y(t)$ as follows (assuming nominal sampling times at $kT, k = 0, 1, 2, \ldots$):

1. Partition the waveform $y(t)$ into successive segments of length T starting from $t = T/2$.
2. Translate each of these waveform segments $[y(t),\ (k+1/2)T \le t \le (k+3/2)T,$ $k = 0, 1, 2, \ldots]$ to the interval $[-T/2, T/2]$, and superimpose.

Ford Motor Co.
Exhibit 1030
Page 152

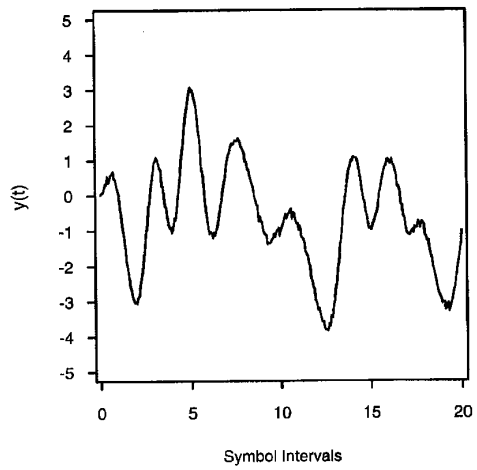


Figure 4.8a Received signal $y(t)$.

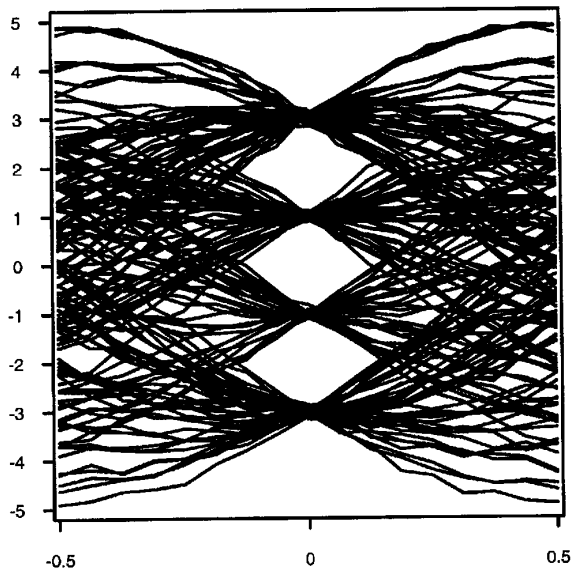


Figure 4.8b Eye diagram for received signal shown in Fig. 4.8a.

The resulting picture is shown in Fig. 4.8b for the $y(t)$ shown in Fig. 4.8a. (Partitioning $y(t)$ into successive segments of length iT, $i > 1$, is also possible. This would result in i successive eye diagrams.) The number of eye openings is one less than the number of transmitted signal levels. In practice, the eye diagram is easily viewed on an oscilloscope by applying the received waveform $y(t)$ to the vertical deflection plates of the oscilloscope and applying a sawtooth waveform at the symbol rate $1/T$ to the horizontal deflection plates. This causes successive symbol intervals to be translated into one interval on the oscilloscope display.

Each waveform segment $y(t)$, $(k + 1/2)T \leq t \leq (k + 3/2)T$, depends on the particular sequence of channel symbols surrounding A_k. The number of channel symbols that affects a

Ford Motor Co.
Exhibit 1030
Page 153

particular waveform segment depends on the extent of the intersymbol interference, shown in Eq. (4.6). This, in turn, depends on the duration of the impulse response $h(t)$. For example, if $h(t)$ has most of its energy in the interval $0 < t < mT$, then each waveform segment depends on approximately m symbols. Assuming binary transmission, this implies that there are a total of 2^m waveform segments that can be superimposed in the eye diagram. (It is possible that only one sequence of channel symbols causes significant intersymbol interference, and this sequence occurs with very low probability.) In current digital wireless applications the impulse response typically spans only a few symbols.

The eye diagram has the following important features which measure the performance of a digital communications system.

4.4.1 Vertical Eye Opening

The vertical openings at any time t_0, $-T/2 \leq t_0 \leq T/2$, represent the separation between signal levels with worst-case intersymbol interference, assuming that $y(t)$ is sampled at times $t = kT + t_0$, $k = 0, 1, 2, \ldots$. It is possible for the intersymbol interference to be large enough so that this vertical opening between some, or all, signal levels disappears altogether. In that case, the eye is said to be closed. Otherwise, the eye is said to be open. A closed eye implies that if the estimated bits are obtained by thresholding the samples $y(kT)$, then the decisions will depend primarily on the intersymbol interference rather than on the desired symbol. The probability of error will, therefore, be close to $1/2$. Conversely, wide vertical spacings between signal levels imply a large degree of immunity to additive noise. In general, $y(t)$ should be sampled at the times $kT + t_0$, $k = 0, 1, 2, \ldots$, where t_0 is chosen to maximize the vertical eye opening.

4.4.2 Horizontal Eye Opening

The width of each opening indicates the sensitivity to timing offset. Specifically, a very narrow eye opening indicates that a small timing offset will result in sampling where the eye is closed. Conversely, a wide horizontal opening indicates that a large timing offset can be tolerated, although the error probability will depend on the vertical opening.

4.4.3 Slope of the Inner Eye

The slope of the inner eye indicates sensitivity to timing jitter or variance in the timing offset. Specifically, a very steep slope means that the eye closes rapidly as the timing offset increases. In this case, a significant amount of jitter in the sampling times significantly increases the probability of error.

The shape of the eye diagram is determined by the pulse shape. In general, the faster the baseband pulse decays, the wider the eye opening. For example, a rectangular pulse produces a box-shaped eye diagram (assuming binary signalling). The minimum bandwidth pulse shape Eq. (4.12) produces an eye diagram which is closed for all t except for $t = 0$. This is because, as shown earlier, an arbitrarily small timing offset can lead to an intersymbol interference term that is arbitrarily large, depending on the data sequence.

4.5 Partial-Response Signalling

To avoid the problems associated with Nyquist signalling over an ideal bandlimited channel, bandwidth and/or power efficiency must be compromised. Raised cosine pulses compromise

Ford Motor Co.
Exhibit 1030
Page 154

bandwidth efficiency to gain robustness with respect to timing errors. Another possibility is to introduce a controlled amount of intersymbol interference at the transmitter, which can be removed at the receiver. This approach is called **partial-response (PR) signalling**. The terminology reflects the fact that the sampled system impulse response does not have the full response given by the Nyquist condition Eq. (4.7).

To illustrate PR signalling, suppose that the Nyquist condition Eq. (4.7) is replaced by the condition

$$h_k = \begin{cases} 1 & k = 0, 1 \\ 0 & \text{all other } k \end{cases} \tag{4.26}$$

The kth received sample is then

$$y_k = A_k + A_{k-1} + \tilde{n}_k \tag{4.27}$$

so that there is intersymbol interference from one neighboring transmitted symbol. For now we focus on the spectral characteristics of PR signalling and defer discussion of how to detect the transmitted sequence $\{A_k\}$ in the presence of intersymbol interference. The equivalent discrete-time transfer function in this case is the discrete Fourier transform of the sequence in Eq. (4.26),

$$\begin{aligned} H_{\text{eq}}(e^{j2\pi fT}) &= \frac{1}{T}\sum_k H\left(f + \frac{k}{T}\right) \\ &= 1 + e^{-j2\pi fT} = 2e^{-j\pi fT}\cos(\pi fT) \end{aligned} \tag{4.28}$$

As in the full-response case, for Eq. (4.28) to be satisfied, the *minimum* bandwidth of the channel $G(f)$ and transmitter filter $P(f)$ is $W = 1/(2T)$. Assuming $P(f)$ has this minimum bandwidth implies

$$H(f) = \begin{cases} 2Te^{-j\pi fT}\cos(\pi fT) & |f| < 1/(2T) \\ 0 & |f| > 1/(2T) \end{cases} \tag{4.29a}$$

and

$$h(t) = T\{\ \text{sinc}\ (t/T) + \ \text{sinc}\ [(t-T)/T]\} \tag{4.29b}$$

where sinc $x = (\sin \pi x)/(\pi x)$. This pulse is called a *duobinary* pulse and is shown along with the associated $H(f)$ in Fig. 4.9. [Notice that $h(t)$ satisfies Eq. (4.26).] Unlike the ideal bandlimited frequency response, the transfer function $H(f)$ in Eq. (4.29a) is continuous and is, therefore, easily approximated by a physically realizable filter. Duobinary PR was first proposed by Lender, [7], and later generalized by Kretzmer, [6].

The main advantage of the duobinary pulse Eq. (4.29b), relative to the minimum bandwidth pulse Eq. (4.12), is that signalling at the Nyquist symbol rate is feasible with zero excess bandwidth. Because the pulse decays much more rapidly than a Nyquist pulse, it is robust with respect to timing errors. Selecting the transmitter and receiver filters so that the overall system response is duobinary is appropriate in situations where the channel frequency response $G(f)$ is near zero or has a rapid rolloff at the Nyquist band edge $f = 1/(2T)$.

Ford Motor Co.
Exhibit 1030
Page 155

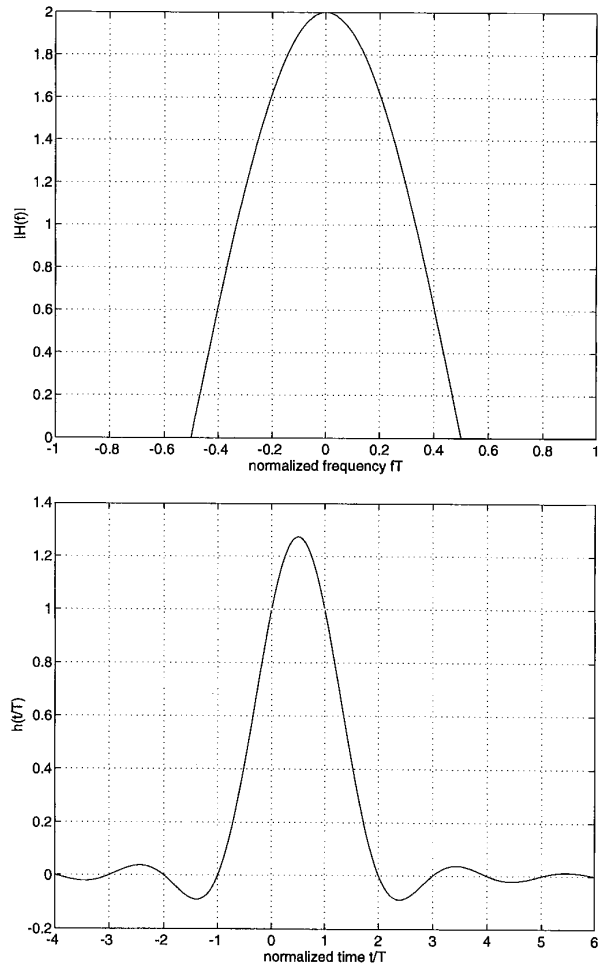


FIGURE 4.9 Duobinary frequency response and minimum bandwidth pulse.

As another example of PR signalling, consider the *modified* duobinary partial response

$$h_k = \begin{cases} 1 & k = -1 \\ -1 & k = 1 \\ 0 & \text{all other } k \end{cases} \tag{4.30}$$

which has equivalent discrete-time transfer function

$$\begin{aligned} H_{\text{eq}}\left(e^{j2\pi fT}\right) &= e^{j2\pi fT} - e^{-j2\pi fT} \\ &= j2\sin(2\pi fT) \end{aligned} \tag{4.31}$$

With zero excess bandwidth, the overall system response is

$$H(f) = \begin{cases} j2T\sin(2\pi fT) & |f| < 1/(2T) \\ 0 & |f| > 1/(2T) \end{cases} \tag{4.32a}$$

and

$$h(t) = T\{\text{sinc } [(t+T)/T] - \text{sinc } [(t-T)/T]\} \tag{4.32b}$$

These functions are plotted in Fig. 4.10. This pulse shape is appropriate when the channel response $G(f)$ is near zero at both DC ($f = 0$) and at the Nyquist band edge. This is often the case for wire (twisted-pair) channels where the transmitted signal is coupled to the channel through a transformer. Like duobinary PR, modified duobinary allows minimum bandwidth signalling at the Nyquist rate.

A particular partial response is often identified by the polynomial

$$\sum_{k=0}^{K} h_k D^k$$

where D (for delay) takes the place of the usual z^{-1} in the z transform of the sequence $\{h_k\}$. For example, duobinary is also referred to as $1 + D$ partial response.

In general, more complicated system responses than those shown in Figs. 4.9 and 4.10 can be generated by choosing more nonzero coefficients in the sequence $\{h_k\}$. This complicates detection, however, because of the additional intersymbol interference that is generated.

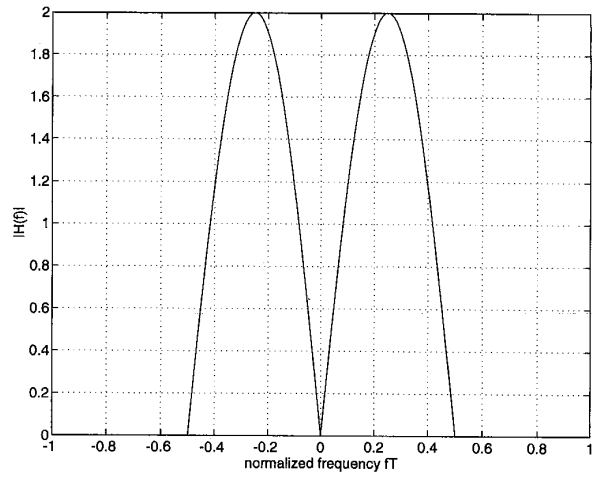
Rather than modulating a PR pulse $h(t)$, a PR signal can also be generated by filtering the sequence of transmitted levels $\{A_i\}$. This is shown in Fig. 4.11. Namely, the transmitted levels are first passed through a discrete-time (digital) filter with transfer function $P_d(e^{j2\pi fT})$ (where the subscript d indicates discrete). [Note that $P_d(e^{j2\pi fT})$ can be selected to be $H_{\text{eq}}(e^{j2\pi fT})$.] The outputs of this filter form the PAM signal, where the pulse shaping filter $P(f) = 1$, $|f| < 1/(2T)$ and is zero elsewhere. If the transmitted levels $\{A_k\}$ are selected independently and are identically distributed, then the transmitted spectrum is $\sigma_A^2|P_d(e^{j2\pi fT})|^2$ for $|f| < 1/(2T)$ and is zero for $|f| > 1/(2T)$, where $\sigma_A^2 = E[|A_k|^2]$.

Shaping the transmitted spectrum to have nulls coincident with nulls in the channel response potentially offers significant performance advantages. By introducing intersymbol interference, however, PR signalling increases the number of received signal levels, which increases the complexity of the detector and may reduce immunity to noise. For example, the set of received signal levels for duobinary signalling is $\{0, \pm 2\}$ from which the transmitted levels $\{\pm 1\}$ must be estimated. The performance of a particular PR scheme depends on the channel characteristics, as well as the type of detector used at the receiver. We now describe a simple suboptimal detection strategy.

4.5.1 Precoding

Consider the received signal sample Eq. (4.27) with duobinary signalling. If the receiver has correctly decoded the symbol A_{k-1}, then in the absence of noise A_k can be decoded by subtracting A_{k-1} from the received sample y_k. If an error occurs, however, then subtracting the preceding symbol estimate from the received sample will cause the error to propagate to successive detected symbols. To avoid this problem, the transmitted levels can be **precoded** in such a way as to compensate for the intersymbol interference introduced by the overall partial response.

Ford Motor Co.
Exhibit 1030
Page 157



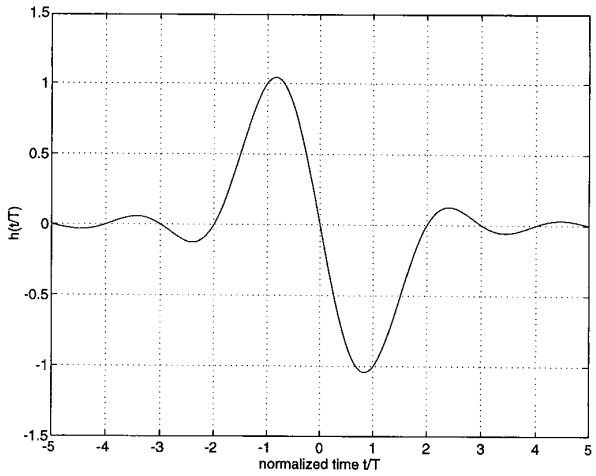


FIGURE 4.10 Modified duobinary frequency response and minimum bandwidth pulse.

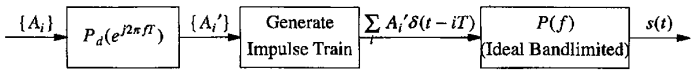


FIGURE 4.11 Generation of PR signal.

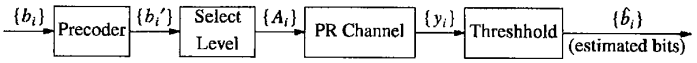


FIGURE 4.12 Precoding for a PR channel.

We first illustrate precoding for duobinary PR. The sequence of operations is illustrated in Fig. 4.12. Let $\{b_k\}$ denote the sequence of source bits where $b_k \in \{0, 1\}$. This sequence is transformed to the sequence $\{b'_k\}$ by the operation

$$b'_k = b_k \oplus b'_{k-1} \tag{4.33}$$

Ford Motor Co.
Exhibit 1030
Page 158

TABLE 4.1 Example of Precoding for Duobinary PR.

$\{b_i\}$:		1	0	0	1	1	1	0	0	1	0
$\{b'_i\}$:	0	1	1	1	0	1	0	0	0	1	1
$\{A_i\}$:	−1	1	1	1	−1	1	−1	−1	−1	1	1
$\{y_i\}$:		0	2	2	0	0	0	−2	−2	0	2

where $\oplus$ denotes modulo 2 addition (exclusive OR). The sequence $\{b'_k\}$ is mapped to the sequence of binary transmitted signal levels $\{A_k\}$ according to

$$A_k = 2b'_k - 1 \tag{4.34}$$

That is, $b'_k = 0$ ($b'_k = 1$) is mapped to the transmitted level $A_k = -1$ ($A_k = 1$). In the absence of noise, the received symbol is then

$$y_k = A_k + A_{k-1} = 2\left(b'_k + b'_{k-1} - 1\right) \tag{4.35}$$

and combining Eqs. (4.33) and (4.35) gives

$$b_k = \left(\frac{1}{2}y_k + 1\right) \bmod 2 \tag{4.36}$$

That is, if $y_k = \pm 2$, then $b_k = 0$, and if $y_k = 0$, then $b_k = 1$. Precoding, therefore, enables the detector to make *symbol-by-symbol* decisions that do not depend on previous decisions. Table 4.1 shows a sequence of transmitted bits $\{b_i\}$, precoded bits $\{b'_i\}$, transmitted signal levels $\{A_i\}$, and received samples $\{y_i\}$.

The preceding precoding technique can be extended to multilevel PAM and to other PR channels. Suppose that the PR is specified by

$$H_{\text{eq}}(D) = \sum_{k=0}^{K} h_k D^k$$

where the coefficients are integers and that the source symbols $\{b_k\}$ are selected from the set $\{0, 1, \ldots, M-1\}$. These symbols are transformed to the sequence $\{b'_k\}$ via the precoding operation

$$b'_k = \left(b_k - \sum_{i=1}^{K} h_i b'_{k-i}\right) \bmod M \tag{4.37}$$

Because of the modulo operation, each symbol b'_k is also in the set $\{0, 1, \ldots, M-1\}$. The kth transmitted signal level is given by

$$A_k = 2b'_k - (M-1) \tag{4.38}$$

so that the set of transmitted levels is $\{-(M-1), \ldots, (M-1)\}$ (i.e., a shifted version of the set of values assumed by b_k). In the absence of noise the received sample is

$$y_k = \sum_{i=0}^{K} h_i A_{k-i} \tag{4.39}$$

and it can be shown that the kth source symbol is given by

$$b_k = \frac{1}{2}\left(y_k + (M-1) \cdot H_{\text{eq}}(1)\right) \bmod M \tag{4.40}$$

Precoding the symbols $\{b_k\}$ in this manner, therefore, enables symbol-by-symbol decisions at the receiver. In the presence of noise, more sophisticated detection schemes (e.g., maximum likelihood) can be used with PR signalling to obtain improvements in performance.

Ford Motor Co.
Exhibit 1030
Page 159

4.6 Additional Considerations

In many applications, bandwidth and intersymbol interference are not the only important considerations for selecting baseband pulses. Here we give a brief discussion of additional practical constraints that may influence this selection.

4.6.1 Average Transmitted Power and Spectral Constraints

The constraint on average transmitted power varies according to the application. For example, low-average power is highly desirable for mobile wireless applications that use battery-powered transmitters. In many applications (e.g., digital subscriber loops, as well as digital radio), constraints are imposed to limit the amount of interference, or crosstalk, radiated into neighboring receivers and communications systems. Because this type of interference is frequency dependent, the constraint may take the form of a spectral mask that specifies the maximum allowable transmitted power as a function of frequency. For example, crosstalk in wireline channels is generally caused by capacitive coupling and increases as a function of frequency. Consequently, to reduce the amount of crosstalk generated at a particular transmitter, the pulse shaping filter generally attenuates high frequencies more than low frequencies.

In radio applications where signals are assigned different frequency bands, constraints on the transmitted spectrum are imposed to limit *adjacent-channel interference*. This interference is generated by transmitters assigned to adjacent frequency bands. Therefore, a constraint is needed to limit the amount of *out-of-band power* generated by each transmitter, in addition to an overall average power constraint. To meet this constraint, the transmitter filter in Fig. 4.3 must have a sufficiently steep rolloff at the edges of the assigned frequency band. (Conversely, if the transmitted signals are time multiplexed, then the duration of the system impulse response must be contained within the assigned time slot.)

4.6.2 Peak-to-Average Power

In addition to a constraint on average transmitted power, a *peak-power* constraint is often imposed as well. This constraint is important in practice for the following reasons:

1. The dynamic range of the transmitter is limited. In particular, saturation of the output amplifier will "clip" the transmitted waveform.
2. Rapid fades can severely distort signals with high peak-to-average power.
3. The transmitted signal may be subjected to nonlinearities. Saturation of the output amplifier is one example. Another example that pertains to wireline applications is the companding process in the voice telephone network [5]. Namely, the compander used to reduce quantization noise for pulse-code modulated voice signals introduces amplitude-dependent distortion in data signals.

The preceding impairments or constraints indicate that the transmitted waveform should have a low **peak-to-average power ratio (PAR)**. For a transmitted waveform $x(t)$, the PAR is defined as

$$\text{PAR} = \frac{\max |x(t)|^2}{E\{|x(t)|^2\}}$$

where $E(\cdot)$ denotes expectation. Using binary signalling with rectangular pulse shapes minimizes the PAR. However, this compromises bandwidth efficiency. In applications where PAR

Ford Motor Co.
Exhibit 1030
Page 160

should be low, binary signalling with rounded pulses are often used. Operating RF power amplifiers with power back-off can also reduce PAR, but leads to inefficient amplification.

For an **orthogonal frequency division multiplexing (OFDM)** system, it is well known that the transmitted signal can exhibit a very high PAR compared to an equivalent single-carrier system. Hence more sophisticated approaches to PAR reduction are required for OFDM. Some proposed approaches are described in [8] and references therein. These include altering the set of transmitted symbols and setting aside certain OFDM tones specifically to minimize PAR.

4.6.3 Channel and Receiver Characteristics

The type of channel impairments encountered and the type of detection scheme used at the receiver can also influence the choice of a transmitted pulse shape. For example, a constant amplitude pulse is appropriate for a fast fading environment with noncoherent detection. The ability to track channel characteristics, such as phase, may allow more bandwidth efficient pulse shapes in addition to multilevel signalling.

High-speed data communications over time-varying channels requires that the transmitter and/or receiver adapt to the changing channel characteristics. Adapting the transmitter to compensate for a time-varying channel requires a feedback channel through which the receiver can notify the transmitter of changes in channel characteristics. Because of this extra complication, adapting the receiver is often preferred to adapting the transmitter pulse shape. However, the following examples are notable exceptions.

1. The current IS-95 air interface for direct-sequence code-division multiple access adapts the transmitter power to control the amount of interference generated and to compensate for channel fades. This can be viewed as a simple form of adaptive transmitter pulse shaping in which a single parameter associated with the pulse shape is varied.
2. Multitone modulation divides the channel bandwidth into small subbands, and the transmitted power and source bits are distributed among these subbands to maximize the information rate. The received signal-to-noise ratio for each subband must be transmitted back to the transmitter to guide the allocation of transmitted bits and power [1].

In addition to multitone modulation, *adaptive precoding* (also known as Tomlinson–Harashima precoding [4, 11]) is another way in which the transmitter can adapt to the channel frequency response. Adaptive precoding is an extension of the technique described earlier for partial-response channels. Namely, the equivalent discrete-time channel impulse response is measured at the receiver and sent back to the transmitter, where it is used in a precoder. The precoder compensates for the intersymbol interference introduced by the channel, allowing the receiver to detect the data by a simple threshhold operation. Both multitone modulation and precoding have been used with wireline channels (voiceband modems and digital subscriber loops).

4.6.4 Complexity

Generation of a bandwidth-efficient signal requires a filter with a sharp cutoff. In addition, bandwidth-efficient pulse shapes can complicate other system functions, such as timing and carrier recovery. If sufficient bandwidth is available, the cost can be reduced by using a rectangular pulse shape with a simple detection strategy (low-pass filter and threshold).

Ford Motor Co.
Exhibit 1030
Page 161

4.6.5 Tolerance to Interference

Interference is one of the primary channel impairments associated with digital radio. In addition to adjacent-channel interference described earlier, *cochannel interference* may be generated by other transmitters assigned to the same frequency band as the desired signal. Cochannel interference can be controlled through frequency (and perhaps time slot) assignments and by pulse shaping. For example, assuming fixed average power, increasing the bandwidth occupied by the signal lowers the power spectral density and decreases the amount of interference into a narrowband system that occupies part of the available bandwidth. Sufficient bandwidth spreading, therefore, enables wideband signals to be overlaid on top of narrowband signals without disrupting either service.

4.6.6 Probability of Intercept and Detection

The broadcast nature of wireless channels generally makes eavesdropping easier than for wired channels. A requirement for most commercial, as well as military applications, is to guarantee the privacy of user conversations (low probability of intercept). An additional requirement, in some applications, is that determining whether or not communications is taking place must be difficult (low probability of detection). Spread-spectrum waveforms are attractive in these applications since spreading the pulse energy over a wide frequency band decreases the power spectral density and, hence, makes the signal less visible. Power-efficient modulation combined with coding enables a further reduction in transmitted power for a target error rate.

4.7 Examples

We conclude this chapter with a brief description of baseband pulse shapes used in existing and emerging standards for digital mobile cellular and Personal Communications Services (PCS).

4.7.1 Global System for Mobile Communications (GSM)

The European GSM standard for digital mobile cellular communications operates in the 900-MHz frequency band, and is based on time-division multiple access (TDMA) [9]. The U.S. version operates at 1900 MHz, and is called PCS-1900. A special variant of binary FSK is used called *Gaussian minimum-shift keying (GMSK)*. The GMSK modulator is illustrated in Fig. 4.13. The input to the modulator is a binary PAM signal $s(t)$, given by Eq. (4.3), where the pulse $p(t)$ is a Gaussian function and $|s(t)| < 1$. This waveform frequency modulates the carrier f_c, so that the (passband) transmitted signal is

$$w(t) = K\cos\left[2\pi f_c t + 2\pi f_d \int_{-\infty}^{t} s(\tau)\,\mathrm{d}\tau\right]$$

The maximum frequency deviation from the carrier is $f_d = 1/(2T)$, which characterizes minimum-shift keying. This technique can be used with a noncoherent receiver that is easy to implement. Because the transmitted signal has a constant envelope, the data can be reliably detected in the presence of rapid fades that are characteristic of mobile radio channels.

Ford Motor Co.
Exhibit 1030
Page 162

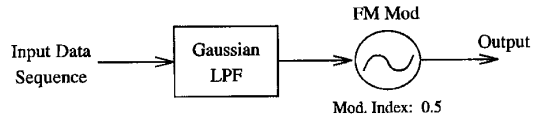


FIGURE 4.13 Generation of GMSK signal; LPF is low-pass filter.

4.7.2 U.S. Digital Cellular (IS-136)

The IS-136 air interface (formerly IS-54) operates in the 800 MHz band and is based on TDMA [3]. There is also a 1900 MHz version of IS-136. The baseband signal is given by Eq. (4.3) where the symbols are complex-valued, corresponding to quadrature phase modulation. The pulse has a square-root raised cosine spectrum with 35% excess bandwidth.

4.7.3 Interim Standard-95

The IS-95 air interface for digital mobile cellular uses spread-spectrum signalling (CDMA) in the 800-MHz band [10]. There is also a 1900 MHz version of IS-95. The baseband transmitted pulse shapes are analogous to those shown in Fig. 4.2, where the number of square pulses (chips) per bit is 128. To improve spectral efficiency the (wideband) transmitted signal is filtered by an approximation to an ideal low-pass response with a small amount of excess bandwidth. This shapes the chips so that they resemble minimum bandwidth pulses.

4.7.4 Personal Access Communications System (PACS)

Both PACS and the Japanese personal handy phone (PHP) system are TDMA systems which have been proposed for personal communications systems (PCS), and operate near 2 GHz [2]. The baseband signal is given by Eq. (4.3) with four complex symbols representing four-phase quadrature modulation. The baseband pulse has a square-root raised cosine spectrum with 50% excess bandwidth.

Defining Terms

Baseband signal: A signal with frequency content centered around DC.

Equivalent discrete-time transfer function: A discrete-time transfer function (z transform) that relates the transmitted amplitudes to received samples in the absence of noise.

Excess bandwidth: That part of the baseband transmitted spectrum which is not contained within the Nyquist band.

Eye diagram: Superposition of segments of a received PAM signal that indicates the amount of intersymbol interference present.

Frequency-shift keying: A digital modulation technique in which the transmitted pulse is sinusoidal, where the frequency is determined by the source bits.

Intersymbol interference: The additive contribution (interference) to a received sample from transmitted symbols other than the symbol to be detected.

Matched filter: The receiver filter with impulse response equal to the time-reversed, complex conjugate impulse response of the combined transmitter filter-channel impulse response.

Ford Motor Co.
Exhibit 1030
Page 163

Nyquist band: The narrowest frequency band that can support a PAM signal without intersymbol interference (the interval $[-1/(2T), 1/(2T)]$ where $1/T$ is the symbol rate).

Nyquist criterion: A condition on the overall frequency response of a PAM system that ensures the absence of intersymbol interference.

Orthogonal frequency division multiplexing (OFDM): Modulation technique in which the transmitted signal is the sum of low-bit-rate narrowband digital signals modulated on orthogonal carriers.

Partial-response signalling: A signalling technique in which a controlled amount of intersymbol interference is introduced at the transmitter in order to shape the transmitted spectrum.

Precoding: A transformation of source symbols at the transmitter that compensates for intersymbol interference introduced by the channel.

Pulse amplitude modulation (PAM): A digital modulation technique in which the source bits are mapped to a sequence of amplitudes that modulate a transmitted pulse.

Raised cosine pulse: A pulse shape with Fourier transform that decays to zero according to a raised cosine; see Eq. (4.18). The amount of excess bandwidth is conveniently determined by a single parameter (α).

Spread spectrum: A signalling technique in which the pulse bandwidth is many times wider than the Nyquist bandwidth.

Zero-forcing criterion: A design constraint which specifies that intersymbol interference be eliminated.

References

[1] Bingham, J.A.C., Multicarrier modulation for data transmission: an idea whose time has come. *IEEE Commun. Mag.*, 28(May), 5–14, 1990.

[2] Cox, D.C., Wireless personal communications: what is it? *IEEE Personal Comm.*, 2(2), 20–35, 1995.

[3] Electronic Industries Association/Telecommunications Industry Association. Recommended minimum performance standards for 800 MHz dual-mode mobile stations. Incorp. EIA/TIA 19B, EIA/TIA Project No. 2216, Mar.,1991

[4] Harashima, H. and Miyakawa, H., Matched-transmission technique for channels with intersymbol interference. *IEEE Trans. on Commun.*, COM-20(Aug.), 774–780, 1972.

[5] Kalet, I. and Saltzberg, B.R., QAM transmission through a companding channel—signal constellations and detection. *IEEE Trans. on Comm.*, 42(2–4), 417–429, 1994.

[6] Kretzmer, E.R., Generalization of a technique for binary data communication. *IEEE Trans. Comm. Tech.*, COM-14 (Feb.), 67, 68, 1966.

[7] Lender, A., The duobinary technique for high-speed data Transmission. *AIEE Trans. on Comm. Electronics*, 82 (March), 214–218, 1963.

[8] Muller, S.H. and Huber, J.B., A comparison of peak power reduction schemes for OFDM. *Proc. GLOBECOM '97*, (Mon.), 1–5, 1997.

[9] Rahnema, M., Overview of the GSM system and protocol architecture. *IEEE Commun. Mag.*, (April), 92–100, 1993.

[10] Telecommunication Industry Association. Mobile station-base station compatibility standard for dual-mode wideband spread spectrum cellular system. TIA/EIA/IS-95-A. May, 1995.

Ford Motor Co.
Exhibit 1030
Page 164

[11] Tomlinson, M., New automatic equalizer employing modulo arithmetic. *Electron. Lett.*, 7 (March), 138, 139, 1971.

Further Information

Baseband signalling and pulse shaping is fundamental to the design of any digital communications system and is, therefore, covered in numerous texts on digital communications. For more advanced treatments see E.A. Lee and D.G. Messerschmitt, *Digital Communication*, Kluwer 1994, and J.G. Proakis, *Digital Communications*, McGraw-Hill 1995.

Ford Motor Co.
Exhibit 1030
Page 165

5

Channel Equalization

5.1 Characterization of Channel Distortion 5-1
5.2 Characterization of Intersymbol Interference 5-2
5.3 Linear Equalizers 5-9
Adaptive Linear Equalizers
5.4 Decision-Feedback Equalizer 5-18
5.5 Maximum-Likelihood Sequence Detection 5-21
5.6 Conclusions 5-22
Defining Terms 5-24
References 5-24
Further Information 5-24

John G. Proakis
Northeastern University

5.1 Characterization of Channel Distortion

Many communication channels, including telephone channels, and some radio channels, may be generally characterized as band-limited linear filters. Consequently, such channels are described by their frequency response $C(f)$, which may be expressed as

$$C(f) = A(f)e^{j\theta(f)} \tag{5.1}$$

where $A(f)$ is called the *amplitude response* and $\theta(f)$ is called the *phase response.* Another characteristic that is sometimes used in place of the phase response is the *envelope delay* or *group delay,* which is defined as

$$\tau(f) = -\frac{1}{2\pi}\frac{\mathrm{d}\theta(f)}{\mathrm{d}f} \tag{5.2}$$

A channel is said to be nondistorting or ideal if, within the bandwidth W occupied by the transmitted signal, $A(f) = \text{const}$ and $\theta(f)$ is a linear function of frequency [or the envelope delay $\tau(f) = \text{const}$]. On the other hand, if $A(f)$ and $\tau(f)$ are not constant within the bandwidth occupied by the transmitted signal, the channel distorts the signal. If $A(f)$ is not constant, the distortion is called *amplitude distortion* and if $\tau(f)$ is not constant, the distortion on the transmitted signal is called *delay distortion.*

As a result of the amplitude and delay distortion caused by the nonideal channel frequency response characteristic $C(f)$, a succession of pulses transmitted through the channel at rates comparable to the bandwidth W are smeared to the point that they are no longer distinguishable as well-defined pulses at the receiving terminal. Instead, they overlap and, thus, we have **intersymbol interference** (**ISI**). As an example of the effect of delay distortion on a transmitted pulse, Fig. 5.1(a) illustrates a band-limited pulse having zeros

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 166

periodically spaced in time at points labeled $\pm T$, $\pm 2T$, etc. If information is conveyed by the pulse amplitude, as in pulse amplitude modulation (PAM), for example, then one can transmit a sequence of pulses, each of which has a peak at the periodic zeros of the other pulses. Transmission of the pulse through a channel modeled as having a linear envelope delay characteristic $\tau(f)$ [quadratic phase $\theta(f)$], however, results in the received pulse shown in Fig. 5.1(b) having zero crossings that are no longer periodically spaced. Consequently a sequence of successive pulses would be smeared into one another, and the peaks of the pulses would no longer be distinguishable. Thus, the channel delay distortion results in intersymbol interference. As will be discussed in this chapter, it is possible to compensate for the nonideal frequency response characteristic of the channel by use of a filter or equalizer at the demodulator. Figure 5.1(c) illustrates the output of a linear equalizer that compensates for the linear distortion in the channel.

The extent of the intersymbol interference on a telephone channel can be appreciated by observing a frequency response characteristic of the channel. Figure 5.2 illustrates the measured average amplitude and delay as a function of frequency for a medium-range (180–725 mi) telephone channel of the switched telecommunications network as given by Duffy and Tratcher, 1971. We observe that the usable band of the channel extends from about 300 Hz to about 3000 Hz. The corresponding impulse response of the average channel is shown in Fig. 5.3. Its duration is about 10 ms. In comparison, the transmitted symbol rates on such a channel may be of the order of 2500 pulses or symbols per second. Hence, intersymbol interference might extend over 20–30 symbols.

Besides telephone channels, there are other physical channels that exhibit some form of time dispersion and, thus, introduce intersymbol interference. Radio channels, such as short-wave ionospheric propagation (HF), tropospheric scatter, and mobile cellular radio are three examples of time-dispersive wireless channels. In these channels, time dispersion and, hence, intersymbol interference is the result of multiple propagation paths with different path delays. The number of paths and the relative time delays among the paths vary with time and, for this reason, these radio channels are usually called time-variant multipath channels. The time-variant multipath conditions give rise to a wide variety of frequency response characteristics. Consequently, the frequency response characterization that is used for telephone channels is inappropriate for time-variant multipath channels. Instead, these radio channels are characterized statistically in terms of the scattering function, which, in brief, is a two-dimensional representation of the average received signal power as a function of relative time delay and Doppler frequency (see Proakis [4]).

For illustrative purposes, a scattering function measured on a medium-range (150 mi) tropospheric scatter channel is shown in Fig. 5.4. The total time duration (multipath spread) of the channel response is approximately 0.7 μs on the average, and the spread between half-power points in Doppler frequency is a little less than 1 Hz on the strongest path and somewhat larger on the other paths. Typically, if one is transmitting at a rate of 10^7 symbols/s over such a channel, the multipath spread of 0.7 μs will result in intersymbol interference that spans about seven symbols.

5.2 Characterization of Intersymbol Interference

In a digital communication system, channel distortion causes intersymbol interference, as illustrated in the preceding section. In this section, we shall present a model that characterizes the ISI. The digital modulation methods to which this treatment applies are PAM, phase-shift keying (PSK) and quadrature amplitude modulation (QAM). The transmitted

Ford Motor Co.
Exhibit 1030
Page 167

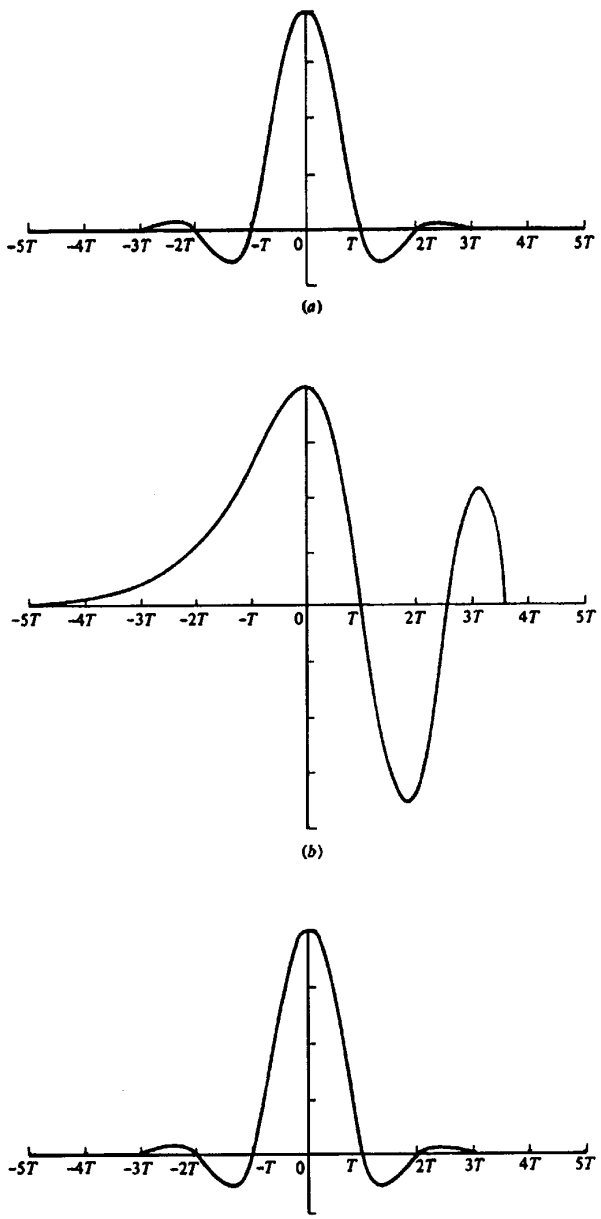


FIGURE 5.1 Effect of channel distortion: (a) channel input, (b) channel output, (c) equalizer output.

Ford Motor Co.
Exhibit 1030
Page 168

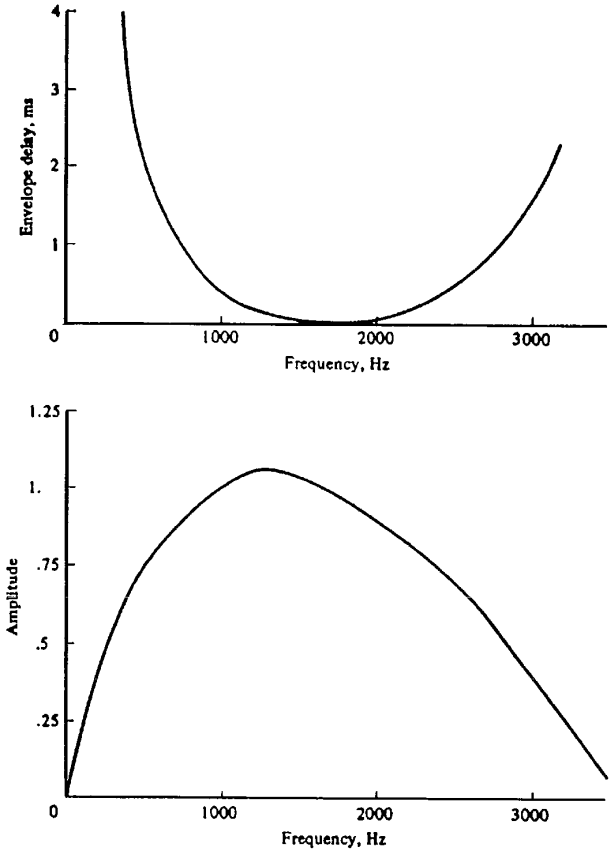


FIGURE 5.2 Average amplitude and delay characteristics of medium-range telephone channel.

signal for these three types of modulation may be expressed as

$$\begin{aligned} s(t) &= v_c(t)\cos 2\pi f_c t - v_s(t)\sin 2\pi f_c t \\ &= \text{Re}\left[v(t)\, e^{j2\pi f_c t}\right] \end{aligned} \tag{5.3}$$

where $v(t) = v_c(t) + jv_s(t)$ is called the *equivalent low-pass signal*, f_c is the carrier frequency, and Re[ ] denotes the real part of the quantity in brackets.

In general, the equivalent low-pass signal is expressed as

$$v(t) = \sum_{n=0}^{\infty} I_n\, g_T(t - nT) \tag{5.4}$$

where $g_T(t)$ is the basic pulse shape that is selected to control the spectral characteristics of the transmitted signal, $\{I_n\}$ the sequence of transmitted information symbols selected from a signal constellation consisting of M points, and T the signal interval ($1/T$ is the symbol rate). For PAM, PSK, and QAM, the values of I_n are points from M-ary signal constellations. Figure 5.5 illustrates the signal constellations for the case of $M = 8$ signal points. Note that for PAM, the signal constellation is one dimensional. Hence, the equivalent low-pass signal $v(t)$ is real valued, i.e., $v_s(t) = 0$ and $v_c(t) = v(t)$. For M-ary ($M > 2$)

Ford Motor Co.
Exhibit 1030
Page 169

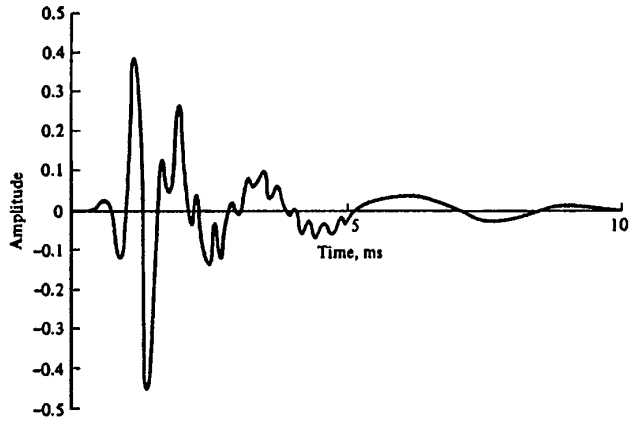


FIGURE 5.3 Impulse response of average channel with amplitude and delay shown in Fig.5.2.

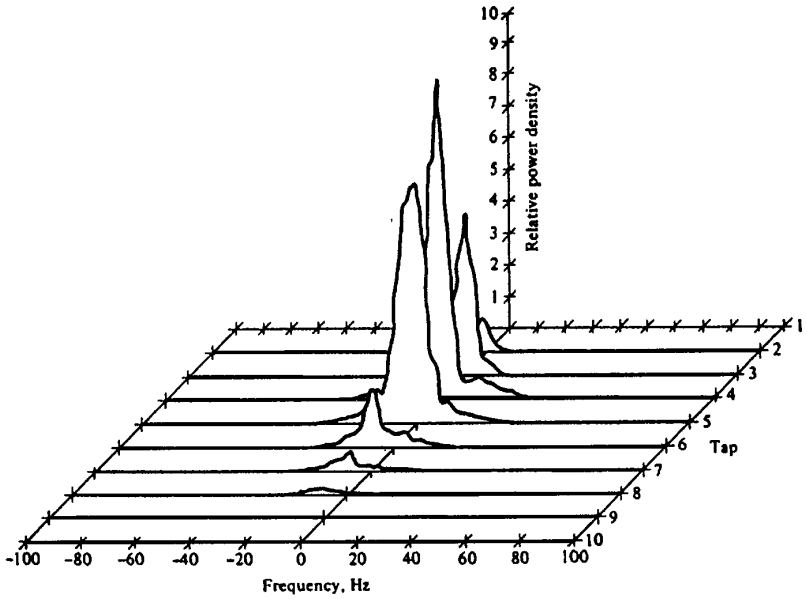


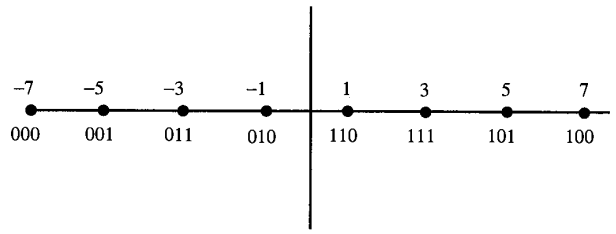
FIGURE 5.4 Scattering function of a medium-range tropospheric scatter channel.

PSK and QAM, the signal constellations are two dimensional and, hence, $v(t)$ is complex valued.

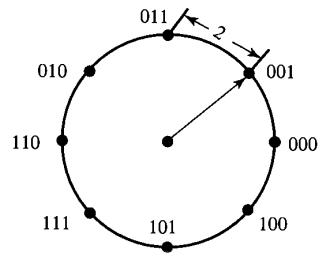
The signal $s(t)$ is transmitted over a bandpass channel that may be characterized by an equivalent low-pass frequency response $C(f)$. Consequently, the equivalent low-pass received signal can be represented as

$$r(t) = \sum_{n=0}^{\infty} I_n \, h(t - nT) + w(t) \tag{5.5}$$

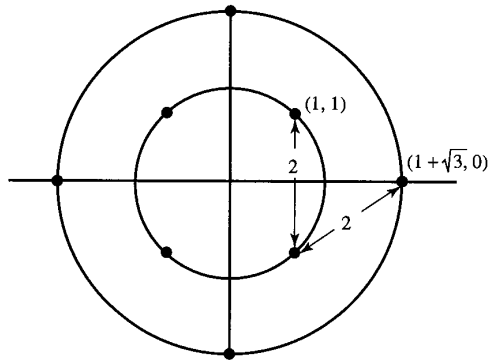
where $h(t) = g_T(t) * c(t)$, and $c(t)$ is the impulse response of the equivalent low-pass channel, the asterisk denotes convolution, and $w(t)$ represents the additive noise in the channel.



(a) PAM



(b) PSK



(c) QAM

FIGURE 5.5 $M = 8$ signal constellations for PAM, PSK, and QAM.

To characterize the ISI, suppose that the received signal is passed through a receiving filter and then sampled at the rate $1/T$ samples/s. In general, the optimum filter at the receiver is matched to the received signal pulse $h(t)$. Hence, the frequency response of this filter is $H^*(f)$. We denote its output as

$$y(t) = \sum_{n=0}^{\infty} I_n x(t - nT) + \nu(t) \tag{5.6}$$

Ford Motor Co.
Exhibit 1030
Page 171

where $x(t)$ is the signal pulse response of the receiving filter, i.e., $X(f) = H(f)H^*(f) = |H(f)|^2$, and $\nu(t)$ is the response of the receiving filter to the noise $w(t)$. Now, if $y(t)$ is sampled at times $t = kT$, $k = 0, 1, 2, \ldots$, we have

$$\begin{aligned} y(kT) \equiv y_k &= \sum_{n=0}^{\infty} I_n x(kT - nT) + \nu(kT) \\ &= \sum_{n=0}^{\infty} I_n x_{k-n} + \nu_k, \qquad k = 0, 1, \ldots \end{aligned} \tag{5.7}$$

The sample values $\{y_k\}$ can be expressed as

$$y_k = x_0 \left(I_k + \frac{1}{x_0} \sum_{\substack{n=0 \\ n \neq k}}^{\infty} I_n x_{k-n} \right) + \nu_k, \qquad k = 0, 1, \ldots \tag{5.8}$$

The term x_0 is an arbitrary scale factor, which we arbitrarily set equal to unity for convenience. Then

$$y_k = I_k + \sum_{\substack{n=0 \\ n \neq k}}^{\infty} I_n x_{k-n} + \nu_k \tag{5.9}$$

The term I_k represents the desired information symbol at the kth sampling instant, the term

$$\sum_{\substack{n=0 \\ n \neq k}}^{\infty} I_n x_{k-n} \tag{5.10}$$

represents the ISI, and ν_k is the additive noise variable at the kth sampling instant.

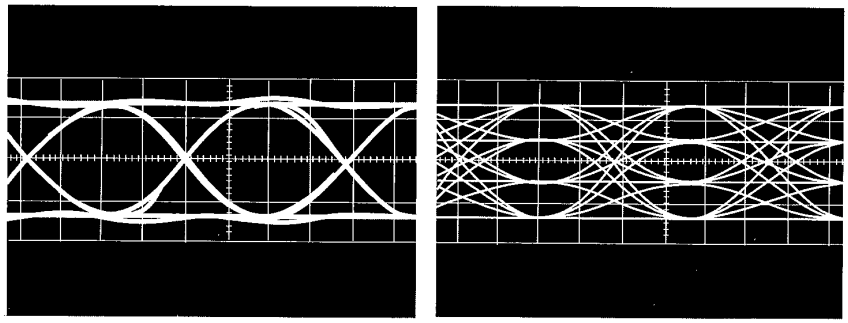
The amount of ISI, and noise in a digital communications system can be viewed on an oscilloscope. For PAM signals, we can display the received signal $y(t)$ on the vertical input with the horizontal sweep rate set at $1/T$. The resulting oscilloscope display is called an *eye pattern* because of its resemblance to the human eye. For example, Fig. 5.6 illustrates the eye patterns for binary and four-level PAM modulation. The effect of ISI is to cause the eye to close, thereby reducing the margin for additive noise to cause errors. Figure 5.7 graphically illustrates the effect of ISI in reducing the opening of a binary eye. Note that intersymbol interference distorts the position of the zero crossings and causes a reduction in the eye opening. Thus, it causes the system to be more sensitive to a synchronization error.

For PSK and QAM it is customary to display the eye pattern as a two-dimensional scatter diagram illustrating the sampled values $\{y_k\}$ that represent the decision variables at the sampling instants. Figure 5.8 illustrates such an eye pattern for an 8-PSK signal. In the absence of intersymbol interference and noise, the superimposed signals at the sampling instants would result in eight distinct points corresponding to the eight transmitted signal phases. Intersymbol interference and noise result in a deviation of the received samples $\{y_k\}$ from the desired 8-PSK signal. The larger the intersymbol interference and noise, the larger the scattering of the received signal samples relative to the transmitted signal points.

In practice, the transmitter and receiver filters are designed for zero ISI at the desired sampling times $t = kT$. Thus, if $G_T(f)$ is the frequency response of the transmitter filter and $G_R(f)$ is the frequency response of the receiver filter, then the product $G_T(f)\,G_R(f)$ is designed to yield zero ISI. For example, the product $G_T(f)\,G_R(f)$ may be selected as

$$G_T(f)G_R(f) = X_{rc}(f) \tag{5.11}$$

Ford Motor Co.
Exhibit 1030
Page 172



BINARY QUATERNARY

FIGURE 5.6 Examples of eye patterns for binary and quaternary amplitude shift keying (or PAM).

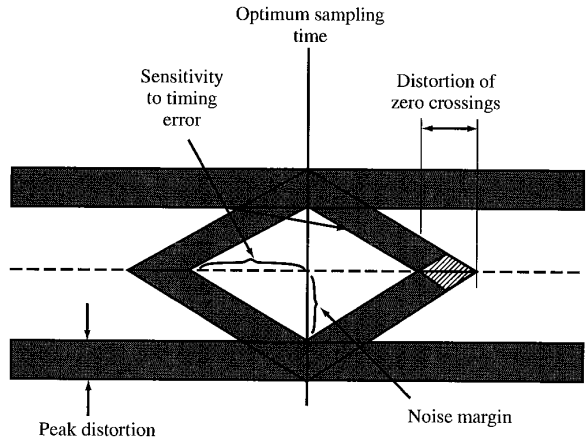


FIGURE 5.7 Effect of intersymbol interference on eye opening.

where $X_{rc}(f)$ is the raised-cosine frequency response characteristic, defined as

$$X_{rc}(f) = \begin{cases} T, & 0 \le |f| \le (1-\alpha)/2T \\ \frac{T}{2}\left[1 + \cos\frac{\pi T}{\alpha}\left(|f| - \frac{1-\alpha}{2T}\right)\right], & \frac{1-\alpha}{2T} \le |f| \le \frac{1+\alpha}{2T} \\ 0, & |f| > \frac{1+\alpha}{2T} \end{cases} \tag{5.12}$$

where α is called the *rolloff* factor, which takes values in the range $0 \le \alpha \le 1$, and $1/T$ is the symbol rate. The frequency response $X_{rc}(f)$ is illustrated in Fig. 5.9(a) for $\alpha = 0, 1/2$, and 1. Note that when $\alpha = 0$, $X_{rc}(f)$ reduces to an ideal brick wall physically nonrealizable frequency response with bandwidth occupancy $1/2T$. The frequency $1/2T$ is called the *Nyquist frequency.* For $\alpha > 0$, the bandwidth occupied by the desired signal $X_{rc}(f)$ beyond the Nyquist frequency $1/2T$ is called the *excess bandwidth,* and is usually expressed as a percentage of the Nyquist frequency. For example, when $\alpha = 1/2$, the excess bandwidth is 50% and when $\alpha = 1$, the excess bandwidth is 100%. The signal pulse $x_{rc}(t)$ having the

Ford Motor Co.
Exhibit 1030
Page 173

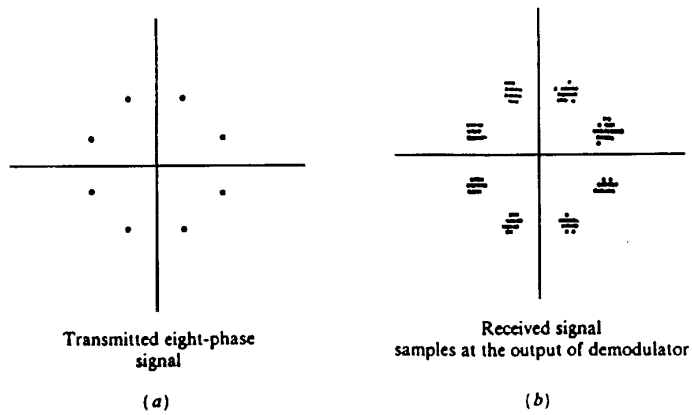


FIGURE 5.8 Two-dimensional digital eye patterns.

raised-cosine spectrum is

$$x_{rc}(t) = \frac{\sin \pi t/T}{\pi t/T} \quad \frac{\cos(\pi \alpha t/T)}{1 - 4\alpha^2 t^2/T^2} \tag{5.13}$$

Figure 5.9(b) illustrates $x_{rc}(t)$ for $\alpha = 0$, $1/2$, and 1. Note that $x_{rc}(t) = 1$ at $t = 0$ and $x_{rc}(t) = 0$ at $t = kT$, $k = \pm 1, \pm 2, \ldots$. Consequently, at the sampling instants $t = kT$, $k \neq 0$, there is no ISI from adjacent symbols when there is no channel distortion. In the presence of channel distortion, however, the ISI given by Eq. (5.10) is no longer zero, and a **channel equalizer** is needed to minimize its effect on system performance.

5.3 Linear Equalizers

The most common type of channel equalizer used in practice to reduce SI is a linear transversal filter with adjustable coefficients $\{c_i\}$, as shown in Fig. 5.10.

On channels whose frequency response characteristics are unknown, but time invariant, we may measure the channel characteristics and adjust the parameters of the equalizer; once adjusted, the parameters remain fixed during the transmission of data. Such equalizers are called **preset equalizers.** On the other hand, **adaptive equalizers** update their parameters on a periodic basis during the transmission of data and, thus, they are capable of tracking a slowly time-varying channel response.

First, let us consider the design characteristics for a linear equalizer from a frequency domain viewpoint. Figure 5.11 shows a block diagram of a system that employs a linear filter as a **channel equalizer.**

The demodulator consists of a receiver filter with frequency response $G_R(f)$ in cascade with a channel equalizing filter that has a frequency response $G_E(f)$. As indicated in the preceding section, the receiver filter response $G_R(f)$ is matched to the transmitter response, i.e., $G_R(f) = G_T^*(f)$, and the product $G_R(f)G_T(f)$ is usually designed so that there is zero ISI at the sampling instants as, for example, when $G_R(t)G_T(f) = X_{rc}(f)$.

For the system shown in Fig. 5.11, in which the channel frequency response is not ideal, the desired condition for zero ISI is

$$G_T(f)C(f)G_R(f)G_E(f) = X_{rc}(f) \tag{5.14}$$

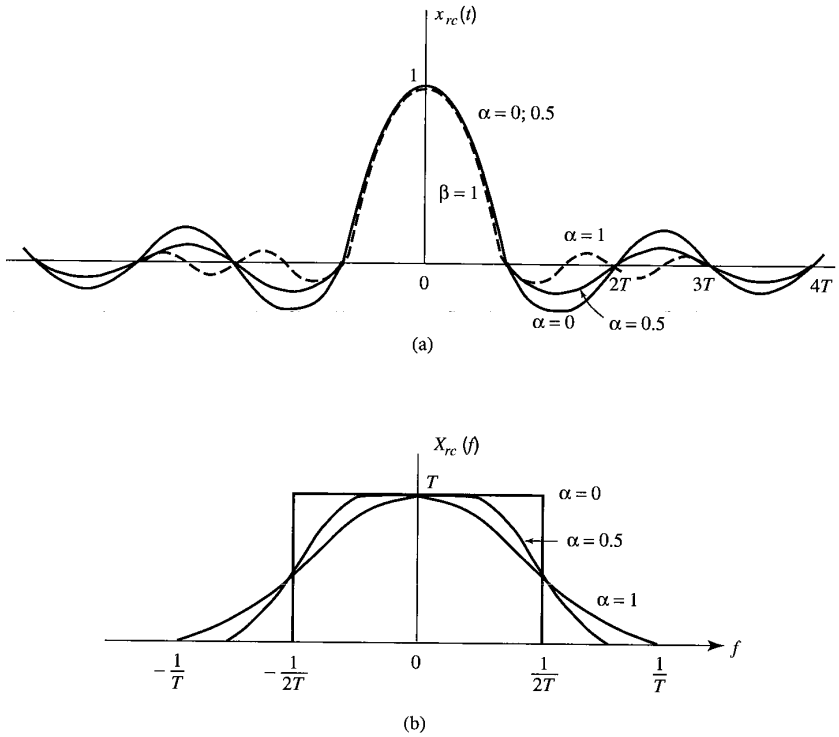


FIGURE 5.9 Pulses having a raised cosine spectrum.

where $X_{rc}(f)$ is the desired raised-cosine spectral characteristic. Since $G_T(f)G_R(f) = X_{rc}(f)$ by design, the frequency response of the equalizer that compensates for the channel distortion is

$$G_E(f) = \frac{1}{C(f)} = \frac{1}{|C(f)|} e^{-j\theta_c(f)} \tag{5.15}$$

Thus, the amplitude response of the equalizer is $|G_E(f)| = 1/|C(f)|$ and its phase response is $\theta_E(f) = -\theta_c(f)$. In this case, the equalizer is said to be the *inverse channel filter* to the channel response.

We note that the inverse channel filter completely eliminates ISI caused by the channel. Since it forces the ISI to be zero at the sampling instants $t = kT$, $k = 0, 1, \ldots$, the equalizer is called a **zero-forcing equalizer.** Hence, the input to the detector is simply

$$z_k = I_k + \eta_k, \qquad k = 0, 1, \ldots \tag{5.16}$$

where η_k represents the additive noise and I_k is the desired symbol.

In practice, the ISI caused by channel distortion is usually limited to a finite number of symbols on either side of the desired symbol. Hence, the number of terms that constitute the ISI in the summation given by Eq. (5.10) is finite. As a consequence, in practice the channel equalizer is implemented as a finite duration impulse response (FIR) filter, or transversal filter, with adjustable tap coefficients $\{c_n\}$, as illustrated in Fig. 5.10. The time delay τ between adjacent taps may be selected as large as T, the symbol interval, in which case the FIR equalizer is called a **symbol-spaced equalizer.** In this case, the input to the

Ford Motor Co.
Exhibit 1030
Page 175

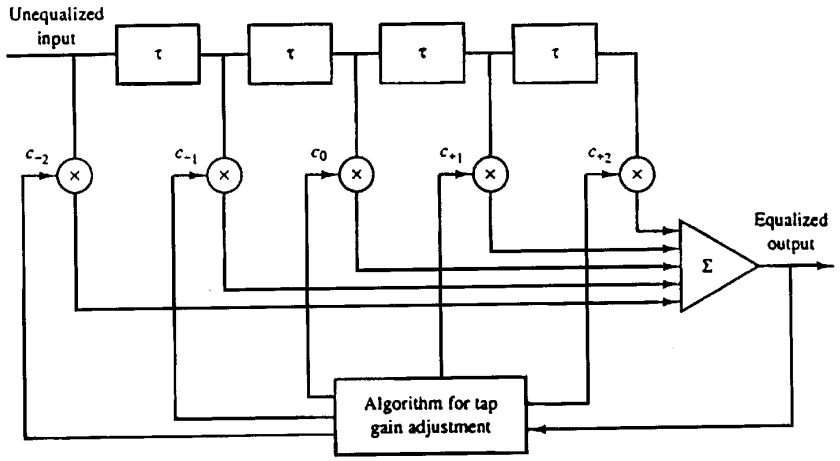


FIGURE 5.10 Linear transversal filter.

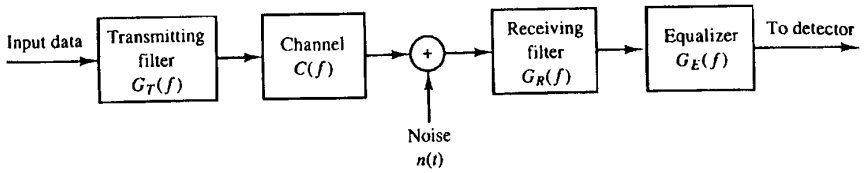


FIGURE 5.11 Block diagram of a system with an equalizer.

equalizer is the sampled sequence given by Eq. (5.7). We note that when the symbol rate $1/T < 2W$, however, frequencies in the received signal above the folding frequency $1/T$ are aliased into frequencies below $1/T$. In this case, the equalizer compensates for the aliased channel-distorted signal.

On the other hand, when the time delay τ between adjacent taps is selected such that $1/\tau \geq 2W > 1/T$, no aliasing occurs and, hence, the inverse channel equalizer compensates for the true channel distortion. Since $\tau < T$, the channel equalizer is said to have *fractionally spaced taps* and it is called a **fractionally spaced equalizer.** In practice, τ is often selected as $\tau = T/2$. Notice that, in this case, the sampling rate at the output of the filter $G_R(f)$ is $2/T$.

The impulse response of the FIR equalizer is

$$g_E(t) = \sum_{n=-N}^{N} c_n \delta(t - n\tau) \tag{5.17}$$

and the corresponding frequency response is

$$G_E(f) = \sum_{n=-N}^{N} c_n e^{-j2\pi f n\tau} \tag{5.18}$$

where $\{c_n\}$ are the $(2N+1)$ equalizer coefficients and N is chosen sufficiently large so that the equalizer spans the length of the ISI, i.e., $2N+1 \geq L$, where L is the number of signal samples spanned by the ISI. Since $X(f) = G_T(f)C(f)G_R(f)$ and $x(t)$ is the signal pulse

corresponding to $X(f)$, then the equalized output signal pulse is

$$q(t) = \sum_{n=-N}^{N} c_n x(t - n\tau) \tag{5.19}$$

The zero-forcing condition can now be applied to the samples of $q(t)$ taken at times $t = mT$. These samples are

$$q(mT) = \sum_{n=-N}^{N} c_n x(mT - n\tau), \qquad m = 0, \pm 1, \ldots, \pm N \tag{5.20}$$

Since there are $2N + 1$ equalizer coefficients, we can control only $2N + 1$ sampled values of $q(t)$. Specifically, we may force the conditions

$$q(mT) = \sum_{n=-N}^{N} c_n x(mT - n\tau) = \begin{cases} 1, & m = 0 \\ 0, & m = \pm 1, \pm 2, \ldots, \pm N \end{cases} \tag{5.21}$$

which may be expressed in matrix form as $\boldsymbol{X}\boldsymbol{c} = \boldsymbol{q}$, where $\boldsymbol{X}$ is a $(2N+1) \times (2N+1)$ matrix with elements $\{x(mT - n\tau)\}$, $\boldsymbol{c}$ is the $(2N + 1)$ coefficient vector and $\boldsymbol{q}$ is the $(2N + 1)$ column vector with one nonzero element. Thus, we obtain a set of $2N + 1$ linear equations for the coefficients of the zero-forcing equalizer.

We should emphasize that the FIR zero-forcing equalizer does not completely eliminate ISI because it has a finite length. As N is increased, however, the residual ISI can be reduced, and in the limit as $N \to \infty$, the ISI is completely eliminated.

EXAMPLE 5.1:

Consider a channel distorted pulse $x(t)$, at the input to the equalizer, given by the expression

$$x(t) = \frac{1}{1 + \left(\frac{2t}{T}\right)^2}$$

where $1/T$ is the symbol rate. The pulse is sampled at the rate $2/T$ and equalized by a zero-forcing equalizer. Determine the coefficients of a five-tap zero-forcing equalizer.

Solution 5.1 According to Eq. (5.21), the zero-forcing equalizer must satisfy the equations

$$q(mT) = \sum_{n=-2}^{2} c_n x\left(mT - nT/2\right) = \begin{cases} 1, & m = 0 \\ 0, & m = \pm 1, \pm 2 \end{cases}$$

The matrix $\boldsymbol{X}$ with elements $x(mT - nT/2)$ is given as

$$\boldsymbol{X} = \begin{bmatrix} \frac{1}{5} & \frac{1}{10} & \frac{1}{17} & \frac{1}{26} & \frac{1}{37} \\ 1 & \frac{1}{2} & \frac{1}{5} & \frac{1}{10} & \frac{1}{17} \\ \frac{1}{5} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{5} \\ \frac{1}{17} & \frac{1}{10} & \frac{1}{5} & \frac{1}{2} & 1 \\ \frac{1}{37} & \frac{1}{26} & \frac{1}{17} & \frac{1}{10} & \frac{1}{5} \end{bmatrix} \tag{5.22}$$

Ford Motor Co.
Exhibit 1030
Page 177

The coefficient vector $\boldsymbol{c}$ and the vector $\boldsymbol{q}$ are given as

$$\boldsymbol{c} = \begin{bmatrix} c_{-2} \\ c_{-1} \\ c_0 \\ c_1 \\ c_2 \end{bmatrix} \qquad \boldsymbol{q} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{5.23}$$

Then, the linear equations $\boldsymbol{X}\boldsymbol{c} = \boldsymbol{q}$ can be solved by inverting the matrix $\boldsymbol{X}$. Thus, we obtain

$$\boldsymbol{c}_{\text{opt}} = \boldsymbol{X}^{-1}\boldsymbol{q} = \begin{bmatrix} -2.2 \\ 4.9 \\ -3 \\ 4.9 \\ -2.2 \end{bmatrix} \tag{5.24}$$

One drawback to the zero-forcing equalizer is that it ignores the presence of additive noise. As a consequence, its use may result in significant noise enhancement. This is easily seen by noting that in a frequency range where $C(f)$ is small, the channel equalizer $G_E(f) = 1/C(f)$ compensates by placing a large gain in that frequency range. Consequently, the noise in that frequency range is greatly enhanced. An alternative is to relax the zero ISI condition and select the channel equalizer characteristic such that the combined power in the residual ISI and the additive noise at the output of the equalizer is minimized. A channel equalizer that is optimized based on the minimum mean square error (MMSE) criterion accomplishes the desired goal.

To elaborate, let us consider the noise corrupted output of the FIR equalizer, which is

$$z(t) = \sum_{n=-N}^{N} c_n y(t - n\tau) \tag{5.25}$$

where $y(t)$ is the input to the equalizer, given by Eq. (5.6). The equalizer output is sampled at times $t = mT$. Thus, we obtain

$$z(mT) = \sum_{n=-N}^{N} c_n y(mT - n\tau) \tag{5.26}$$

The desired response at the output of the equalizer at $t = mT$ is the transmitted symbol I_m. The error is defined as the difference between I_m and $z(mT)$. Then, the mean square error (*MSE*) between the actual output sample $z(mT)$ and the desired values I_m is

$$\begin{aligned} MSE &= E\left|z(mT) - I_m\right|^2 \\ &= E\left[\left|\sum_{n=-N}^{N} c_n y(mT - n\tau) - I_m\right|^2\right] \\ &= \sum_{n=-N}^{N}\sum_{k=-N}^{N} c_n c_k R_Y(n-k) \\ &\quad - 2\sum_{k=-N}^{N} c_k R_{IY}(k) + E\left(|I_m|^2\right) \end{aligned} \tag{5.27}$$

where the correlations are defined as

$$\begin{aligned} R_Y(n-k) &= E\left[y^*(mT-n\tau)y(mT-k\tau)\right] \\ R_{IY}(k) &= E\left[y(mT-k\tau)I_m^*\right] \end{aligned} \tag{5.28}$$

and the expectation is taken with respect to the random information sequence $\{I_m\}$ and the additive noise.

The minimum *MSE* solution is obtained by differentiating Eq. (5.27) with respect to the equalizer coefficients $\{c_n\}$. Thus, we obtain the necessary conditions for the minimum *MSE* as

$$\sum_{n=-N}^{N} c_n R_Y(n-k) = R_{IY}(k), \qquad k = 0, \pm 1, 2, \ldots, \pm N \tag{5.29}$$

These are the $(2N+1)$ linear equations for the equalizer coefficients. In contrast to the zero-forcing solution already described, these equations depend on the statistical properties (the autocorrelation) of the noise as well as the ISI through the autocorrelation $R_Y(n)$.

In practice, the autocorrelation matrix $R_Y(n)$ and the crosscorrelation vector $R_{IY}(n)$ are unknown a priori. These correlation sequences can be estimated, however, by transmitting a test signal over the channel and using the time-average estimates

$$\begin{aligned} \hat{R}_Y(n) &= \frac{1}{K}\sum_{k=1}^{K} y^*(kT-n\tau)y(kT) \\ \hat{R}_{IY}(n) &= \frac{1}{K}\sum_{k=1}^{K} y(kT-n\tau)I_k^* \end{aligned} \tag{5.30}$$

in place of the ensemble averages to solve for the equalizer coefficients given by Eq. (5.29).

5.3.1 Adaptive Linear Equalizers

We have shown that the tap coefficients of a linear equalizer can be determined by solving a set of linear equations. In the zero-forcing optimization criterion, the linear equations are given by Eq. (5.21). On the other hand, if the optimization criterion is based on minimizing the *MSE,* the optimum equalizer coefficients are determined by solving the set of linear equations given by Eq. (5.29).

In both cases, we may express the set of linear equations in the general matrix form

$$\boldsymbol{B}\boldsymbol{c} = \boldsymbol{d} \tag{5.31}$$

where $\boldsymbol{B}$ is a $(2N+1)\times(2N+1)$ matrix, $\boldsymbol{c}$ is a column vector representing the $2N+1$ equalizer coefficients, and $\boldsymbol{d}$ a $(2N+1)$-dimensional column vector. The solution of Eq. (5.31) yields

$$\boldsymbol{c}_{\text{opt}} = \boldsymbol{B}^{-1}\boldsymbol{d} \tag{5.32}$$

In practical implementations of equalizers, the solution of Eq. (5.31) for the optimum coefficient vector is usually obtained by an iterative procedure that avoids the explicit computation of the inverse of the matrix $\boldsymbol{B}$. The simplest iterative procedure is the method of steepest descent, in which one begins by choosing arbitrarily the coefficient vector $\boldsymbol{c}$, say $\boldsymbol{c}_0$. This initial choice of coefficients corresponds to a point on the criterion function that is being optimized. For example, in the case of the *MSE* criterion, the initial guess $\boldsymbol{c}_0$

Ford Motor Co.
Exhibit 1030
Page 179

corresponds to a point on the quadratic *MSE* surface in the $(2N+1)$-dimensional space of coefficients. The gradient vector, defined as $\boldsymbol{g}_0$, which is the derivative of the *MSE* with respect to the $2N+1$ filter coefficients, is then computed at this point on the criterion surface, and each tap coefficient is changed in the direction opposite to its corresponding gradient component. The change in the jth tap coefficient is proportional to the size of the jth gradient component.

For example, the gradient vector denoted as $\boldsymbol{g}_k$, for the MSE criterion, found by taking the derivatives of the MSE with respect to each of the $2N+1$ coefficients, is

$$\boldsymbol{g}_k = \boldsymbol{B}\boldsymbol{c}_k - \boldsymbol{d}, \qquad k = 0, 1, 2, \ldots \tag{5.33}$$

Then the coefficient vector $\boldsymbol{c}_k$ is updated according to the relation

$$\boldsymbol{c}_{k+1} = \boldsymbol{c}_k - \Delta \boldsymbol{g}_k \tag{5.34}$$

where Δ is the *step-size parameter* for the iterative procedure. To ensure convergence of the iterative procedure, Δ is chosen to be a small positive number. In such a case, the gradient vector $\boldsymbol{g}_k$ converges toward zero, i.e., $\boldsymbol{g}_k \to \boldsymbol{0}$ as $k \to \infty$, and the coefficient vector $\boldsymbol{c}_k \to \boldsymbol{c}_{\text{opt}}$ as illustrated in Fig. 5.12 based on two-dimensional optimization. In general, convergence of the equalizer tap coefficients to $\boldsymbol{c}_{\text{opt}}$ cannot be attained in a finite number of iterations with the steepest-descent method. The optimum solution $\boldsymbol{c}_{\text{opt}}$, however, can be approached as closely as desired in a few hundred iterations. In digital communication systems that employ channel equalizers, each iteration corresponds to a time interval for sending one symbol and, hence, a few hundred iterations to achieve convergence to $\boldsymbol{c}_{\text{opt}}$ corresponds to a fraction of a second.

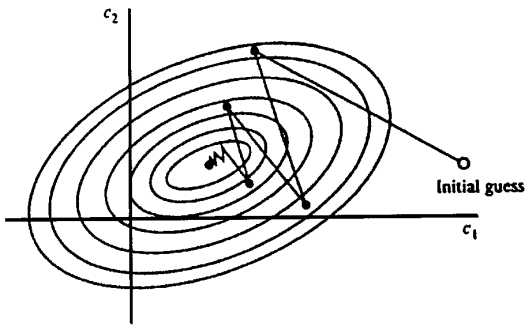


FIGURE 5.12 Examples of convergence characteristics of a gradient algorithm.

Adaptive channel equalization is required for channels whose characteristics change with time. In such a case, the ISI varies with time. The channel equalizer must track such time variations in the channel response and adapt its coefficients to reduce the ISI. In the context of the preceding discussion, the optimum coefficient vector $\boldsymbol{c}_{\text{opt}}$ varies with time due to time variations in the matrix $\boldsymbol{B}$ and, for the case of the *MSE* criterion, time variations in the vector $\boldsymbol{d}$. Under these conditions, the iterative method described can be modified to use estimates of the gradient components. Thus, the algorithm for adjusting the equalizer tap coefficients may be expressed as

$$\hat{\boldsymbol{c}}_{k+1} = \hat{\boldsymbol{c}}_k - \Delta \hat{\boldsymbol{g}}_k \tag{5.35}$$

where $\hat{\boldsymbol{g}}_k$ denotes an estimate of the gradient vector $\boldsymbol{g}_k$ and $\hat{\boldsymbol{c}}_k$ denotes the estimate of the tap coefficient vector.

Ford Motor Co.
Exhibit 1030
Page 180

In the case of the *MSE* criterion, the gradient vector $\boldsymbol{g}_k$ given by Eq. (5.33) may also be expressed as

$$\boldsymbol{g}_k = -E\left(e_k \boldsymbol{y}_k^*\right)$$

An estimate $\hat{\boldsymbol{g}}_k$ of the gradient vector at the kth iteration is computed as

$$\hat{\boldsymbol{g}}_k = -e_k \boldsymbol{y}_k^* \tag{5.36}$$

where e_k denotes the difference between the desired output from the equalizer at the kth time instant and the actual output $z(kT)$, and $\boldsymbol{y}_k$ denotes the column vector of $2N+1$ received signal values contained in the equalizer at time instant k. The *error signal* e_k is expressed as

$$e_k = I_k - z_k \tag{5.37}$$

where $z_k = z(kT)$ is the equalizer output given by Eq. (5.26) and I_k is the desired symbol. Hence, by substituting Eq. (5.36) into Eq. (5.35), we obtain the adaptive algorithm for optimizing the tap coefficients (based on the *MSE* criterion) as

$$\hat{\boldsymbol{c}}_{k+1} = \hat{\boldsymbol{c}}_k + \Delta e_k \boldsymbol{y}_k^* \tag{5.38}$$

Since an estimate of the gradient vector is used in Eq. (5.38) the algorithm is called a **stochastic gradient algorithm;** it is also known as the **LMS algorithm.**

A block diagram of an adaptive equalizer that adapts its tap coefficients according to Eq. (5.38) is illustrated in Fig. 5.13. Note that the difference between the desired output I_k and the actual output z_k from the equalizer is used to form the error signal e_k. This error is scaled by the step-size parameter Δ, and the scaled error signal Δe_k multiplies the received signal values $\{y(kT - n\tau)\}$ at the $2N+1$ taps. The products $\Delta e_k y^*(kT - n\tau)$ at the $(2N+1)$ taps are then added to the previous values of the tap coefficients to obtain the updated tap coefficients, according to Eq. (5.38). This computation is repeated as each new symbol is received. Thus, the equalizer coefficients are updated at the symbol rate.

Initially, the adaptive equalizer is trained by the transmission of a known pseudo-random sequence $\{I_m\}$ over the channel. At the demodulator, the equalizer employs the known sequence to adjust its coefficients. Upon initial adjustment, the adaptive equalizer switches from a **training mode** to a **decision-directed mode,** in which case the decisions at the output of the detector are sufficiently reliable so that the error signal is formed by computing the difference between the detector output and the equalizer output, i.e.,

$$e_k = \tilde{I}_k - z_k \tag{5.39}$$

where $\tilde{I}_k$ is the output of the detector. In general, decision errors at the output of the detector occur infrequently and, consequently, such errors have little effect on the performance of the tracking algorithm given by Eq. (5.38).

A rule of thumb for selecting the step-size parameter so as to ensure convergence and good tracking capabilities in slowly varying channels is

$$\Delta = \frac{1}{5(2N+1)P_R} \tag{5.40}$$

where P_R denotes the received signal-plus-noise power, which can be estimated from the received signal (see Proakis [4]).

The convergence characteristic of the stochastic gradient algorithm in Eq. (5.38) is illustrated in Fig. 5.14. These graphs were obtained from a computer simulation of an 11-tap adaptive equalizer operating on a channel with a rather modest amount of ISI. The input

Ford Motor Co.
Exhibit 1030
Page 181

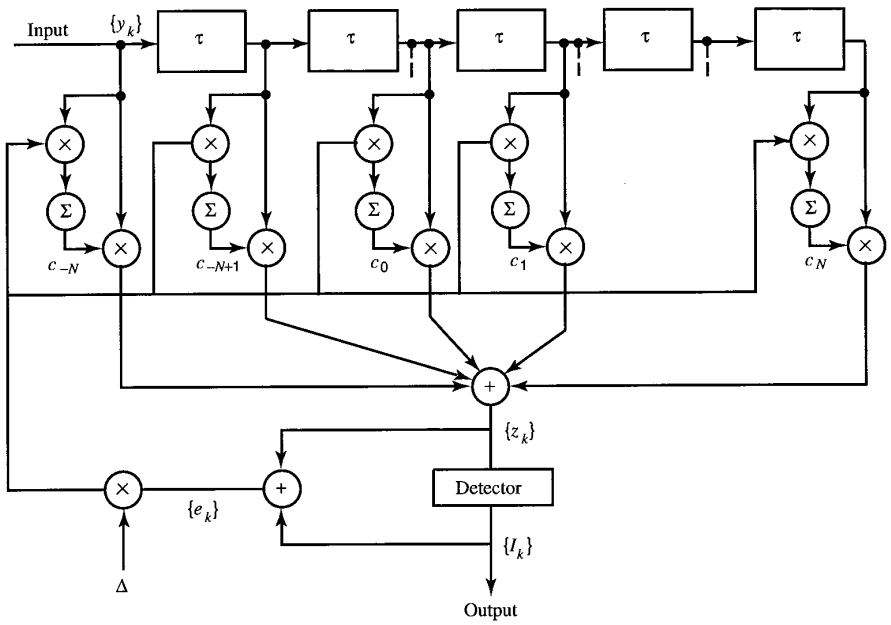


FIGURE 5.13 Linear adaptive equalizer based on the *MSE* criterion.

signal-plus-noise power P_R was normalized to unity. The rule of thumb given in Eq. (5.40) for selecting the step size gives $\Delta = 0.018$. The effect of making Δ too large is illustrated by the large jumps in *MSE* as shown for $\Delta = 0.115$. As Δ is decreased, the convergence is slowed somewhat, but a lower MSE is achieved, indicating that the estimated coefficients are closer to $\boldsymbol{c}_{\text{opt}}$.

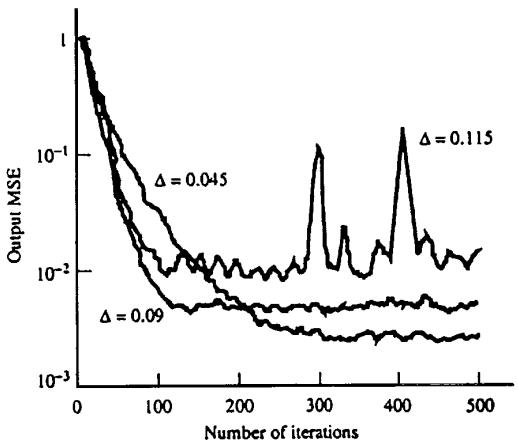


FIGURE 5.14 Initial convergence characteristics of the *LMS* algorithm with different step sizes.

Although we have described in some detail the operation of an adaptive equalizer that is optimized on the basis of the MSE criterion, the operation of an adaptive equalizer based on the zero-forcing method is very similar. The major difference lies in the method for estimating the gradient vectors $\boldsymbol{g}_k$ at each iteration. A block diagram of an adaptive zero-forcing equalizer is shown in Fig. 5.15.

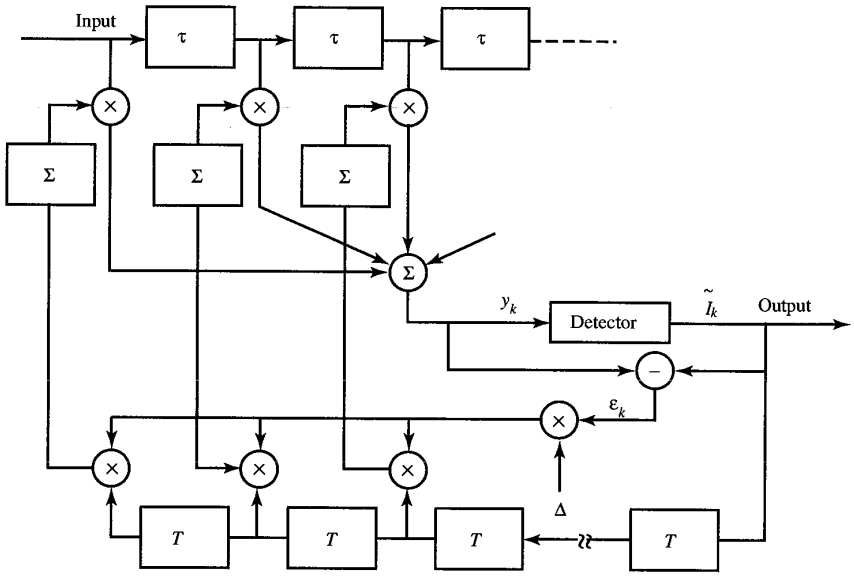


FIGURE 5.15 An adaptive zero-forcing equalizer.

For more details on the tap coefficient update method for a zero-forcing equalizer, the reader is referred to the papers by Lucky [2, 3], and the text by Proakis [4].

5.4 Decision-Feedback Equalizer

The linear filter equalizers described in the preceding section are very effective on channels, such as wire line telephone channels, where the ISI is not severe. The severity of the ISI is directly related to the spectral characteristics and not necessarily to the time span of the ISI. For example, consider the ISI resulting from the two channels that are illustrated in Fig. 5.16. The time span for the ISI in channel A is 5 symbol intervals on each side of the desired signal component, which has a value of 0.72. On the other hand, the time span for the ISI in channel B is one symbol interval on each side of the desired signal component, which has a value of 0.815. The energy of the total response is normalized to unity for both channels.

In spite of the shorter ISI span, channel B results in more severe ISI. This is evidenced in the frequency response characteristics of these channels, which are shown in Fig. 5.17. We observe that channel B has a spectral null [the frequency response $C(f) = 0$ for some frequencies in the band $|f| \leq W$] at $f = 1/2T$, whereas this does not occur in the case of channel A. Consequently, a linear equalizer will introduce a large gain in its frequency response to compensate for the channel null. Thus, the noise in channel B will be enhanced

Ford Motor Co.
Exhibit 1030
Page 183

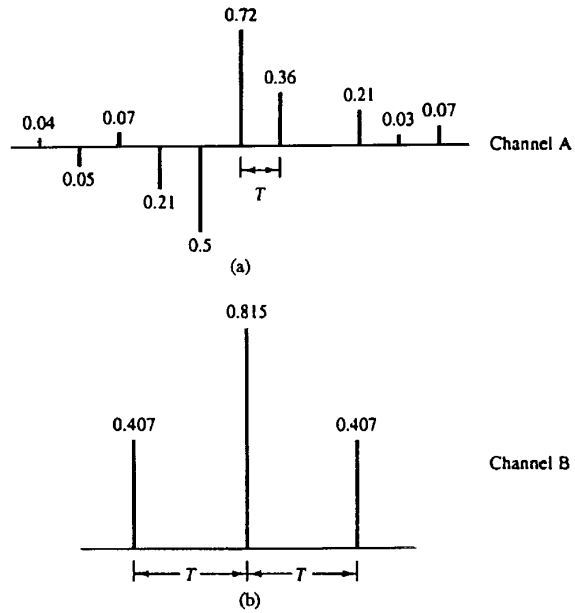


FIGURE 5.16 Two channels with ISI.

much more than in channel A. This implies that the performance of the linear equalizer for channel B will be sufficiently poorer than that for channel A. This fact is borne out by the computer simulation results for the performance of the two linear equalizers shown in Fig. 5.18. Hence, the basic limitation of a linear equalizer is that it performs poorly on channels having spectral nulls. Such channels are often encountered in radio communications, such as ionospheric transmission at frequencies below 30 MHz and mobile radio channels, such as those used for cellular radio communications.

A **decision-feedback equalizer** (**DFE**) is a nonlinear equalizer that employs previous decisions to eliminate the ISI caused by previously detected symbols on the current symbol to be detected. A simple block diagram for a DFE is shown in Fig. 5.19. The DFE consists of two filters. The first filter is called a *feedforward filter* and it is generally a fractionally spaced FIR filter with adjustable tap coefficients. This filter is identical in form to the linear equalizer already described. Its input is the received filtered signal $y(t)$ sampled at some rate that is a multiple of the symbol rate, e.g., at rate $2/T$. The second filter is a *feedback filter*. It is implemented as an FIR filter with symbol-spaced taps having adjustable coefficients. Its input is the set of previously detected symbols. The output of the feedback filter is subtracted from the output of the feedforward filter to form the input to the detector. Thus, we have

$$z_m = \sum_{n=-N_1}^{0} c_n y(mT - n\tau) - \sum_{n=1}^{N_2} b_n \tilde{I}_{m-n} \tag{5.41}$$

where $\{c_n\}$ and $\{b_n\}$ are the adjustable coefficients of the feedforward and feedback filters, respectively, $\tilde{I}_{m-n}, n = 1, 2, \ldots, N_2$ are the previously detected symbols, N_1+1 is the length of the feedforward filter, and N_2 is the length of the feedback filter. Based on the input z_m, the detector determines which of the possible transmitted symbols is closest in distance to the input signal I_m. Thus, it makes its decision and outputs $\tilde{I}_m$. What makes the

Ford Motor Co.
Exhibit 1030
Page 184

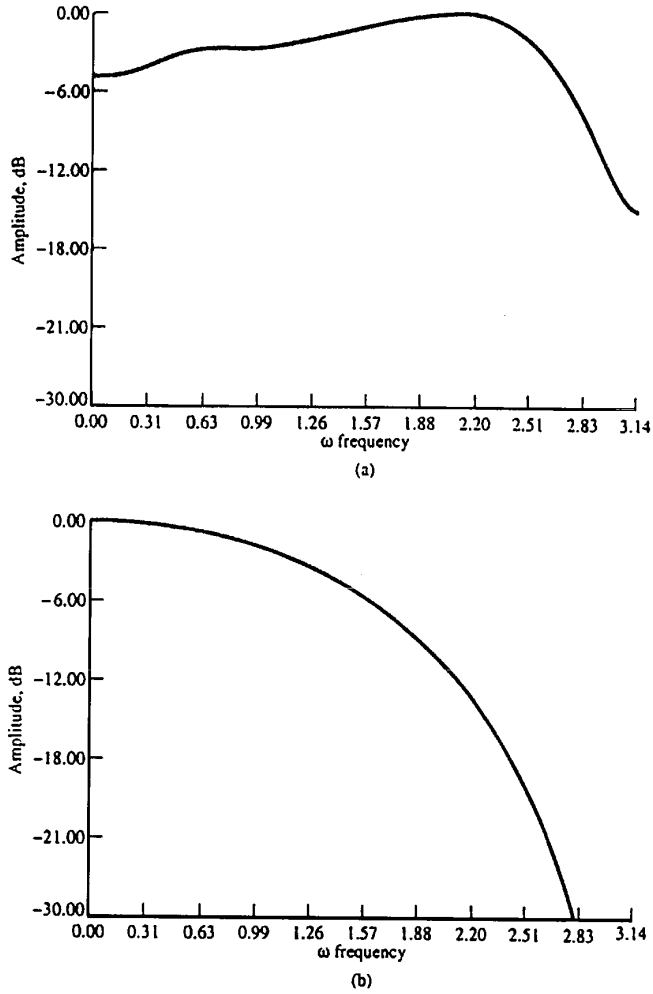


FIGURE 5.17 Amplitude spectra for (a) channel A shown in Fig.5.16(a) and (b) channel B shown in Fig.5.16(b).

DFE nonlinear is the nonlinear characteristic of the detector that provides the input to the feedback filter.

The tap coefficients of the feedforward and feedback filters are selected to optimize some desired performance measure. For mathematical simplicity, the *MSE* criterion is usually applied, and a stochastic gradient algorithm is commonly used to implement an adaptive DFE. Figure 5.20 illustrates the block diagram of an adaptive DFE whose tap coefficients are adjusted by means of the LMS stochastic gradient algorithm. Figure 5.21 illustrates the probability of error performance of the DFE, obtained by computer simulation, for binary PAM transmission over channel B. The gain in performance relative to that of a linear equalizer is clearly evident.

We should mention that decision errors from the detector that are fed to the feedback filter have a small effect on the performance of the DFE. In general, a small loss in performance

Ford Motor Co.
Exhibit 1030
Page 185

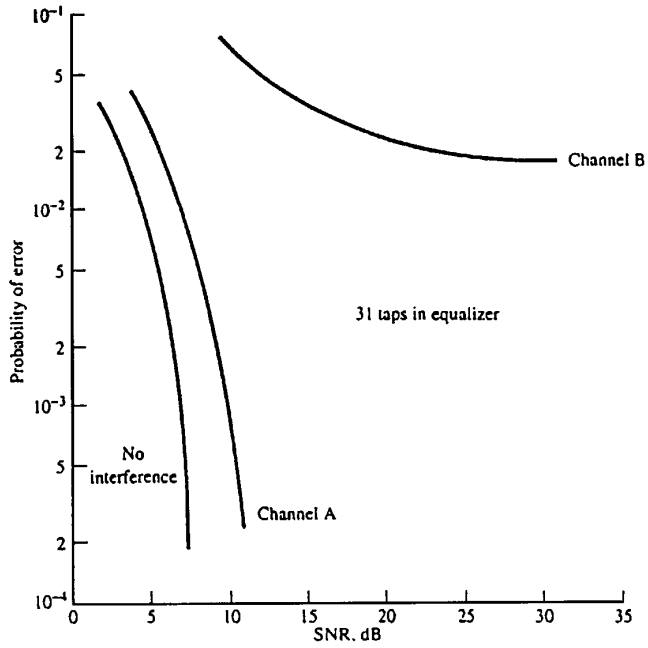


FIGURE 5.18 Error-rate performance of linear *MSE* equalizer.

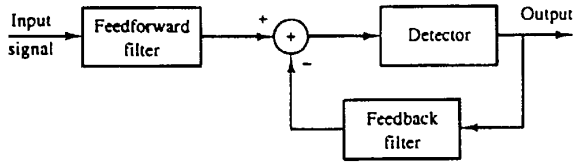


FIGURE 5.19 Block diagram of DFE.

of one to two decibels is possible at error rates below 10^{-2}, as illustrated in Fig. 5.21, but the decision errors in the feedback filters are not catastrophic.

5.5 Maximum-Likelihood Sequence Detection

Although the DFE outperforms a linear equalizer, it is not the optimum equalizer from the viewpoint of minimizing the probability of error in the detection of the information sequence $\{I_k\}$ from the received signal samples $\{y_k\}$ given in Eq. (5.5). In a digital communication system that transmits information over a channel that causes ISI, the optimum detector is a maximum-likelihood symbol sequence detector which produces at its output the most probable symbol sequence $\{\tilde{I}_k\}$ for the given received sampled sequence $\{y_k\}$. That is, the detector finds the sequence $\{\tilde{I}_k\}$ that maximizes the *likelihood function*

$$\Lambda(\{I_k\}) = \ln p(\{y_k\} \mid \{I_k\}) \tag{5.42}$$

where $p(\{y_k\} \mid \{I_k\})$ is the joint probability of the received sequence $\{y_k\}$ conditioned on $\{I_k\}$. The sequence of symbols $\{\tilde{I}_k\}$ that maximizes this joint conditional probability is called the **maximum-likelihood sequence detector.**

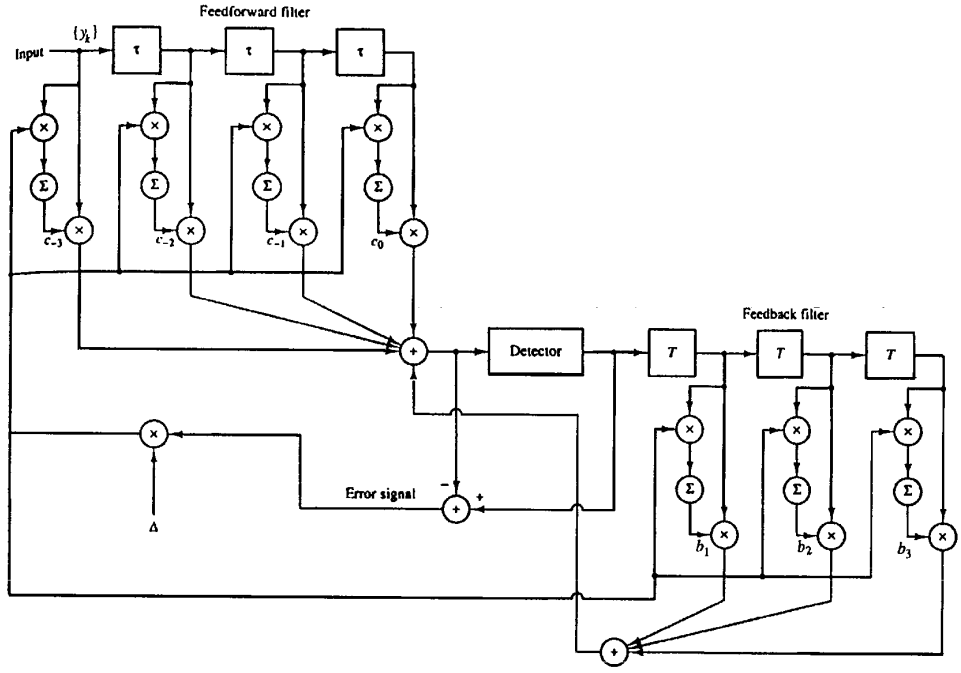


FIGURE 5.20 Adaptive DFE.

An algorithm that implements maximum-likelihood sequence detection (MLSD) is the Viterbi algorithm, which was originally devised for decoding convolutional codes. For a description of this algorithm in the context of sequence detection in the presence of ISI, the reader is referred to the paper by Forney [1] and the text by Proakis [4].

The major drawback of MLSD for channels with ISI is the exponential behavior in computational complexity as a function of the span of the ISI. Consequently, MLSD is practical only for channels where the ISI spans only a few symbols and the ISI is severe, in the sense that it causes a severe degradation in the performance of a linear equalizer or a decision-feedback equalizer. For example, Fig. 5.22 illustrates the error probability performance of the Viterbi algorithm for a binary PAM signal transmitted through channel B (see Fig. 5.16). For purposes of comparison, we also illustrate the probability of error for a DFE. Both results were obtained by computer simulation. We observe that the performance of the maximum likelihood sequence detector is about 4.5 dB better than that of the DFE at an error probability of 10^{-4}. Hence, this is one example where the ML sequence detector provides a significant performance gain on a channel with a relatively short ISI span.

5.6 Conclusions

Channel equalizers are widely used in digital communication systems to mitigate the effects of ISI caused by channel distortion. Linear equalizers are widely used for high-speed modems that transmit data over telephone channels. For wireless (radio) transmission, such as in mobile cellular communications and interoffice communications, the multipath propagation of the transmitted signal results in severe ISI. Such channels require more powerful equalizers to combat the severe ISI. The decision-feedback equalizer and the MLSD are two nonlinear channel equalizers that are suitable for radio channels with severe ISI.

Ford Motor Co.
Exhibit 1030
Page 187

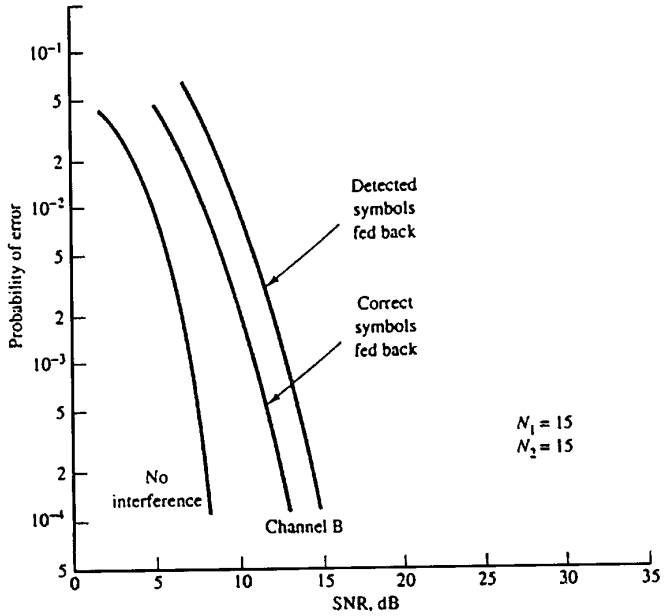


FIGURE 5.21 Performance of DFE with and without error propagation.

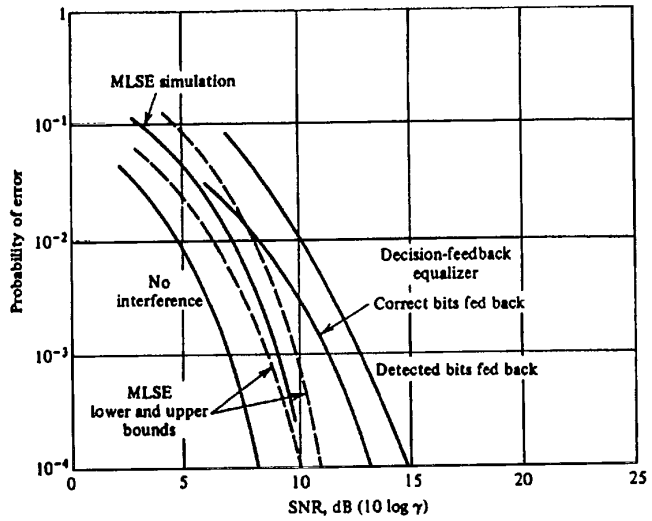


FIGURE 5.22 Comparison of performance between MLSE and decision-feedback equalization for channel B of Fig.5.16.

Ford Motor Co.
Exhibit 1030
Page 188

Defining Terms

Adaptive equalizer: A channel equalizer whose parameters are updated automatically and adaptively during transmission of data.

Channel equalizer: A device that is used to reduce the effects of channel distortion in a received signal.

Decision-directed mode: Mode for adjustment of the equalizer coefficient adaptively based on the use of the detected symbols at the output of the detector.

Decision-feedback equalizer (DFE): An adaptive equalizer that consists of a feedforward filter and a feedback filter, where the latter is fed with previously detected symbols that are used to eliminate the intersymbol interference due to the tail in the channel impulse response.

Fractionally spaced equalizer: A tapped-delay line channel equalizer in which the delay between adjacent taps is less than the duration of a transmitted symbol.

Intersymbol interference: Interference in a received symbol from adjacent (nearby) transmitted symbols caused by channel distortion in data transmission.

LMS algorithm: See stochastic gradient algorithm.

Maximum-likelihood sequence detector: A detector for estimating the most probable sequence of data symbols by maximizing the likelihood function of the received signal.

Preset equalizer: A channel equalizer whose parameters are fixed (time-invariant) during transmission of data.

Stochastic gradient algorithm: An algorithm for adaptively adjusting the coefficients of an equalizer based on the use of (noise-corrupted) estimates of the gradients.

Symbol-spaced equalizer: A tapped-delay line channel equalizer in which the delay between adjacent taps is equal to the duration of a transmitted symbol.

Training mode: Mode for adjustment of the equalizer coefficients based on the transmission of a known sequence of transmitted symbols.

Zero-forcing equalizer: A channel equalizer whose parameters are adjusted to completely eliminate intersymbol interference in a sequence of transmitted data symbols.

References

[1] Forney, G.D., Jr., Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *IEEE Trans. Inform. Theory,* IT-18, 363–378, May 1972.

[2] Lucky, R.W., Automatic equalization for digital communications. *Bell Syst. Tech. J.,* 44, 547–588, Apr. 1965.

[3] Lucky, R.W., Techniques for adaptive equalization of digital communication. *Bell Syst. Tech. J.,* 45, 255–286, Feb. 1966.

[4] Proakis, J.G., *Digital Communications,* 3rd ed., McGraw-Hill, New York, 1995.

Further Information

For a comprehensive treatment of adaptive equalization techniques and their performance characteristics, the reader may refer to the book by Proakis [4]. The two papers by

Ford Motor Co.
Exhibit 1030
Page 189

Lucky [2, 3], provide a treatment on linear equalizers based on the zero-forcing criterion. Additional information on decision-feedback equalizers may be found in the journal papers "An Adaptive Decision-Feedback Equalizer" by D.A. George, R.R. Bowen, and J.R. Storey, *IEEE Transactions on Communications Technology,* Vol. COM-19, pp. 281–293, June 1971, and "Feedback Equalization for Fading Dispersive Channels" by P. Monsen, *IEEE Transactions on Information Theory,* Vol. IT-17, pp. 56–64, January 1971. A through treatment of channel equalization based on maximum-likelihood sequence detection is given in the paper by Forney [1].

Ford Motor Co.
Exhibit 1030
Page 190

6

Line Coding

6.1 Introduction 6-1
6.2 Common Line Coding Formats 6-2
Unipolar NRZ (Binary On-Off Keying) • Unipolar RZ • Polar NRZ • Polar RZ [Bipolar, Alternate Mark Inversion (AMI), or Pseudoternary] • Manchester Coding (Split Phase or Digital Biphase)
6.3 Alternate Line Codes 6-10
Delay Modulation (Miller Code) • Split Phase (Mark) • Biphase (Mark) • Code Mark Inversion (CMI) • NRZ (I) • Binary N Zero Substitution (BNZS) • High-Density Bipolar N (HDBN) • Ternary Coding
6.4 Multilevel Signalling, Partial Response Signalling, and Duobinary Coding 6-16
Multilevel Signalling • Partial Response Signalling and Duobinary Coding
6.5 Bandwidth Comparison 6-18
6.6 Concluding Remarks 6-18
Defining Terms 6-18
References 6-19

Joseph L. LoCicero
Illinois Institute of Technology

Bhasker P. Patel
Illinois Institute of Technology

6.1 Introduction

The terminology **line coding** originated in telephony with the need to transmit digital information across a copper telephone *line;* more specifically, binary data over a digital repeatered line. The concept of line coding, however, readily applies to any transmission line or channel. In a digital communication system, there exists a known set of symbols to be transmitted. These can be designated as $\{m_i\}$, $i = 1, 2, \ldots, N$, with a probability of occurrence $\{p_i\}$, $i = 1, 2, \ldots, N$, where the sequentially transmitted symbols are generally assumed to be statistically independent. The conversion or *coding* of these abstract symbols into real, temporal waveforms to be transmitted in baseband is the process of line coding. Since the most common type of line coding is for binary data, such a waveform can be succinctly termed a direct format for serial bits. The concentration in this section will be line coding for binary data.

Different channel characteristics, as well as different applications and performance requirements, have provided the impetus for the development and study of various types of line coding [1, 2]. For example, the channel might be ac coupled and, thus, could not support a line code with a dc component or large dc content. Synchronization or timing

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 191

recovery requirements might necessitate a discrete component at the data rate. The channel bandwidth and **crosstalk** limitations might dictate the type of line coding employed. Even such factors as the complexity of the encoder and the economy of the decoder could determine the line code chosen. Each line code has its own distinct properties. Depending on the application, one property may be more important than the other. In what follows, we describe, in general, the most desirable features that are considered when choosing a line code.

It is commonly accepted [1, 2, 5, 8] that the dominant considerations effecting the choice of a line code are: 1) timing, 2) dc content, 3) power spectrum, 4) performance monitoring, 5) probability of error, and 6) transparency. Each of these are detailed in the following paragraphs.

1) *Timing:* The waveform produced by a line code should contain enough timing information such that the receiver can synchronize with the transmitter and decode the received signal properly. The timing content should be relatively independent of source statistics, i.e., a long string of **1**s or **0**s should not result in loss of timing or jitter at the receiver.

2) *DC content:* Since the repeaters used in telephony are ac coupled, it is desirable to have zero dc in the waveform produced by a given line code. If a signal with significant dc content is used in ac coupled lines, it will cause **dc wander** in the received waveform. That is, the received signal baseline will vary with time. Telephone lines do not pass dc due to ac coupling with transformers and capacitors to eliminate dc ground loops. Because of this, the telephone channel causes a droop in constant signals. This causes dc wander. It can be eliminated by dc restoration circuits, feedback systems, or with specially designed line codes.

3) *Power spectrum:* The power spectrum and bandwidth of the transmitted signal should be matched to the frequency response of the channel to avoid significant distortion. Also, the power spectrum should be such that most of the energy is contained in as small bandwidth as possible. The smaller is the bandwidth, the higher is the transmission efficiency.

4) *Performance monitoring:* It is very desirable to detect errors caused by a noisy transmission channel. The error detection capability in turn allows performance monitoring while the channel is in use (i.e., without elaborate testing procedures that require suspending use of the channel).

5) *Probability of error:* The average error probability should be as small as possible for a given transmitter power. This reflects the reliability of the line code.

6) *Transparency:* A line code should allow all the possible patterns of **1**s and **0**s. If a certain pattern is undesirable due to other considerations, it should be mapped to a unique alternative pattern.

6.2 Common Line Coding Formats

A line coding format consists of a formal definition of the line code that specifies how a string of binary digits are converted to a line code waveform. There are two major classes of binary line codes: **level codes** and **transition codes.** Level codes carry information in their voltage level, which may be high or low for a full bit period or part of the bit period. Level codes are usually instantaneous since they typically encode a binary digit into a distinct waveform, independent of any past binary data. However, some level codes do exhibit memory. Transition codes carry information in the change in level appearing in the line code waveform. Transition codes may be instantaneous, but they generally have memory, using past binary data to dictate the present waveform. There are two common forms of level line codes: one is called **return to zero (RZ)** and the other is called

Ford Motor Co.
Exhibit 1030
Page 192

nonreturn to zero (NRZ). In RZ coding, the level of the pulse returns to zero for a portion of the bit interval. In NRZ coding, the level of the pulse is maintained during the entire bit interval.

Line coding formats are further classified according to the polarity of the voltage levels used to represent the data. If only one polarity of voltage level is used, i.e., positive or negative (in addition to the zero level) then it is called **unipolar** signalling. If both positive and negative voltage levels are being used, with or without a zero voltage level, then it is called **polar** signalling. The term **bipolar** signalling is used by some authors to designate a specific line coding scheme with positive, negative, and zero voltage levels. This will be described in detail later in this section. The formal definition of five common line codes is given in the following along with a representative waveform, the *power spectral density* (PSD), the probability of error, and a discussion of advantages and disadvantages. In some cases specific applications are noted.

6.2.1 Unipolar NRZ (Binary On-Off Keying)

In this line code, a binary **1** is represented by a non-zero voltage level and a binary **0** is represented by a zero voltage level as shown in Fig. 6.1(a). This is an instantaneous level code. The PSD of this code with equally likely **1**s and **0**s is given by [5, 8]

$$S_1(f) = \frac{V^2T}{4}\left(\frac{\sin \pi fT}{\pi fT}\right)^2 + \frac{V^2}{4}\delta(f) \tag{6.1}$$

where V is the binary **1** voltage level, $T = 1/R$ is the bit duration, and R is the bit rate in bits per second. The spectrum of unipolar NRZ is plotted in Fig. 6.2a. This PSD is a two-sided even spectrum, although only half of the plot is shown for efficiency of presentation. If the probability of a binary **1** is p, and the probability of a binary **0** is $(1-p)$, then the PSD of this code, in the most general case, is $4p(1-p)\,S_1(f)$. Considering the frequency of the first spectral null as the bandwidth of the waveform, the bandwidth of unipolar NRZ is R in hertz. The error rate performance of this code, for equally likely data, with additive white Gaussian noise (AWGN) and optimum, i.e., matched filter, detection is given by [1, 5]

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right) \tag{6.2}$$

where E_b/N_0 is a measure of the signal-to-noise ratio (SNR) of the received signal. In general, E_b is the energy per bit and $N_0/2$ is the two-sided PSD of the AWGN. More specifically, for unipolar NRZ, E_b is the energy in a binary **1**, which is V^2T. The performance of the unipolar NRZ code is plotted in Fig. 6.3

The principle advantages of unipolar NRZ are ease of generation, since it requires only a single power supply, and a relatively low bandwidth of R Hz. There are quite a few disadvantages of this line code. A loss of synchronization and timing jitter can result with a long sequence of **1**s or **0**s because no pulse transition is present. The code has no error detection capability and, hence, performance cannot be monitored. There is a significant dc component as well as a dc content. The error rate performance is not as good as that of polar line codes.

6.2.2 Unipolar RZ

In this line code, a binary **1** is represented by a nonzero voltage level during a portion of the bit duration, usually for half of the bit period, and a zero voltage level for rest of the bit

Ford Motor Co.
Exhibit 1030
Page 193

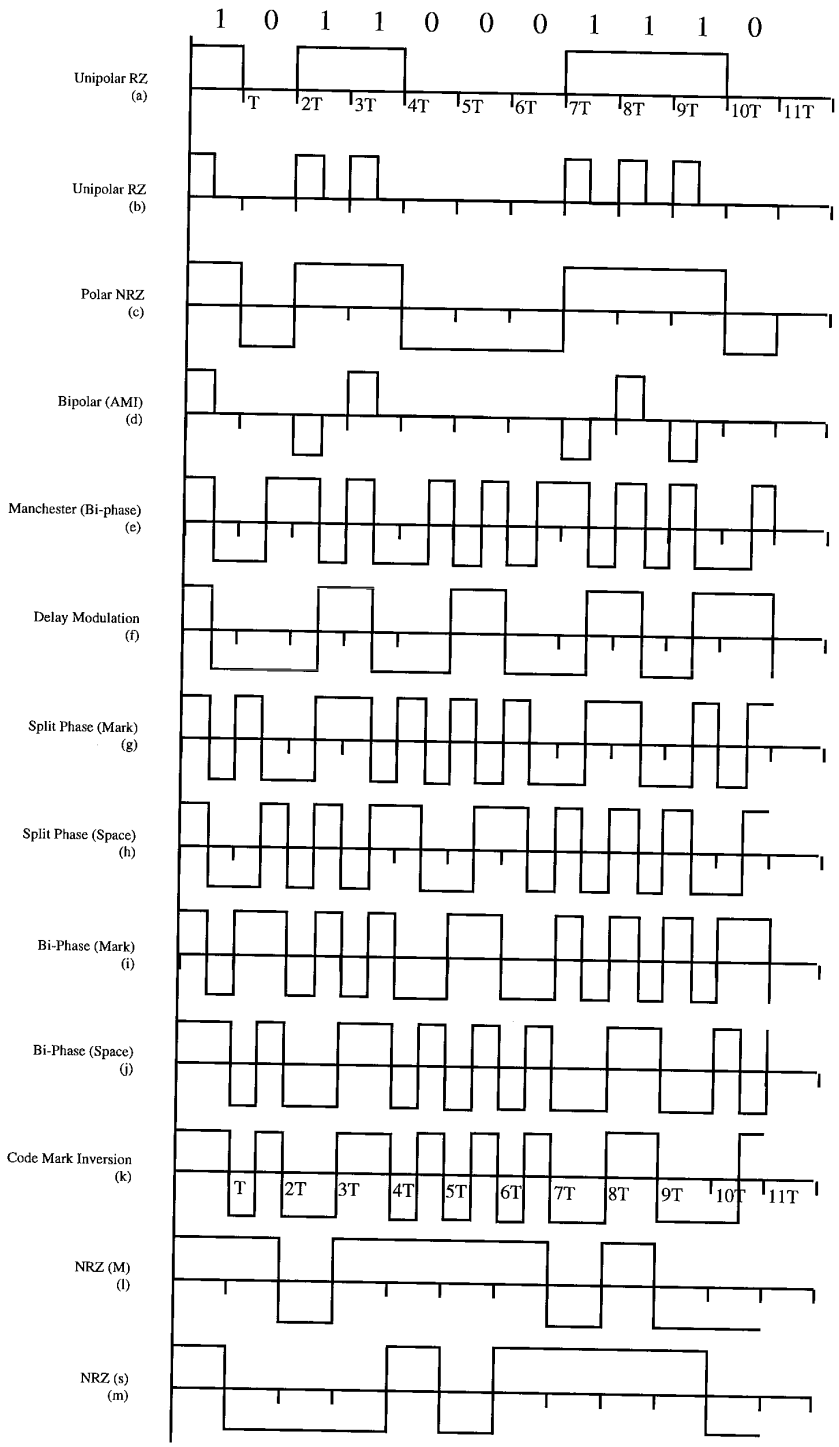


FIGURE 6.1 Waveforms for different line codes.

Ford Motor Co.
Exhibit 1030
Page 194

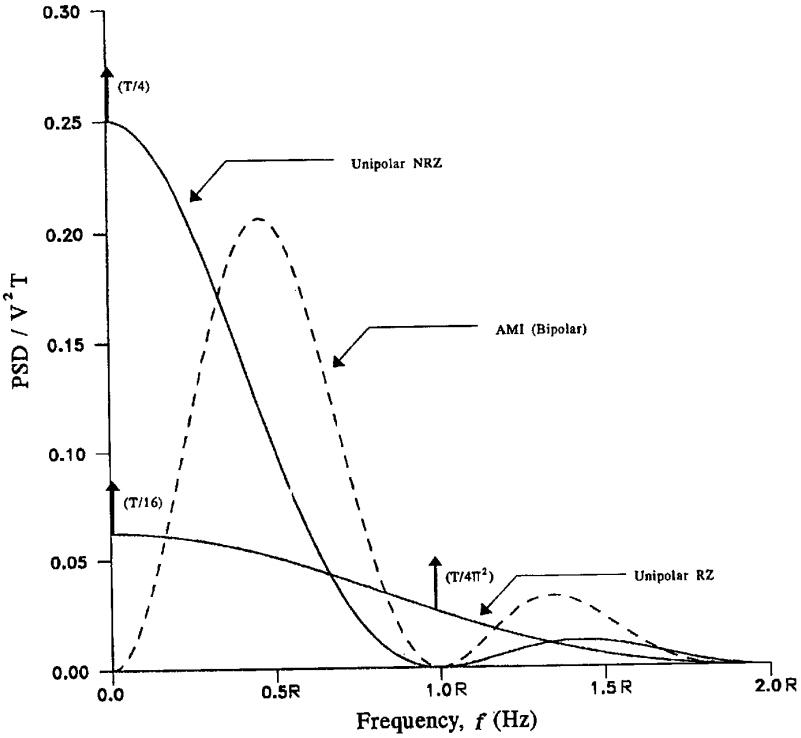


Figure 6.2a Power spectral density of different line codes, where $R = 1/\mathrm{T}$ is the bit rate.

duration. A binary **0** is represented by a zero voltage level during the entire bit duration. Thus, this is an instantaneous level code. Figure 6.1(b) illustrates a unipolar RZ waveform in which the **1** is represented by a nonzero voltage level for half the bit period. The PSD of this line code, with equally likely binary digits, is given by [5, 6, 8]

$$S_2(f) = \frac{V^2T}{16}\left(\frac{\sin \pi fT/2}{\pi fT/2}\right)^2 + \frac{V^2}{4\pi^2}\left[\frac{\pi^2}{4}\delta(f) + \sum_{n=-\infty}^{\infty}\frac{1}{(2n+1)^2}\delta(f-(2n+1)R)\right] \tag{6.3}$$

where again V is the binary **1** voltage level, and $T = 1/R$ is the bit period. The spectrum of this code is drawn in Fig. 6.2a. In the most general case, when the probability of a **1** is p, the continuous portion of the PSD in Eq. (6.3) is scaled by the factor $4p(1-p)$ and the discrete portion is scaled by the factor $4p^2$. The first null bandwidth of unipolar RZ is $2R$ Hz. The error rate performance of this line code is the same as that of the unipolar NRZ provided we increase the voltage level of this code such that the energy in binary **1**, E_b, is the same for both codes. The probability of error is given by Eq. (6.2) and identified in Fig. 6.3. If the voltage level and bit period are the same for unipolar NRZ and unipolar RZ, then the energy in a binary **1** for unipolar RZ will be $V^2T/2$ and the probability of error is worse by 3 dB.

The main advantages of unipolar RZ are, again, ease of generation since it requires a single power supply and the presence of a discrete spectral component at the symbol rate, which

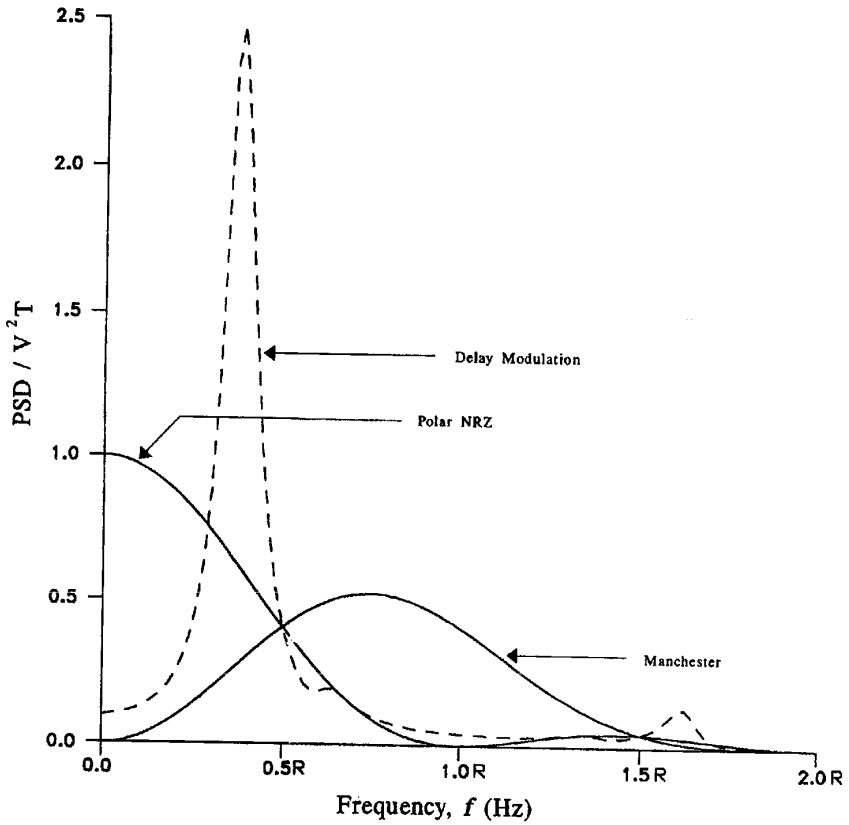


Figure 6.2b Power spectral density of different line codes, where $R = 1/\mathrm{T}$ is the bit rate.

allows simple timing recovery. A number of disadvantages exist for this line code. It has a nonzero dc component and nonzero dc content, which can lead to dc wander. A long string of **0**s will lack pulse transitions and could lead to loss of synchronization. There is no error detection capability and, hence, performance monitoring is not possible. The bandwidth requirement ($2R$ Hz) is higher than that of NRZ signals. The error rate performance is worse than that of polar line codes.

Unipolar NRZ as well as unipolar RZ are examples of pulse/no-pulse type of signalling. In this type of signalling, the pulse for a binary **0**, $g_2(t)$, is zero and the pulse for a binary **1** is specified generically as $g_1(t) = g(t)$. Using $G(f)$ as the Fourier transform of $g(t)$, the PSD of pulse/no-pulse signalling is given as [6, 7, 10]

$$S_{\mathrm{PNP}}(f) = p(1-p)R|G(f)|^2 + p^2R^2 \sum_{n=-\infty}^{\infty} |G(nR)|^2 \delta(f - nR) \tag{6.4}$$

where p is the probability of a binary **1**, and R is the bit rate.

6.2.3 Polar NRZ

In this line code, a binary **1** is represented by a positive voltage $+V$ and a binary **0** is represented by a negative voltage $-V$ over the full bit period. This code is also referred to

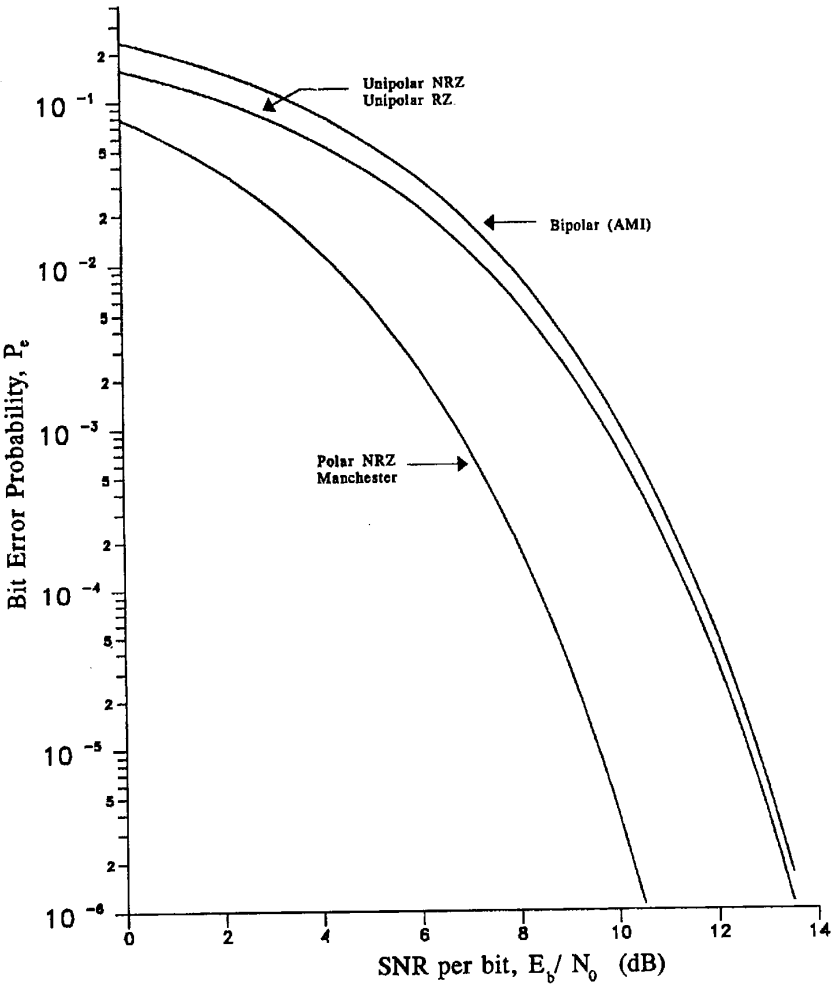


FIGURE 6.3 Bit error probability for different line codes.

as NRZ (L), since a bit is represented by maintaining a level (L) during its entire period. A polar NRZ waveform is shown in Fig. 6.1(c). This is again an instantaneous level code. Alternatively, a **1** may be represented by a $-V$ voltage level and a **0** by a $+V$ voltage level, without changing the spectral characteristics and performance of the line code. The PSD of this line code with equally likely bits is given by [5, 8]

$$S_3(f) = V^2T\left(\frac{\sin \pi fT}{\pi fT}\right)^2 \tag{6.5}$$

This is plotted in Fig. 6.2b. When the probability of a **1** is p, and p is not 0.5, a dc component exists, and the PSD becomes [10]

$$S_{3p}(f) = 4V^2Tp(1-p)\left(\frac{\sin \pi fT}{\pi fT}\right)^2 + V^2(1-2p)^2\delta(f) \tag{6.6}$$

The first null bandwidth for this line code is again R Hz, independent of p. The probability of error of this line code when $p = 0.5$ is given by [1, 5]

$$P_e = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{6.7}$$

The performance of polar NRZ is plotted in Fig. 6.3. This is better than the error performance of the unipolar codes by 3 dB.

The advantages of polar NRZ include a low-bandwidth requirement, R Hz, comparable to unipolar NRZ, very good error probability, and greatly reduced dc because the waveform has a zero dc component when $p = 0.5$ even though the dc content is never zero. A few notable disadvantages are that there is no error detection capability, and that a long string of **1**s or **0**s could result in loss of synchronization, since there are no transitions during the string duration. Two power supplies are required to generate this code.

6.2.4 Polar RZ [Bipolar, Alternate Mark Inversion (AMI), or Pseudoternary]

In this scheme, a binary **1** is represented by alternating the positive and negative voltage levels, which return to zero for a portion of the bit duration, generally half the bit period. A binary **0** is represented by a zero voltage level during the entire bit duration. This line coding scheme is often called **alternate mark inversion (AMI)** since **1**s (marks) are represented by alternating positive and negative pulses. It is also called *pseudoternary* since three different voltage levels are used to represent binary data. Some authors designate this line code as bipolar RZ (BRZ). An AMI waveform is shown in Fig. 6.1(d). Note that this is a level code with memory. The AMI code is well known for its use in telephony. The PSD of this line code with memory is given by [1, 2, 7]

$$S_{4p}(f) = 2p(1-p)R|G(f)|^2\left(\frac{1-\cos 2\pi fT}{1+(2p-1)^2+2(2p-1)\cos 2\pi fT}\right) \tag{6.8}$$

where $G(f)$ is the Fourier transform of the pulse used to represent a binary **1**, and p is the probability of a binary **1**. When $p = 0.5$ and square pulses with amplitude $\pm V$ and duration $T/2$ are used to represent binary **1**s, the PSD becomes

$$S_4(f) = \frac{V^2T}{4}\left(\frac{\sin \pi fT/2}{\pi fT/2}\right)^2 \sin^2(\pi fT) \tag{6.9}$$

This PSD is plotted in Fig. 6.2a. The first null bandwidth of this waveform is R Hz. This is true for RZ rectangular pulses, independent of the value of p in Eq. (6.8). The error rate performance of this line code for equally likely binary data is given by [5]

$$P_e \approx \frac{3}{4}\text{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right), \qquad E_b/N_0 > 2 \tag{6.10}$$

This curve is plotted in Fig. 6.3 and is seen to be no more than 0.5 dB worse than the unipolar codes.

The advantages of polar RZ (or AMI, as it is most commonly called) outweigh the disadvantages. This code has no dc component and zero dc content, completely avoiding the dc wander problem. Timing recovery is rather easy since squaring, or full-wave rectifying, this type of signal yields a unipolar RZ waveform with a discrete component at the bit rate, R Hz. Because of the alternating polarity pulses for binary **1**s, this code has error detection

Ford Motor Co.
Exhibit 1030
Page 198

and, hence, performance monitoring capability. It has a low-bandwidth requirement, R Hz, comparable to unipolar NRZ. The obvious disadvantage is that the error rate performance is worse than that of the unipolar and polar waveforms. A long string of **0**s could result in loss of synchronization, and two power supplies are required for this code.

6.2.5 Manchester Coding (Split Phase or Digital Biphase)

In this coding, a binary **1** is represented by a pulse that has positive voltage during the first-half of the bit duration and negative voltage during second-half of the bit duration. A binary **0** is represented by a pulse that is negative during the first-half of the bit duration and positive during the second-half of the bit duration. The negative or positive midbit transition indicates a binary **1** or binary **0**, respectively. Thus, a Manchester code is classified as an instantaneous transition code; it has no memory. The code is also called diphase because a square wave with a 0° phase is used to represent a binary **1** and a square wave with a phase of 180° used to represent a binary **0**; or vice versa. This line code is used in Ethernet local area networks (LANs). The waveform for Manchester coding is shown in Fig. 6.1(e). The PSD of a Manchester waveform with equally likely bits is given by [5, 8]

$$S_5(f) = V^2T\left(\frac{\sin \pi fT/2}{\pi fT/2}\right)^2 \sin^2(\pi fT/2) \tag{6.11}$$

where $\pm V$ are used as the positive/negative voltage levels for this code. Its spectrum is plotted in Fig. 6.2b. When the probability p of a binary **1**, is not equal to one-half, the continuous portion of the PSD is reduced in amplitude and discrete components appear at integer multiples of the bit rate, $R = 1/T$. The resulting PSD is [6, 10]

$$\begin{aligned} S_{5p}(f) &= V^2T4p(1-p)\left(\frac{\sin \pi fT/2}{\pi fT/2}\right)^2 \sin^2\frac{\pi fT}{2} \\ &\quad + V^2(1-2p)^2 \sum_{n=-\infty, n\neq 0}^{\infty} \left(\frac{2}{n\pi}\right)^2 \delta(f-nR) \end{aligned} \tag{6.12}$$

The first null bandwidth of the waveform generated by a Manchester code is $2R$ Hz. The error rate performance of this waveform when $p = 0.5$ is the same as that of polar NRZ, given by Eq. (6.9), and plotted in Fig. 6.3.

The advantages of this code include a zero dc content on an individual pulse basis, so no pattern of bits can cause dc buildup; midbit transitions are always present making it is easy to extract timing information; and it has good error rate performance, identical to polar NRZ. The main disadvantage of this code is a larger bandwidth than any of the other common codes. Also, it has no error detection capability and, hence, performance monitoring is not possible.

Polar NRZ and Manchester coding are examples of the use of pure polar signalling where the pulse for a binary **0**, $g_2(t)$ is the negative of the pulse for a binary **1**, i.e., $g_2(t) = -g_1(t)$. This is also referred to as an antipodal signal set. For this broad type of polar binary line code, the PSD is given by [10]

$$S_{\mathrm{BP}}(f) = 4p(1-p)R|G(f)|^2 + (2p-1)^2R^2 \sum_{n=-\infty}^{\infty} |G(nR)|^2 \delta(f-nR) \tag{6.13}$$

where $|G(f)|$ is the magnitude of the Fourier transform of either $g_1(t)$ or $g_2(t)$.

Ford Motor Co.
Exhibit 1030
Page 199

A further generalization of the PSD of binary line codes can be given, wherein a continuous spectrum and a discrete spectrum is evident. Let a binary **1,** with probability p, be represented by $g_1(t)$ over the $T = 1/R$ second bit interval; and let a binary **0,** with probability $1 - p$, be represented by $g_2(t)$ over the same T second bit interval. The two-sided PSD for this general binary line code is [10]

$$\begin{aligned} S_{GB}(f) &= p(1-p)R|G_1(f) - G_2(f)|^2 \\ &\quad + R^2 \sum_{n=-\infty}^{\infty} |pG_1(nR) + (1-p)G_2(nR)|^2 \, \delta(f - nR) \end{aligned} \tag{6.14}$$

where the Fourier transform of $g_1(t)$ and $g_2(t)$ are given by $G_1(f)$ and $G_2(f)$, respectively.

6.3 Alternate Line Codes

Most of the line codes discussed thus far were instantaneous level codes. Only AMI had memory, and Manchester was an instantaneous transition code. The alternate line codes presented in this section all have memory. The first four are transition codes, where binary data is represented as the presence or absence of a transition, or by the direction of transition, i.e., positive to negative or vice versa. The last four codes described in this section are level line codes with memory.

6.3.1 Delay Modulation (Miller Code)

In this line code, a binary **1** is represented by a transition at the midbit position, and a binary **0** is represented by no transition at the midbit position. If a **0** is followed by another **0,** however, the signal transition also occurs at the end of the bit interval, that is, between the two **0**s. An example of delay modulation is shown in Fig. 6.1(f). It is clear that delay modulation is a transition code with memory. This code achieves the goal of providing good timing content without sacrificing bandwidth. The PSD of the Miller code for equally likely data is given by [10]

$$\begin{aligned} S_6(f) &= \frac{V^2 T}{2(\pi fT)^2(17 + 8\cos 2\pi fT)} \\ &\quad \times (23 - 2\cos \pi fT - 22\cos 2\pi fT \\ &\quad - 12\cos 3\pi fT + 5\cos 4\pi fT + 12\cos 5\pi fT \\ &\quad + 2\cos 6\pi fT - 8\cos 7\pi fT + 2\cos 8\pi fT) \end{aligned} \tag{6.15}$$

This spectrum is plotted in Fig. 6.2b. The advantages of this code are that it requires relatively low bandwidth, most of the energy is contained in less than $0.5R$. However, there is no distinct spectral null within the $2R$-Hz band. It has low dc content and no dc component. It has very good timing content, and carrier tracking is easier than Manchester coding. Error rate performance is comparable to that of the common line codes. One important disadvantage is that it has no error detection capability and, hence, performance cannot be monitored.

6.3.2 Split Phase (Mark)

This code is similar to Manchester in the sense that there are always midbit transitions. Hence, this code is relatively easy to synchronize and has no dc. Unlike Manchester, however, split phase (mark) encodes a binary digit into a midbit transition dependent on the

Ford Motor Co.
Exhibit 1030
Page 200

midbit transition in the previous bit period [12]. Specifically, a binary **1** produces a reversal of midbit transition relative to the previous midbit transition. A binary **0** produces no reversal of the midbit transition. Certainly this is a transition code with memory. An example of a split phase (mark) coded waveform is shown in Fig. 6.1(g), where the waveform in the first bit period is chosen arbitrarily. Since this method encodes bits differentially, there is no 180°-phase ambiguity associated with some line codes. This phase ambiguity may not be an issue in most baseband links but is important if the line code is modulated. Split phase (space) is very similar to split phase (mark), where the role of the binary **1** and binary **0** are interchanged. An example of a split phase (space) coded waveform is given in Fig. 6.1(h); again, the first bit waveform is arbitrary.

6.3.3 Biphase (Mark)

This code, designated as Bi ϕ-M, is similar to a Miller code in that a binary **1** is represented by a midbit transition, and a binary **0** has no midbit transition. However, this code always has a transition at the beginning of a bit period [10]. Thus, the code is easy to synchronize and has no dc. An example of Bi ϕ-M is given in Fig. 6.1(i), where the direction of the transition at $t = 0$ is arbitrarily chosen. Biphase (space) or Bi ϕ-S is similar to Bi ϕ-M, except the role of the binary data is reversed. Here a binary **0** (space) produces a midbit transition, and a binary **1** does not have a midbit transition. A waveform example of Bi ϕ-S is shown in Fig. 6.1(j). Both Bi ϕ-S and Bi ϕ-M are transition codes with memory.

6.3.4 Code Mark Inversion (CMI)

This line code is used as the interface to a Consultative Committee on International Telegraphy and Telephony (CCITT) multiplexer and is very similar to Bi ϕ-S. A binary **1** is encoded as an NRZ pulse with alternate polarity, $+V$ or $-V$. A binary **0** is encoded with a definitive midbit transition (or square wave phase) [1]. An example of this waveform is shown in Fig. 6.1(k) where a negative to positive transition (or 180° phase) is used for a binary **0.** The voltage level of the first binary **1** in this example is chosen arbitrarily. This example waveform is identical to Bi ϕ-S shown in Fig. 6.1(j), except for the last bit. CMI has good synchronization properties and has no dc.

6.3.5 NRZ (I)

This type of line code uses an inversion (I) to designate binary digits, specifically, a change in level or no change in level. There are two variants of this code, NRZ mark (M) and NRZ space (S) [5, 12]. In NRZ (M), a change of level is used to indicate a binary **1,** and no change of level is used to indicate a binary **0.** In NRZ (S) a change of level is used to indicate a binary **0,** and no change of level is used to indicate a binary **1.** Waveforms for NRZ (M) and NRZ (S) are depicted in Fig. 6.1(l) and Fig. 6.1(m), respectively, where the voltage level of the first binary **1** in the example is chosen arbitrarily. These codes are level codes with memory. In general, line codes that use differential encoding, like NRZ (I), are insensitive to 180° phase ambiguity. Clock recovery with NRZ (I) is not particularly good, and dc wander is a problem as well. Its bandwidth is comparable to polar NRZ.

6.3.6 Binary N Zero Substitution (BNZS)

The common bipolar code AMI has many desirable properties of a line code. Its major limitation, however, is that a long string of zeros can lead to loss of synchronization and

Ford Motor Co.
Exhibit 1030
Page 201

timing jitter because there are no pulses in the waveform for relatively long periods of time. **Binary N zero substitution (BNZS)** attempts to improve AMI by substituting a special code of length N for all strings of N zeros. This special code contains pulses that look like binary **1**s but purposely produce violations of the AMI pulse convention. Two consecutive pulses of the same polarity violate the AMI pulse convention, independent of the number of zeros between the two consecutive pulses. These violations can be detected at the receiver, and the special code replaced by N zeros. The special code contains pulses facilitating synchronization even when the original data has long string of zeros. The special code is chosen such that the desirable properties of AMI coding are retained despite the AMI pulse convention violations, i.e., dc balance and error detection capability. The only disadvantage of BNZS compared to AMI is a slight increase in crosstalk due to the increased number of pulses and, hence, an increase in the average energy in the code.

Choosing different values of N yields different BNZS codes. The value of N is chosen to meet the timing requirements of the application. In telephony, there are three commonly used BNZS codes: B6ZS, B3ZS, and B8ZS. All BNZS codes are level codes with memory.

In a B6ZS code, a string of six consecutive zeros is replaced by one of two the special codes according to the rule:

If the last pulse was positive (+), the special code is: 0 + − 0 − + .

If the last pulse was negative (−), the special code is: 0 − + 0 + − .

Here a zero indicates a zero voltage level for the bit period; a plus designates a positive pulse; and a minus indicates a negative pulse.

This special code causes two AMI pulse violations: in its second bit position and in its fifth bit position. These violations are easily detected at the receiver and zeros resubstituted. If the number of consecutive zeros is $12, 18, 24, \ldots$, the substitution is repeated $2, 3, 4, \ldots$ times. Since the number of violations is even, the B6ZS waveform is the same as the AMI waveform outside the special code, i.e., between special code sequences.

There are four pulses introduced by the special code that facilitates timing recovery. Also, note that the special code is dc balanced. An example of the B6ZS code is given as follows, where the special code is indicated by the bold characters.

Original data:	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	1
B6ZS format:	0	+	**0**	+	−	**0**	−	+	−	+	0	−	**0**	−	+	**0**	+	−	+	−

The computation of the PSD of a B6ZS code is tedious. Its shape is given in Fig. 6.4, for comparison purposes with AMI, for the case of equally likely data.

In a B3ZS code, a string of three consecutive zeros is replaced by either $B\mathbf{0}V$ or $\mathbf{00V}$, where B denotes a pulse obeying the AMI (bipolar) convention and V denotes a pulse violating the AMI convention. $B\mathbf{0}V$ or $\mathbf{00}V$ is chosen such that the number of bipolar (B) pulses between the violations is odd. The B3ZS rules are summarized in Table 6.1.

TABLE 6.1 B3ZS Substitution Rules

Number of B Pulses Since Last Violation	Polarity of Last B Pulse	Substitution Code	Substitution Code Form
Odd	Negative (−)	**00−**	$\mathbf{00}V$
Odd	Positive (+)	**00+**	$\mathbf{00}V$
Even	Negative (−)	**+0+**	$B\mathbf{0}V$
Even	Positive (+)	**−0−**	$B\mathbf{0}V$

Ford Motor Co.
Exhibit 1030
Page 202

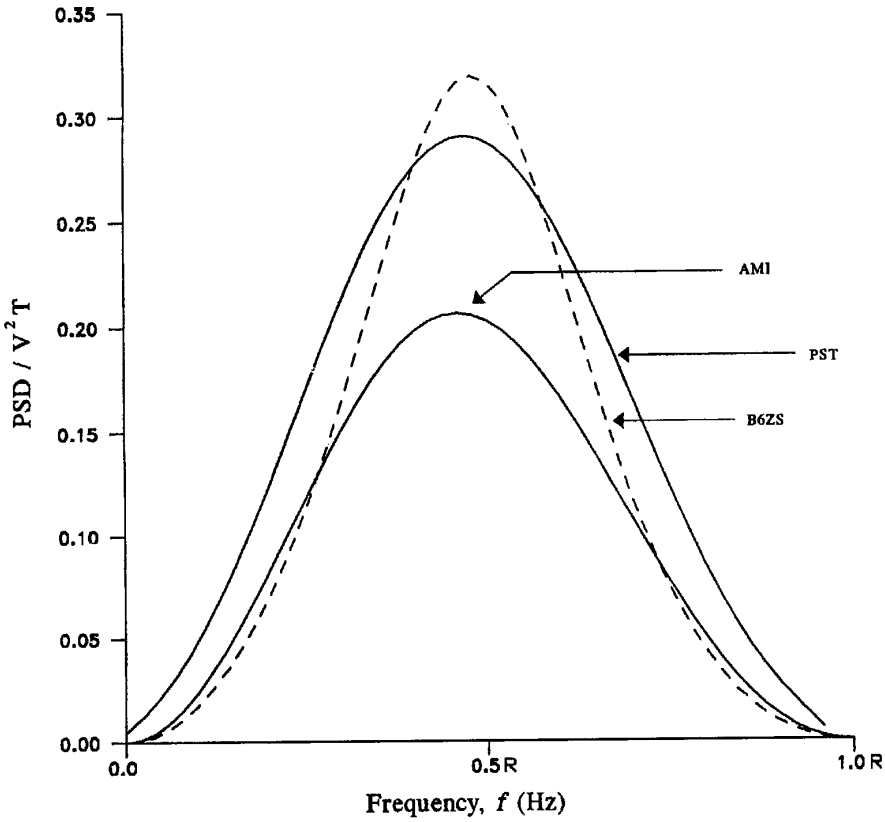


FIGURE 6.4 Power spectral density of different line codes, where $R = 1/\mathrm{T}$ is the bit rate.

Observe that the violation always occurs in the third bit position of the substitution code, and so it can be easily detected and zero replacement made at the receiver. Also, the substitution code selection maintains dc balance. There is either one or two pulses in the substitution code, facilitating synchronization. The error detection capability of AMI is retained in B3ZS because a single channel error would make the number of bipolar pulses between violations even instead of being odd. Unlike B6ZS, the B3ZS waveform between violations may not be the same as the AMI waveform. B3ZS is used in the digital signal-3 (DS-3) signal interface in North America and also in the long distance-4 (LD-4) coaxial transmission system in Canada. Next is an example of a B3ZS code, using the same symbol meaning as in the B6ZS code.

Original data:	1	0	0	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1
B3ZS format:																		
Even No. of B pulses:	+	0	0	−	+	**0**	+	−	+	−	**0**	−	0	+	**0**	**0**	+	−
Odd No. of B pulses:	+	0	0	−	**0**	**0**	−	+	−	+	**0**	+	0	−	**0**	**0**	−	+

The last BNZS code considered here uses $N = 8$. A B8ZS code is used to provide transparent channels for the Integrated Services Digital Network (ISDN) on T1 lines and is similar to the B6ZS code. Here a string of eight consecutive zeros is replaced by one of two

special codes according to the following rule:

If the last pulse was positive (+), the special code is: 0 0 0 + − 0 − +.
If the last pulse was negative (−), the special code is: 0 0 0 − + 0 + −.

There are two bipolar violations in the special codes, at the fourth and seventh bit positions. The code is dc balanced, and the error detection capability of AMI is retained. The waveform between substitutions is the same as that of AMI. If the number of consecutive zeros is $16, 24, \ldots$, then the substitution is repeated $2, 3, \ldots$, times.

6.3.7 High-Density Bipolar N (HDBN)

This coding algorithm is a CCITT standard recommended by the Conference of European Posts and Telecommunications Administrations (CEPT), a European standards body. It is quite similar to BNZS coding. It is thus a level code with memory. Whenever there is a string of $N + 1$ consecutive zeros, they are replaced by a special code of length $N + 1$ containing AMI violations. Specific codes can be constructed for different values of N. A specific **high-density bipolar N (HDBN)** code, HDB3, is implemented as a CEPT primary digital signal. It is very similar to the B3ZS code. In this code, a string of four consecutive zeros is replaced by either $B\mathbf{00}V$ or $\mathbf{000}V$. $B\mathbf{00}V$ or $\mathbf{000}V$ is chosen such that the number of bipolar (B) pulses between violations is odd. The HDB3 rules are summarized in Table 6.2.

TABLE 6.2 HDB3 Substitution Rules

Number of B Pulses Since Last Violation	Polarity of Last B Pulse	Substitution Code	Substitution Code Form
Odd	Negative (−)	**000−**	$\mathbf{000}V$
Odd	Positive (+)	**000+**	$\mathbf{000}V$
Even	Negative (−)	**+00+**	$B\mathbf{00}V$
Even	Positive (+)	**−00−**	$B\mathbf{00}V$

Here the violation always occurs in the fourth bit position of the substitution code, so that it can be easily detected and zero replacement made at the receiver. Also, the substitution code selection maintains dc balance. There is either one or two pulses in the substitution code facilitating synchronization. The error detection capability of AMI is retained in HDB3 because a single channel error would make the number of bipolar pulses between violations even instead of being odd.

6.3.8 Ternary Coding

Many line coding schemes employ three symbols or levels to represent only one bit of information, like AMI. Theoretically, it should be possible to transmit information more efficiently with three symbols, specifically the maximum efficiency is $\log_2 3 = 1.58$ bits per symbol. Alternatively, the redundancy in the code signal space can be used to provide better error control. Two examples of ternary coding are described next [1, 2]: **pair selected ternary (PST)** and **4 binary 3 ternary (4B3T).** The PST code has many of the desirable properties of line codes, but its transmission efficiency is still 1 bit per symbol. The 4B3T code also has many of the desirable properties of line codes, and it has increased transmission efficiency.

Ford Motor Co.
Exhibit 1030
Page 204

In the PST code, two consecutive bits, termed a binary pair, are grouped together to form a word. These binary pairs are assigned codewords consisting of two ternary symbols, where each ternary symbol can be +, −, or 0, just as in AMI. There are nine possible ternary codewords. Ternary codewords with identical elements, however, are avoided, i.e., ++, −−, and 00. The remaining six codewords are transmitted using two modes called + mode and − mode. The modes are switched whenever a codeword with a single pulse is transmitted. The PST code and mode switching rules are summarized in Table 6.3.

TABLE 6.3 PST Codeword Assignment and Mode Switching Rules

	Ternary Codewords		Mode
Binary Pair	+ Mode	− Mode	Switching
11	+ −	+ −	No
10	+ 0	− 0	Yes
01	0 +	0 −	Yes
00	− +	− +	No

PST is designed to maintain dc balance and include a strong timing component. One drawback of this code is that the bits must be framed into pairs. At the receiver, an *out-of-frame* condition is signalled when unused ternary codewords (++, −−, and 00) are detected. The mode switching property of PST provides error detection capability. PST can be classified as a level code with memory.

If the original data for PST coding contains only **1**s or **0**s, an alternating sequence of +− +− ⋯ is transmitted. As a result, an out-of-frame condition can not be detected. This problem can be minimized by using the modified PST code as shown in Table 6.4.

TABLE 6.4 Modified PST Codeword Assignment and Mode Switching Rules

	Ternary Codewords		Mode
Binary Pair	+ Mode	− Mode	Switching
11	+ 0	0 −	Yes
10	+ −	+ −	No
01	− +	− +	No
00	0 +	− 0	Yes

It is tedious to derive the PSD of a PST coded waveform. Again, Fig. 6.4 shows the PSD of the PST code along with the PSD of AMI and B6ZS for comparison purposes, all for equally likely binary data. Observe that PST has more power than AMI and, thus, a larger amount of energy per bit, which translates into slightly increased crosstalk.

In 4B3T coding, words consisting of four binary digits are mapped into three ternary symbols. Four bits imply $2^4 = 16$ possible binary words, whereas three ternary symbols allow $3^3 = 27$ possible ternary codewords. The binary-to-ternary conversion in 4B3T insures dc balance and a strong timing component. The specific codeword assignment is as shown in Table 6.5.

Ford Motor Co.
Exhibit 1030
Page 205

TABLE 6.5 4B3T Codeword Assignment

	Ternary Codewords		
Binary Words	Column 1	Column 2	Column 3
0000	− − −		+ + +
0001	− − 0		+ + 0
0010	− 0 −		+ 0 +
0011	0 − −		0 + +
0100	− − +		+ + −
0101	− + −		+ − +
0110	+ − −		− + +
0111	− 0 0		+ 0 0
1000	0 − 0		0 + 0
1001	0 0 −		0 0 +
1010		0 + −	
1011		0 − +	
1100		+ 0 −	
1101		− 0 +	
1110		+ − 0	
1111		− + 0	

There are three types of codewords in Table 6.5, organized into three columns. The codewords in the first column have negative dc, codewords in the second column have zero dc, and those in the third column have positive dc. The encoder monitors the integer variable

$$I = N_p - N_n \ , \tag{6.16}$$

where N_p is the number of positive pulses transmitted and N_n are the number of negative pulses transmitted. Codewords are chosen according to following rule:

If $I < 0$, choose the ternary codeword from columns 1 and 2.

If $I > 0$, choose the ternary codeword from columns 2 and 3.

If $I = 0$, choose the ternary word from column 2, and from column 1 if the previous $I > 0$ or from column 3 if the previous $I < 0$.

Note that the ternary codeword 000 is not used, but the remaining 26 codewords are used in a complementary manner. For example, the column 1 codeword for 0001 is $--0$, whereas the column 3 codeword is $++0$. The maximum transmission efficiency for the 4B3T code is 1.33 bits per symbol compared to 1 bit per symbol for the other line codes. The disadvantages of 4B3T are that framing is required and that performance monitoring is complicated. The 4B3T code is used in the T148 span line developed by ITT Telecommunications. This code allows transmission of 48 channels using only 50% more bandwidth than required by T1 lines, instead of 100% more bandwidth.

6.4 Multilevel Signalling, Partial Response Signalling, and Duobinary Coding

Ternary coding, such as 4B3T, is an example of the use of more than two levels to improve the transmission efficiency. To increase the transmission efficiency further, more

Ford Motor Co.
Exhibit 1030
Page 206

levels and/or more signal processing is needed. Multilevel signalling allows an improvement in the transmission efficiency at the expense of an increase in the error rate, i.e., more transmitter power will be required to maintain a given probability of error. In partial response signalling, intersymbol interference is deliberately introduced by using pulses that are wider and, hence, require less bandwidth. The controlled amount of interference from each pulse can be removed at the receiver. This improves the transmission efficiency, at the expense of increased complexity. **Duobinary coding,** a special case of partial response signalling, requires only the minimum theoretical bandwidth of $0.5R$ Hz. In what follows these techniques are discussed in slightly more detail.

6.4.1 Multilevel Signalling

The number of levels that can be used for a line code is not restricted to two or three. Since more levels or symbols allow higher transmission efficiency, multilevel signalling can be considered in bandwidth-limited applications. Specifically, if the signalling rate or baud rate is R_s and the number of levels used is L, the equivalent transmission bit rate R_b is given by

$$R_b = R_s \log_2[L] \ . \tag{6.17}$$

Alternatively, multilevel signalling can be used to reduce the baud rate, which in turn can reduce crosstalk for the same equivalent bit rate. The penalty, however, is that the SNR must increase to achieve the same error rate. The T1G carrier system of AT&T uses multilevel signalling with $L = 4$ and a baud rate of 3.152 mega-symbols/s to double the capacity of the T1C system from 48 channels to 96 channels. Also, a four level signalling scheme at 80-kB is used to achieve 160 kb/s as a basic rate in a digital subscriber loop (DSL) for ISDN.

6.4.2 Partial Response Signalling and Duobinary Coding

This class of signalling is also called *correlative* coding because it purposely introduces a controlled or correlated amount of intersymbol interference in each symbol. At the receiver, the known amount of interference is effectively removed from each symbol. The advantage of this signalling is that wider pulses can be used requiring less bandwidth, but the SNR must be increased to realize a given error rate. Also, errors can propagate unless *precoding* is used.

There are many commonly used partial response signalling schemes, often described in terms of the delay operator D, which represents one signalling interval delay. For example, in $(1 + D)$ signalling the current pulse and the previous pulse are added. The T1D system of AT&T uses $(1 + D)$ signalling with precoding, referred to as duobinary signalling, to convert binary (two level) data into ternary (three level) data at the same rate. This requires the minimum theoretical channel bandwidth without the deleterious effects of intersymbol interference and avoids error propagation. Complete details regarding duobinary coding are found in Lender, 1963 and Schwartz, 1980. Some partial response signalling schemes, such as $(1 - D)$, are used to shape the bandwidth rather than control it. Another interesting example of duobinary coding is a $(1 - D^2)$, which can be analyzed as the product $(1 - D)$ $(1 + D)$. It is used by GTE in its modified T carrier system. AT&T also uses $(1 - D^2)$ with four input levels to achieve an equivalent data rate of 1.544 Mb/s in only a 0.5-MHz bandwidth.

Ford Motor Co.
Exhibit 1030
Page 207

6.5 Bandwidth Comparison

We have provided the PSD expressions for most of the commonly used line codes. The actual bandwidth requirement, however, depends on the pulse shape used and the definition of bandwidth itself. There are many ways to define bandwidth, for example, as a percentage of the total power or the sidelobe suppression relative to the main lobe. Using the first null of the PSD of the code as the definition of bandwidth, Table 6.6 provides a useful bandwidth comparison.

TABLE 6.6 First Null Bandwidth Comparison

Bandwidth	Codes	
	Unipolar NRZ	BNZS
R	Polar NRZ	HDBN
	Polar RZ (AMI)	PST
$2R$	Unipolar RZ	Split Phase
	Manchester	CMI

The notable omission in Table 6.6 is delay modulation (Miller code). It does not have a first null in the $2R$-Hz band, but most of its power is contained in less than $0.5R$ Hz.

6.6 Concluding Remarks

An in-depth presentation of line coding, particularly applicable to telephony, has been included in this chapter. The most desirable characteristics of line codes were discussed. We introduced five common line codes and eight alternate line codes. Each line code was illustrated by an example waveform. In most cases expressions for the PSD and the probability of error were given and plotted. Advantages and disadvantages of all codes were included in the discussion, and some specific applications were noted. Line codes for optical fiber channels and networks built around them, such as fiber distributed data interface (FDDI) were not included in this section. A discussion of line codes for optical fiber channels, and other new developments in this topic area can be found in [1, 3, 4].

Defining Terms

Alternate mark inversion (AMI): A popular name for bipolar line coding using three levels: zero, positive, and negative.

Binary N zero substitution (BNZS): A class of coding schemes that attempts to improve AMI line coding.

Bipolar: A particular line coding scheme using three levels: zero, positive, and negative.

Crosstalk: An unwanted signal from an adjacent channel.

DC wander: The dc level variation in the received signal due to a channel that cannot support dc.

Duobinary coding: A coding scheme with binary input and ternary output requiring the minimum theoretical channel bandwidth.

Ford Motor Co.
Exhibit 1030
Page 208

4 Binary 3 Ternary (4B3T): A line coding scheme that maps four binary digits into three ternary symbols.

High-density bipolar *N* (HDBN): A class of coding schemes that attempts to improve AMI.

Level codes: Line codes carrying information in their voltage levels.

Line coding: The process of converting abstract symbols into real, temporal waveforms to be transmitted through a baseband channel.

Nonreturn to zero (NRZ): A signal that stays at a nonzero level for the entire bit duration.

Pair selected ternary (PST): A coding scheme based on selecting a pair of three level symbols.

Polar: A line coding scheme using both polarity of voltages, with or without a zero level.

Return to zero (RZ): A signal that returns to zero for a portion of the bit duration.

Transition codes: Line codes carrying information in voltage level transitions.

Unipolar: A line coding scheme using only one polarity of voltage, in addition to a zero level.

References

[1] Bellamy, J., *Digital Telephony,* John Wiley & Sons, New York, NY, 1991.

[2] Bell Telephone Laboratories Technical Staff Members. *Transmission Systems for Communications,* 4th ed., Western Electric Company, Technical Publications, Winston-Salem, NC, 1970.

[3] Bic, J.C., Duponteil, D., and Imbeaux, J.C., *Elements of Digital Communication,* John Wiley & Sons, New York, NY, 1991.

[4] Bylanski, P., *Digital Transmission Systems,* Peter Peregrinus, Herts, England, 1976.

[5] Couch, L.W., *Modern Communication Systems: Principles and Applications,* Prentice-Hall, Englewood Cliffs, NJ, 1994.

[6] Feher, K., *Digital Modulation Techniques in an Interference Environment,* EMC Encyclopedia Series, Vol. IX. Don White Consultants, Germantown, MD, 1977.

[7] Gibson, J.D., *Principles of Analog and Digital Communications,* MacMillan Publishing, New York, NY, 1993.

[8] Lathi, B.P., *Modern Digital and Analog Communication Systems,* Holt, Rinehart and Winston, Philadelphia, PA, 1989.

[9] Lender, A., Duobinary Techniques for High Speed Data Transmission, *IEEE Trans. Commun. Electron.,* CE-82, 214–218, May 1963.

[10] Lindsey, W.C. and Simon, M.K., *Telecommunication Systems Engineering,* Prentice-Hall, Englewood Cliffs, NJ, 1973.

[11] Schwartz, M., *Information Transmission, Modulation, and Noise,* McGraw-Hill, New York, NY, 1980.

[12] Stremler, F.G., *Introduction to Communication Systems,* Addison-Wesley Publishing, Reading, MA, 1990.

Ford Motor Co.
Exhibit 1030
Page 209

7

Echo Cancellation

7.1 Introduction 7-1
7.2 Echo Cancellation for PAM Systems 7-2
7.3 Echo Cancellation for QAM Systems 7-9
7.4 Echo Cancellation for OFDM Systems 7-11
7.5 Summary and Conclusions 7-14
References 7-14
Further Information 7-15

Giovanni Cherubini
IBM Zurich Research Laboratory

7.1 Introduction

Full-duplex data transmission over a single twisted-pair cable permits the simultaneous flow of information in two directions when the same frequency band is used. Examples of applications of this technique are found in digital communications systems that operate over the telephone network. In a digital subscriber loop, at each end of the full-duplex link, a circuit known as a hybrid separates the two directions of transmission. To avoid signal reflections at the near- and far-end hybrid, a precise knowledge of the line impedance would be required. Since the line impedance depends on line parameters that, in general, are not exactly known, an attenuated and distorted replica of the transmit signal leaks to the receiver input as an echo signal. Data-driven adaptive echo cancellation mitigates the effects of impedance mismatch.

A similar problem is caused by crosstalk in transmission systems over voice-grade unshielded twisted-pair cables for local-area network applications, where multipair cables are used to physically separate the two directions of transmission. Crosstalk is a statistical phenomenon due to randomly varying differential capacitive and inductive coupling between adjacent two-wire transmission lines. At the rates of several megabits per second that are usually considered for local-area network applications, near-end crosstalk (NEXT) represents the dominant disturbance; hence adaptive NEXT cancellation must be performed to ensure reliable communications.

In voiceband data modems, the model for the echo channel is considerably different from the echo model adopted in baseband transmission. The transmitted signal is a passband signal obtained by quadrature amplitude modulation (QAM), and the far-end echo may exhibit significant carrier-phase jitter and carrier-frequency shift, which are caused by signal processing at intermediate points in the telephone network. Therefore, a digital adaptive echo canceller for voiceband modems needs to embody algorithms that account for the presence of such additional impairments.

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 210

In this chapter, we describe the echo channel models and adaptive echo canceller structures that are obtained for various digital communications systems, which are classified according to the employed modulation techniques. We also address the tradeoffs between complexity, speed of adaptation, and accuracy of cancellation in adaptive echo cancellers.

7.2 Echo Cancellation for Pulse–Amplitude Modulation (PAM) Systems

The model of a full-duplex baseband data transmission system employing pulse–amplitude modulation (PAM) and adaptive echo cancellation is shown in Fig. 7.1. To describe system operations, we consider one end of the full-duplex link. The configuration of an echo canceller for a PAM transmission system is shown in Fig. 7.2. The transmitted data consist of a sequence $\{a_n\}$ of independent and identically distributed (i.i.d.) real-valued symbols from the M-ary alphabet $\mathcal{A} = \{\pm 1, \pm 3, \ldots, \pm(M-1)\}$. The sequence $\{a_n\}$ is converted into an analog signal by a digital-to-analog (D/A) converter. The conversion to a staircase signal by a zero-order hold D/A converter is described by the frequency response $H_{\mathrm{D/A}}(f) = T\sin(\pi fT)/(\pi fT)$, where T is the modulation interval. The D/A converter output is filtered by the analog transmit filter and is input to the channel through the hybrid.

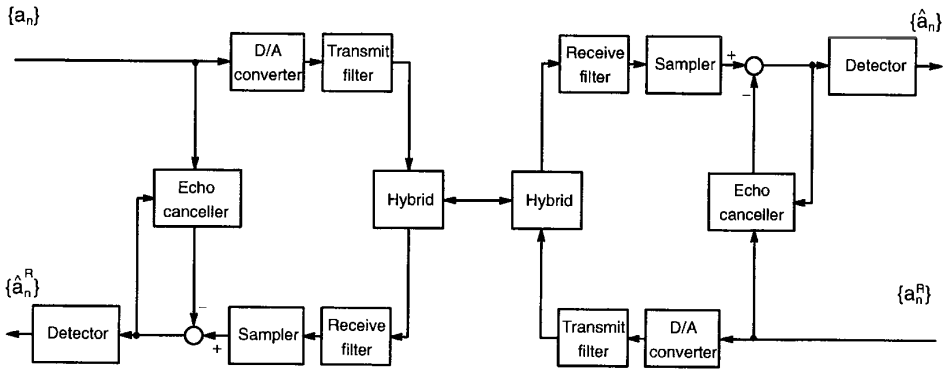


FIGURE 7.1 Model of a full-duplex PAM transmission system.

The signal $x(t)$ at the output of the low-pass analog receive filter has three components, namely, the signal from the far-end transmitter $r(t)$, the echo $u(t)$, and additive Gaussian noise $w(t)$. The signal $x(t)$ is given by

$$\begin{aligned} x(t) &= r(t) + u(t) + w(t) \\ &= \sum_{n=-\infty}^{\infty} a_n^R h(t-nT) + \sum_{n=-\infty}^{\infty} a_n h_E(t-nT) + w(t) \ , \end{aligned} \tag{7.1}$$

where $\{a_n^R\}$ is the sequence of symbols from the remote transmitter, and $h(t)$ and $h_E(t) = \{h_{\mathrm{D/A}} \otimes g_E\}(t)$ are the impulse responses of the overall channel and the echo channel, respectively. In the expression of $h_E(t)$, the function $h_{\mathrm{D/A}}(t)$ is the inverse Fourier transform of $H_{\mathrm{D/A}}(f)$, and the operator $\otimes$ denotes convolution. The signal obtained after echo cancellation is processed by a detector that outputs the sequence of estimated symbols $\{\hat{a}_n^R\}$. In

Ford Motor Co.
Exhibit 1030
Page 211

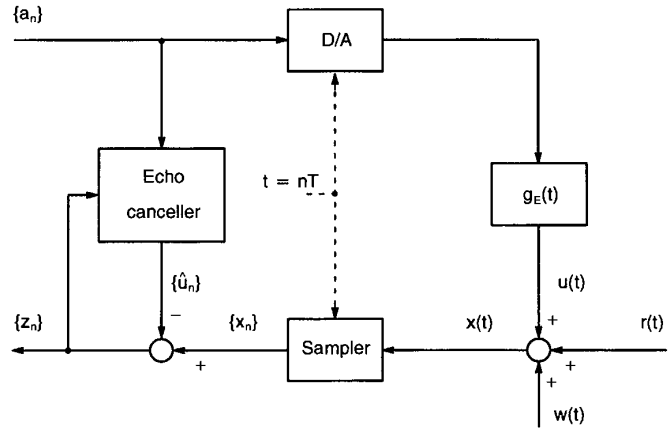


FIGURE 7.2 Configuration of an echo canceller for a PAM transmission system.

the case of full-duplex PAM data transmission over multi-pair cables for local-area network applications, where NEXT represents the main disturbance, the configuration of a digital NEXT canceller is obtained from Fig. 7.2, with the echo channel replaced by the crosstalk channel. For these applications, however, instead of *mono-duplex* transmission, where one pair is used to transmit only in one direction and the other pair to transmit only in the reverse direction, *dual-duplex* transmission may be adopted. Bi-directional transmission at rate ϱ over two pairs is then accomplished by full-duplex transmission of data streams at rate $\varrho/2$ over each of the two pairs. The lower modulation rate and/or spectral efficiency required per pair for achieving an aggregate rate equal to ϱ represents an advantage of dual-duplex over mono-duplex transmission. Dual-duplex transmission requires two transmitters and two receivers at each end of a link, as well as separation of the simultaneously transmitted and received signals on each pair, as illustrated in Fig. 7.3. In dual-duplex transceivers it is therefore necessary to suppress echoes returning from the hybrids and impedance discontinuities in the cable, as well as self NEXT, by adaptive digital echo and NEXT cancellation [3]. Although a dual-duplex scheme might appear to require higher implementation complexity than a mono-duplex scheme, it turns out that the two schemes are equivalent in terms of the number of multiply-and-add operations per second that are needed to perform the various filtering operations.

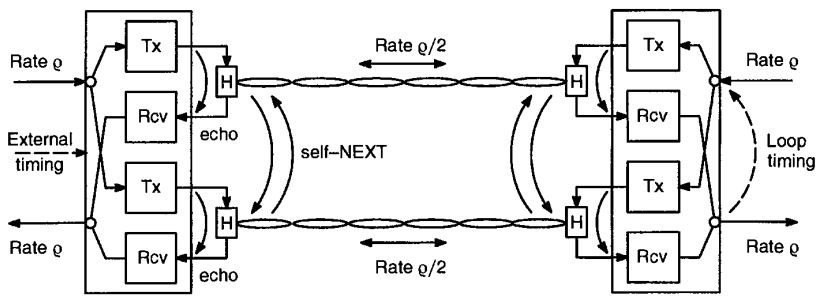


FIGURE 7.3 Model of a dual-duplex transmission system.

One of the transceivers in a full-duplex link will usually employ an externally provided reference clock for its transmit and receive operations. The other transceiver will extract timing from the received signal, and use this timing for its transmitter operations. This is known as *loop timing,* also illustrated in Fig. 7.3. If signals were transmitted in opposite directions with independent clocks, signals received from the remote transmitter would generally shift in phase relative to the also received echo signals. To cope with this effect, some form of interpolation would be required that can significantly increase the transceiver complexity [2].

In general, we consider baseband signalling techniques such that the signal at the output of the overall channel has nonnegligible excess bandwidth, i.e., nonnegligible spectral components at frequencies larger than half of the modulation rate, $|f| \geq 1/2T$. Therefore, to avoid aliasing, the signal $x(t)$ is sampled at twice the modulation rate or at a higher sampling rate. Assuming a sampling rate equal to $m/T, m > 1$, the ith sample during the nth modulation interval is given by

$$\begin{aligned} x\left[(nm+i)\frac{T}{m}\right] &= x_{nm+i} = r_{nm+i} + u_{nm+i} + w_{nm+i}, \qquad i = 0, \ldots, m-1 \\ &= \sum_{k=-\infty}^{\infty} h_{km+i} a^R_{n-k} + \sum_{k=-\infty}^{\infty} h_{E,km+i} a_{n-k} + w_{nm+i}\ , \end{aligned} \tag{7.2}$$

where $\{h_{nm+i},\ i = 0, \ldots, m-1\}$ and $\{h_{E,nm+i},\ i = 0, \ldots, m-1\}$ are the discrete-time impulse responses of the overall channel and the echo channel, respectively, and $\{w_{nm+i}, i = 0, \ldots, m-1\}$ is a sequence of Gaussian noise samples with zero mean and variance σ_w^2. Equation (7.2) suggests that the sequence of samples $\{x_{nm+i},\ i = 0, \ldots, m-1\}$ be regarded as a set of m interleaved sequences, each with a sampling rate equal to the modulation rate. Similarly, the sequence of echo samples $\{u_{nm+i},\ i = 0, \ldots, m-1\}$ can be regarded as a set of m interleaved sequences that are output by m independent echo channels with discrete-time impulse responses $\{h_{E,nm+i}\}$, $i = 0, \ldots, m-1$, and an identical sequence $\{a_n\}$ of input symbols [7]. Hence, echo cancellation can be performed by m interleaved echo cancellers, as shown in Fig. 7.4. Since the performance of each canceller is independent of the other $m-1$ units, in the remaining part of this section we will consider the operations of a single echo canceller.

The echo canceller generates an estimate $\hat{u}_n$ of the echo signal. If we consider a transversal filter realization, $\hat{u}_n$ is obtained as the inner product of the vector of filter coefficients at time $t = nT$, $\boldsymbol{c}_n = (c_{n,0}, \ldots, c_{n,N-1})'$ and the vector of signals stored in the echo canceller delay line at the same instant, $\boldsymbol{a}_n = (a_n, \ldots, a_{n-N+1})'$, expressed by

$$\hat{u}_n = \boldsymbol{c}_n' \boldsymbol{a}_n = \sum_{k=0}^{N-1} c_{n,k} a_{n-k} \tag{7.3}$$

where $\boldsymbol{c}_n'$ denotes the transpose of the vector $\boldsymbol{c}_n$. The estimate of the echo is subtracted from the received signal. The result is defined as the cancellation error signal

$$z_n = x_n - \hat{u}_n = x_n - \boldsymbol{c}_n' \boldsymbol{a}_n\ . \tag{7.4}$$

The echo attenuation that must be provided by the echo canceller to achieve proper system operation depends on the application. For example, for the Integrated Services Digital Network (ISDN) U-Interface transceiver, the echo attenuation must be larger than 55 dB [10]. It is then required that the echo signals outside of the time span of the echo canceller delay line be negligible, i.e., $h_{E,n} \approx 0$ for $n < 0$ and $n > N-1$. As a measure

Ford Motor Co.
Exhibit 1030
Page 213

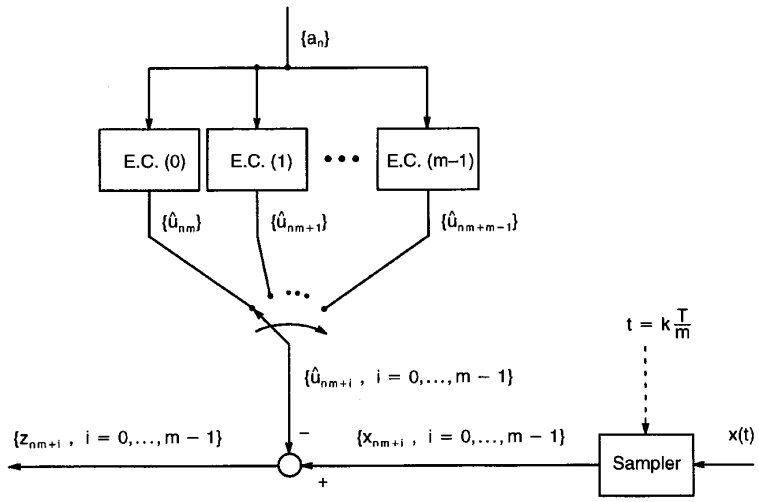


FIGURE 7.4 A set of m interleaved echo cancellers.

of system performance, we consider the mean square error ε_n^2 at the output of the echo canceller at time $t = nT$, defined by

$$\varepsilon_n^2 = E\left\{z_n^2\right\} \ , \tag{7.5}$$

where $\{z_n\}$ is the error sequence and $E\{\cdot\}$ denotes the expectation operator. For a particular coefficient vector $\boldsymbol{c}_n$, substitution of Eq. (7.4) into Eq. (7.5) yields

$$\varepsilon_n^2 = E\left\{x_n^2\right\} - 2\boldsymbol{c}_n'\boldsymbol{q} + \boldsymbol{c}_n'\boldsymbol{R}\boldsymbol{c}_n \ , \tag{7.6}$$

where $\boldsymbol{q} = E\{x_n\boldsymbol{a}_n\}$ and $\boldsymbol{R} = E\{\boldsymbol{a}_n\boldsymbol{a}_n'\}$. With the assumption of i.i.d. transmitted symbols, the correlation matrix $\boldsymbol{R}$ is diagonal. The elements on the diagonal are equal to the variance of the transmitted symbols, $\sigma_a^2 = (M^2 - 1)/3$. The minimum mean square error is given by

$$\varepsilon_{\min}^2 = E\left\{x_n^2\right\} - \boldsymbol{c}_{\text{opt}}'\boldsymbol{R}\boldsymbol{c}_{\text{opt}} \ , \tag{7.7}$$

where the optimum coefficient vector is $\boldsymbol{c}_{\text{opt}} = \boldsymbol{R}^{-1}\boldsymbol{q}$. We note that proper system operation is achieved only if the transmitted symbols are uncorrelated with the symbols from the remote transmitter. If this condition is satisfied, the optimum filter coefficients are given by the values of the discrete-time echo channel impulse response, i.e., $c_{\text{opt},k} = h_{E,k}, k = 0, \ldots, N-1$.

By the decision-directed stochastic gradient algorithm, also known as the least mean square (LMS) algorithm, the coefficients of the echo canceller converge in the mean to $\boldsymbol{c}_{\text{opt}}$. The LMS algorithm for an N-tap adaptive linear transversal filter is formulated as follows:

$$\boldsymbol{c}_{n+1} = \boldsymbol{c}_n - \frac{1}{2}\alpha\nabla_{\boldsymbol{c}}\left\{z_n^2\right\} = \boldsymbol{c}_n + \alpha z_n\boldsymbol{a}_n \ , \tag{7.8}$$

where α is the adaptation gain and

$$\nabla_{\boldsymbol{c}}\left\{z_n^2\right\} = \left(\frac{\partial z_n^2}{\partial c_{n,0}}, \ldots, \frac{\partial z_n^2}{\partial c_{n,N-1}}\right)' = -2z_n\boldsymbol{a}_n$$

Ford Motor Co.
Exhibit 1030
Page 214

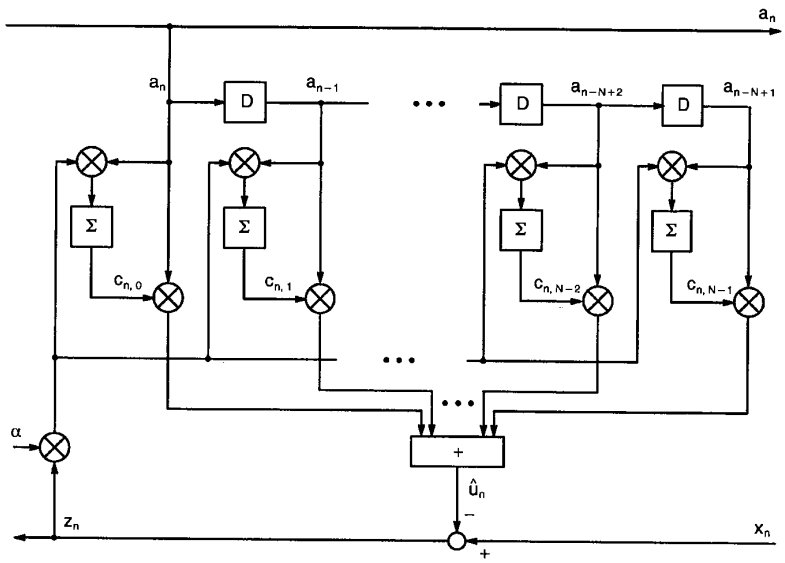


FIGURE 7.5 Block diagram of an adaptive transversal filter echo canceller.

is the gradient of the squared error with respect to the vector of coefficients. The block diagram of an adaptive transversal filter echo canceller is shown in Fig. 7.5.

If we define the vector $\boldsymbol{p}_n = \boldsymbol{c}_{\text{opt}} - \boldsymbol{c}_n$, the mean square error can be expressed as

$$\varepsilon_n^2 = \varepsilon_{\min}^2 + \boldsymbol{p}_n' \boldsymbol{R} \boldsymbol{p}_n \ , \tag{7.9}$$

where the term $\boldsymbol{p}_n' \boldsymbol{R} \boldsymbol{p}_n$ represents an 'excess mean square distortion' due to the misadjustment of the filter settings. The analysis of the convergence behavior of the excess mean square distortion was first proposed for adaptive equalizers [13] and later extended to adaptive echo cancellers [9]. Under the assumption that the vectors $\boldsymbol{p}_n$ and $\boldsymbol{a}_n$ are statistically independent, the dynamics of the mean square error are given by

$$E\left\{\varepsilon_n^2\right\} = \varepsilon_0^2 \left[1 - \alpha\sigma_a^2 \left(2 - \alpha N \sigma_a^2\right)\right]^n + \frac{2\varepsilon_{\min}^2}{2 - \alpha N \sigma_a^2} \ , \tag{7.10}$$

where ε_0^2 is determined by the initial conditions. The mean square error converges to a finite steady-state value ε_∞^2 if the stability condition $0 < \alpha < 2/(N\sigma_a^2)$ is satisfied. The optimum adaptation gain that yields fastest convergence at the beginning of the adaptation process is $\alpha_{\text{opt}} = 1/(N\sigma_a^2)$. The corresponding time constant and asymptotic mean square error are $\tau_{\text{opt}} = N$ and $\varepsilon_\infty^2 = 2\varepsilon_{\min}^2$, respectively.

We note that a fixed adaptation gain equal to α_{opt} could not be adopted in practice, since after echo cancellation the signal from the remote transmitter would be embedded in a residual echo having approximately the same power. If the time constant of the convergence mode is not a critical system parameter, an adaptation gain smaller than α_{opt} will be adopted to achieve an asymptotic mean square error close to $\varepsilon_{\min}^2$. On the other hand, if fast convergence is required, a variable gain will be chosen.

Several techniques have been proposed to increase the speed of convergence of the LMS algorithm. In particular, for echo cancellation in data transmission, the speed of adaptation is reduced by the presence of the signal from the remote transmitter in the cancellation error.

Ford Motor Co.
Exhibit 1030
Page 215

To mitigate this problem, the data signal can be adaptively removed from the cancellation error by a decision-directed algorithm [5].

Modified versions of the LMS algorithm have been also proposed to reduce system complexity. For example, the sign algorithm suggests that only the sign of the error signal be used to compute an approximation of the stochastic gradient [4]. An alternative means to reduce the implementation complexity of an adaptive echo canceller consists in the choice of a filter structure with a lower computational complexity than the transversal filter.

At high data rates, very large scale integration (VLSI) technology is needed for the implementation of transceivers for full-duplex data transmission. High-speed echo cancellers and near-end crosstalk cancellers that do not require multiplications represent an attractive solution because of their low complexity. As an example of an architecture suitable for VLSI implementation, we consider echo cancellation by a distributed-arithmetic filter, where multiplications are replaced by table lookup and shift-and-add operations [12]. By segmenting the echo canceller into filter sections of shorter lengths, various tradeoffs concerning the number of operations per modulation interval and the number of memory locations needed to store the lookup tables are possible. Adaptivity is achieved by updating the values stored in the lookup tables by the LMS algorithm.

To describe the principles of operations of a distributed-arithmetic echo canceller, we assume that the number of elements in the alphabet of input symbols is a power of two, $M = 2^W$. Therefore, each symbol is represented by the vector $(a_n^{(0)}, \ldots, a_n^{(W-1)})$, where $a_n^{(i)}, i = 0, \ldots, W-1$, are independent binary random variables, i.e.,

$$a_n = \sum_{w=0}^{W-1} \left(2a_n^{(w)} - 1\right) 2^w = \sum_{w=0}^{W-1} b_n^{(w)} 2^w \ , \tag{7.11}$$

where $b_n^{(w)} = (2a_n^{(w)} - 1) \in \{-1, +1\}$. By substituting Eq. (7.11) into Eq. (7.1) and segmenting the delay line of the echo canceller into L sections with $K = N/L$ delay elements each, we obtain

$$\hat{u}_n = \sum_{\ell=0}^{L-1} \sum_{w=0}^{W-1} 2^w \left[\sum_{k=0}^{K-1} b_{n-\ell K-k}^{(w)} c_{n,\ell K+k} \right] \ . \tag{7.12}$$

Equation (7.12) suggests that the filter output can be computed using a set of $L2^K$ values that are stored in L tables with 2^K memory locations each. The binary vectors $\boldsymbol{a}_{n,\ell}^{(w)} = (a_{n-(\ell+1)K+1}^{(w)}, \ldots, a_{n-\ell K}^{(w)})$, $w = 0, \ldots, W-1$, $\ell = 0, \ldots, L-1$, determine the addresses of the memory locations where the values that are needed to compute the filter output are stored. The filter output is obtained by WL table lookup and shift-and-add operations.

We observe that $\boldsymbol{a}_{n,\ell}^{(w)}$ and its binary complement $\bar{\boldsymbol{a}}_{n,\ell}^{(w)}$ select two values that differ only in their sign. This symmetry is exploited to halve the number of values to be stored. To determine the output of a distributed-arithmetic filter with reduced memory size, we reformulate Eq. (7.12) as

$$\hat{u}_n = \sum_{\ell=0}^{L-1} \sum_{w=0}^{W-1} 2^w b_{n-\ell K-k_0}^{(w)} \left[c_{n,\ell K+k_0} + b_{n-\ell K-k_0}^{(w)} \sum_{\substack{k=0 \\ k \neq k_0}}^{K-1} b_{n-\ell K-k}^{(w)} c_{n,\ell K+k} \right] \ , \tag{7.13}$$

where k_0 can be any element of the set $\{0, \ldots, K-1\}$. In the following, we take $k_0 = 0$. Then the binary symbols $b_{n-\ell K}^{(w)}$ determine whether the selected values are to be added or

Ford Motor Co.
Exhibit 1030
Page 216

subtracted. Each table has now 2^{K-1} memory locations, and the filter output is given by

$$\hat{u}_n = \sum_{\ell=0}^{L-1} \sum_{w=0}^{W-1} 2^w b_{n-\ell K}^{(w)} d_n \left(i_{n,\ell}^{(w)}, \ell \right) , \tag{7.14}$$

where $d_n(k,\ell)$, $k = 0, \ldots, 2^{K-1}-1$, $\ell = 0, \ldots, L-1$, are the look up values, and $i_{n,\ell}^{(w)}$, $w = 0, \ldots, W-1$, $\ell = 0, \ldots, L-1$, are the look up indices computed as follows:

$$i_{n,\ell}^{(w)} = \begin{cases} \displaystyle\sum_{k=1}^{K-1} a_{n-\ell K-k}^{(w)} 2^{k-1} & \text{if } a_{n-\ell K}^{(w)} = 1 \\ \displaystyle\sum_{k=1}^{K-1} \bar{a}_{n-\ell K-k}^{(w)} 2^{k-1} & \text{if } a_{n-\ell K}^{(w)} = 0 \end{cases} . \tag{7.15}$$

We note that, as long as Eqs. (7.12) and (7.13) hold for some coefficient vector $(c_{n,0}, \ldots, c_{n,N-1})$, the distributed-arithmetic filter emulates the operation of a linear transversal filter. For arbitrary values $d_n(k,\ell)$, however, a nonlinear filtering operation results.

The expression of the LMS algorithm to update the values of a distributed-arithmetic echo canceller takes the form

$$\boldsymbol{d}_{n+1} = \boldsymbol{d}_n - \frac{1}{2}\alpha \nabla_{\boldsymbol{d}} \left\{ z_n^2 \right\} = \boldsymbol{d}_n + \alpha z_n \boldsymbol{y}_n , \tag{7.16}$$

where $\boldsymbol{d}'_n = [\boldsymbol{d}'_n(0), \ldots, \boldsymbol{d}'_n(L-1)]$, with $\boldsymbol{d}'_n(\ell) = [d_n(0,\ell), \ldots, d_n(2^{K-1}-1,\ell)]$, and $\boldsymbol{y}'_n = [\boldsymbol{y}'_n(0), \ldots, \boldsymbol{y}'_n(L-1)]$, with

$$\boldsymbol{y}'_n(\ell) = \sum_{w=0}^{W-1} 2^w b_{n-\ell K}^{(w)} \left(\delta_{0,i_{n,\ell}^{(w)}}, \ldots, \delta_{2^{K-1}-1,i_{n,\ell}^{(w)}} \right) ,$$

are $L2^{K-1} \times 1$ vectors and where $\delta_{i,j}$ is the Kronecker delta. We note that at each iteration only those values that are selected to generate the filter output are updated. The block diagram of an adaptive distributed-arithmetic echo canceller with input symbols from a quaternary alphabet is shown in Fig. 7.6.

The analysis of the mean square error convergence behavior and steady-state performance has been extended to adaptive distributed-arithmetic echo cancellers [1]. The dynamics of the mean square error are given by

$$E\left\{\varepsilon_n^2\right\} = \varepsilon_0^2 \left[1 - \frac{\alpha\sigma_a^2}{2^{K-1}} \left(2 - \alpha L \sigma_a^2 \right) \right]^n + \frac{2\varepsilon_{\min}^2}{2 - \alpha L \sigma_a^2} . \tag{7.17}$$

The stability condition for the echo canceller is $0 < \alpha < 2/(L\sigma_a^2)$. For a given adaptation gain, echo canceller stability depends on the number of tables and on the variance of the transmitted symbols. Therefore, the time span of the echo canceller can be increased without affecting system stability, provided that the number L of tables is kept constant. In that case, however, mean square error convergence will be slower. From Eq. (7.17), we find that the optimum adaptation gain that permits the fastest mean square error convergence at the beginning of the adaptation process is $\alpha_{\text{opt}} = 1/(L\sigma_a^2)$. The time constant of the convergence mode is $\tau_{\text{opt}} = L2^{K-1}$. The smallest achievable time constant is proportional to the total number of values. The realization of a distributed-arithmetic echo canceller can be further simplified by updating at each iteration only the values that are addressed by the most significant bits of the symbols stored in the delay line. The complexity required for adaptation can thus be reduced at the price of a slower rate of convergence.

Ford Motor Co.
Exhibit 1030
Page 217

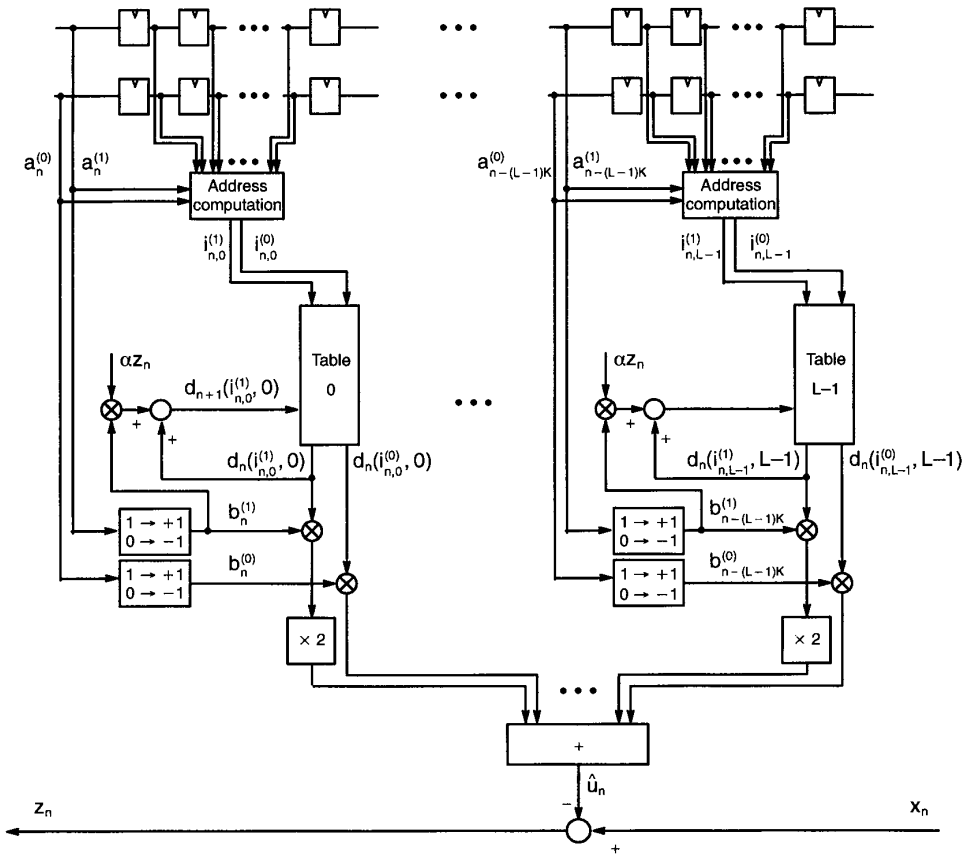


FIGURE 7.6 Block diagram of an adaptive distributed-arithmetic echo canceller.

7.3 Echo Cancellation for Quadrature Amplitude Modulation (QAM) Systems

Although most of the concepts presented in the preceding sections can be readily extended to echo cancellation for communications systems employing QAM, the case of full-duplex transmission over a voiceband data channel requires a specific discussion. We consider the system model shown in Fig. 7.7. The transmitter generates a sequence $\{a_n\}$ of i.i.d. complex-valued symbols from a two-dimensional constellation $\mathcal{A}$, which are modulated by the carrier $e^{j2\pi f_c nT}$, where T and f_c denote the modulation interval and the carrier frequency, respectively. The discrete-time signal at the output of the transmit Hilbert filter may be regarded as an analytic signal, which is generated at the rate of m/T samples/s, $m > 1$. The real part of the analytic signal is converted into an analog signal by a D/A converter and input to the channel. We note that by transmitting the real part of a complex-valued signal positive- and negative-frequency components become folded. The image band attenuation of the transmit Hilbert filter thus determines the achievable echo suppression. In fact, the receiver cannot extract aliasing image-band components from desired passband frequency components, and the echo canceller is able to suppress only echo arising from

transmitted passband components.

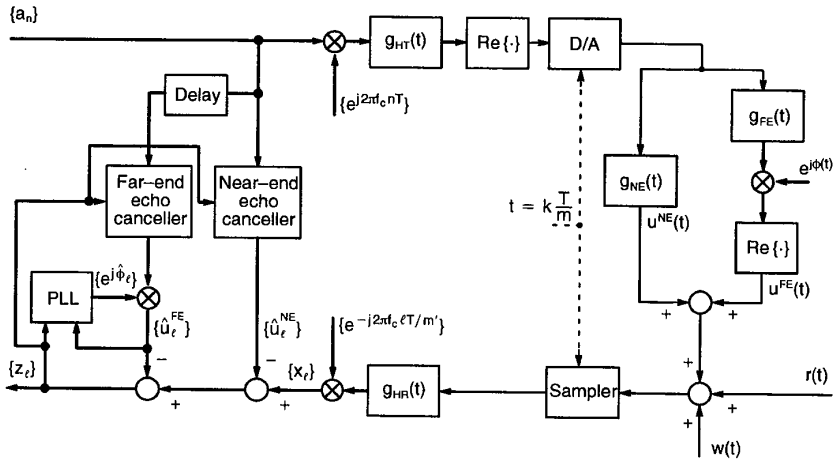


FIGURE 7.7 Configuration of an echo canceller for a QAM transmission system.

The output of the echo channel is represented as the sum of two contributions. The near-end echo $u^{\mathrm{NE}}(t)$ arises from the impedance mismatch between the hybrid and the transmission line, as in the case of baseband transmission. The far-end echo $u^{\mathrm{FE}}(t)$ represents the contribution due to echos that are generated at intermediate points in the telephone network. These echos are characterized by additional impairments, such as jitter and frequency shift, which are accounted for by introducing a carrier-phase rotation of an angle $\phi(t)$ in the model of the far-end echo.

At the receiver, samples of the signal at the channel output are obtained synchronously with the transmitter timing, at the sampling rate of m/T samples/s. The discrete-time received signal is converted to a complex-valued baseband signal $\{x_{nm'+i}, i = 0, \ldots, m'-1\}$, at the rate of m'/T samples/s, $1 < m' < m$, through filtering by the receive Hilbert filter, decimation, and demodulation. From delayed transmit symbols, estimates of the near- and far-end echo signals after demodulation, $\{\hat{u}^{\mathrm{NE}}_{nm'+i},\ i = 0, \ldots, m'-1\}$ and $\{\hat{u}^{\mathrm{FE}}_{nm'+i},\ i = 0, \ldots, m'-1\}$, respectively, are generated using m' interleaved near- and far-end echo cancellers. The cancellation error is given by

$$z_\ell = x_\ell - \hat{u}^{\mathrm{NE}}_\ell - \hat{u}^{\mathrm{FE}}_\ell \ . \tag{7.18}$$

A different model is obtained if echo cancellation is accomplished before demodulation. In this case, two equivalent configurations for the echo canceller may be considered. In one configuration, the modulated symbols are input to the transversal filter, which approximates the passband echo response. Alternatively, the modulator can be placed after the transversal filter, which is then called a baseband transversal filter [14].

In the considered realization, the estimates of the echo signals after demodulation are given by

$$\hat{u}^{\mathrm{NE}}_{nm'+i} = \sum_{k=0}^{N_{\mathrm{NE}}-1} c^{\mathrm{NE}}_{n,km'+i} a_{n-k}, \qquad i = 0, \ldots, m'-1 \ , \tag{7.19}$$

Ford Motor Co.
Exhibit 1030
Page 219

and

$$\hat{u}^{\mathrm{FE}}_{nm'+i} = \left[\sum_{k=0}^{N_{\mathrm{FE}}-1} c^{\mathrm{FE}}_{n,km'+i} a_{n-k-D_{\mathrm{FE}}} \right] e^{j\hat{\phi}_{nm'+i}}, \qquad i = 0, \ldots, m'-1 \ , \tag{7.20}$$

where $(c^{\mathrm{NE}}_{n,0}, \ldots, c^{\mathrm{NE}}_{n,m'N_{\mathrm{NE}}-1})$ and $(c^{\mathrm{FE}}_{n,0}, \ldots, c^{\mathrm{FE}}_{n,m'N_{\mathrm{FE}}-1})$ are the coefficients of the m' interleaved near- and far-end echo cancellers, respectively, $\{\hat{\phi}_{nm'+i}, i = 0, \ldots, m'-1\}$ is the sequence of far-end echo phase estimates, and D_{FE} denotes the bulk delay accounting for the round-trip delay from the transmitter to the point of echo generation. To prevent overlap of the time span of the near-end echo canceller with the time span of the far-end echo canceller, the condition $D_{\mathrm{FE}} > N_{\mathrm{NE}}$ must be satisfied. We also note that, because of the different nature of near- and far-end echo generation, the time span of the far-end echo canceller needs to be larger than the time span of the near-end echo canceller, i.e., $N_{\mathrm{FE}} > N_{\mathrm{NE}}$.

Adaptation of the filter coefficients in the near- and far-end echo cancellers by the LMS algorithm leads to

$$\begin{aligned} c^{\mathrm{NE}}_{n+1,km'+i} &= c^{\mathrm{NE}}_{n,km'+i} + \alpha z_{nm'+i}(a_{n-k})^* \\ k &= 0, \ldots, N_{\mathrm{NE}}-1, \qquad i = 0, \ldots, m'-1 \ , \end{aligned} \tag{7.21}$$

and

$$\begin{aligned} c^{\mathrm{FE}}_{n+1,km'+i} &= c^{\mathrm{FE}}_{n,km'+i} + \alpha z_{nm'+i}(a_{n-k-D_{\mathrm{FE}}})^* e^{-j\hat{\phi}_{nm'+i}} \\ k &= 0, \ldots, N_{\mathrm{FE}}-1, \qquad i = 0, \ldots, m'-1 \ , \end{aligned} \tag{7.22}$$

respectively, where the asterisk denotes complex conjugation.

The far-end echo phase estimate is computed by a second-order phase-lock loop algorithm, where the following stochastic gradient approach is adopted:

$$\begin{cases} \hat{\phi}_{\ell+1} = \hat{\phi}_\ell - \frac{1}{2}\gamma_{\mathrm{FE}} \nabla_{\hat{\phi}} |z_\ell|^2 + \Delta\phi_\ell \pmod{2\pi} \\ \Delta\phi_{\ell+1} = \Delta\phi_\ell - \frac{1}{2}\zeta_{\mathrm{FE}} \nabla_{\hat{\phi}} |z_\ell|^2 \end{cases} , \tag{7.23}$$

where $\ell = nm' + i$, $i = 0, \ldots, m'-1$, γ_{FE} and ζ_{FE} are step-size parameters, and

$$\nabla_{\hat{\phi}} |z_\ell|^2 = \frac{\partial |z_\ell|^2}{\partial \hat{\phi}_\ell} = -2\mathrm{Im}\left\{ z_\ell \left(\hat{u}^{\mathrm{FE}}_\ell\right)^* \right\} . \tag{7.24}$$

We note that algorithm (7.23) requires m' iterations per modulation interval, i.e., we cannot resort to interleaving to reduce the complexity of the computation of the far-end echo phase estimate.

7.4 Echo Cancellation for Orthogonal Frequency Division Multiplexing (OFDM) Systems

Orthogonal frequency division multiplexing (OFDM) is a modulation technique whereby blocks of M symbols are transmitted in parallel over M subchannels by employing M orthogonal subcarriers. We consider a real-valued discrete-time channel impulse response $\{h_i, i = 0, \ldots, L\}$ having length $L + 1 \ll M$. To illustrate the basic principles of OFDM systems, let us consider a noiseless ideal channel with impulse response given by $\{h_i\} = \{\delta_i\}$, where $\{\delta_i\}$ is defined as the discrete-time delta function. Modulation of the complex-valued

Ford Motor Co.
Exhibit 1030
Page 220

input symbols at the n-th modulation interval, denoted by the vector $\boldsymbol{A}_n = \{A_n(i), i = 0, \ldots, M-1\}$, is performed by an inverse discrete Fourier transform (IDFT), as shown in Fig. 7.8. We assume that M is even, and that each block of symbols satisfies the Hermitian symmetry conditions, i.e., $A_n(0)$ and $A_n(M/2)$ are real valued, and $A_n(i) = A_n^*(M-i)$, $i = 1, \ldots, M/2-1$. Then the signals $\boldsymbol{a}_n = \{a_n(i), i = 0, \ldots, M-1\}$ obtained at the output of the IDFT are real valued. After parallel-to-serial conversion, the M signals are sent over the channel at the given transmission rate M/T, where T denotes the modulation interval. At the output of the channel, the noiseless signals are received without distortion. Serial-to-parallel conversion yields blocks of M elements, with boundaries placed such that each block obtained at the modulator output is also presented at the demodulator input. Then demodulation performed by a discrete Fourier transform (DFT) will reproduce the blocks of M input symbols. The overall input-output relationship is therefore equivalent to that of a bank of M parallel, independent subchannels.

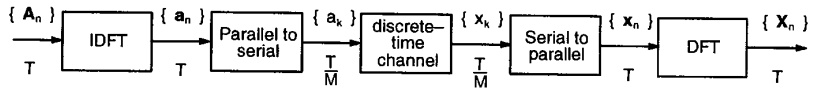


FIGURE 7.8 Block diagram of an OFDM system.

In the general case of a noisy channel with impulse response having length greater than one, M independent subchannels are obtained by a variant of OFDM that is also known as discrete multitone modulation (DMT) [11]. In a DMT system, modulation by the IDFT is performed at the rate $1/T' = M/(M+L)T < 1/T$. After modulation, each block of M signals is cyclically extended by copying the last L signals in front of the block, and converted from parallel to serial. The resulting $L+M$ signals are sent over the channel. At the receiver, blocks of samples with length $L+M$ are taken. Block boundaries are placed such that the last M samples depend only on the elements of one cyclically extended block of signals. The first L samples are discarded, and the vector $\boldsymbol{x}_n$ of the last M samples of the block received at the n-th modulation interval can be expressed as

$$\boldsymbol{x}_n = \Gamma_n \boldsymbol{h} + \boldsymbol{w}_n \ , \tag{7.25}$$

where $\boldsymbol{h}$ is the vector of the impulse response extended with $M-L-1$ zeros, $\boldsymbol{w}_n$ is a vector of additive white Gaussian noise samples, and Γ_n is a $M \times M$ circulant matrix given by

$$\Gamma_n = \begin{bmatrix} a_n(0) & a_n(M-1) & \ldots & a_n(1) \\ a_n(1) & a_n(0) & \ldots & a_n(2) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_n(M-1) & a_n(M-2) & \ldots & a_n(0) \end{bmatrix} . \tag{7.26}$$

Recalling that $\mathcal{F}_M \Gamma_n \mathcal{F}_M^{-1} = diag(\boldsymbol{A}_n)$, where $\mathcal{F}_M$ is the $M \times M$ DFT matrix defined as $\mathcal{F}_M = [(e^{-\frac{j2\pi}{M}})^{km}], k, m = 0, \ldots, M-1$, and $diag(\boldsymbol{A}_n)$ denotes the diagonal matrix with elements on the diagonal given by $\boldsymbol{A}_n$, we find that the output of the demodulator is given by

$$\boldsymbol{X}_n = diag(\boldsymbol{A}_n)\boldsymbol{H} + \boldsymbol{W}_n \ , \tag{7.27}$$

where $\boldsymbol{H}$ denotes the DFT of the vector $\boldsymbol{h}$, and $\boldsymbol{W}_n$ is a vector of independent Gaussian random variables. Equation (7.27) indicates that the sequence of transmitted symbol vectors

Ford Motor Co.
Exhibit 1030
Page 221

can be detected by assuming a bank of M independent subchannels, at the price of a decrease in the data rate by a factor $(M + L)/M$. Note that in practice the computationally more efficient inverse fast Fourier transform and fast Fourier transform are used instead of IDFT and DFT.

We discuss echo cancellation for OFDM with reference to a DMT system [6], as shown in Fig. 7.9. The real-valued discrete-time echo impulse response is $\{h_{E,i}, i = 0, \ldots, N-1\}$,

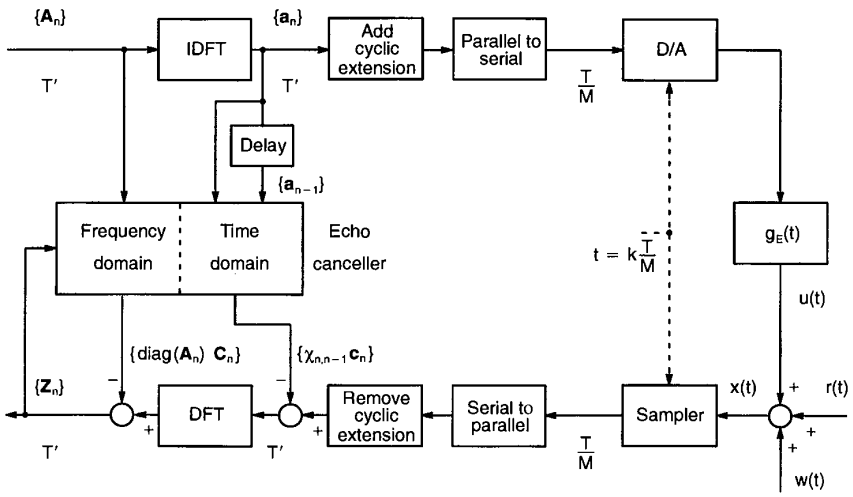


FIGURE 7.9 Configuration of an echo canceller for a DMT transmission system.

having length $N < M$. We initially assume $N \leq L + 1$. Furthermore, we assume that the boundaries of the received blocks are placed such that the last M samples of the n-th received block are expressed by the vector

$$\boldsymbol{x}_n = \Gamma_n^R \boldsymbol{h} + \Gamma_n \boldsymbol{h}_E + \boldsymbol{w}_n \ , \tag{7.28}$$

where Γ_n^R is the circulant matrix with elements given by the signals from the remote transmitter, and $\boldsymbol{h}_E$ is the vector of the echo impulse response extended with $M - N$ zeros. In the frequency domain, the echo is expressed as $\boldsymbol{U}_n = diag(\boldsymbol{A}_n)\boldsymbol{H}_E$, where $\boldsymbol{H}_E$ denotes the DFT of the vector $\boldsymbol{h}_E$. In this case, the echo canceller provides an echo estimate that is given by $\hat{\boldsymbol{U}}_n = diag(\boldsymbol{A}_n)\boldsymbol{C}_n$, where $\boldsymbol{C}_n$ denotes the DFT of the vector $\boldsymbol{c}_n$ of the N coefficients of the echo canceller filter extended with $M - N$ zeros. In practice, however, we need to consider the case $N > L + 1$. The expression of the cancellation error is then given by

$$\boldsymbol{z}_n = \boldsymbol{x}_n - \Psi_{n,n-1}\boldsymbol{c}_n \ , \tag{7.29}$$

where the vector of the last M elements of the n-th received block is now $\boldsymbol{x}_n = \Gamma_n^R \boldsymbol{h} + \Psi_{n,n-1}\boldsymbol{h}_E + \boldsymbol{w}_n$, and $\Psi_{n,n-1}$ is a $M \times M$ Toeplitz matrix given by

$$\Psi_{n,n-1} =$$

Ford Motor Co.
Exhibit 1030
Page 222

$$\begin{bmatrix} a_n(0) & a_n(M-1) & \cdots & a_n(M-L) & a_{n-1}(M-1) & \cdots & a_{n-1}(L+1) \\ a_n(1) & a_n(0) & \cdots & a_n(M-L+1) & a_n(M-L) & \cdots & a_{n-1}(L+2) \\ \vdots & & & \ddots & & & \vdots \\ a_n(M-1) & a_n(M-2) & \cdots & a_n(M-L-1) & a_n(M-L-2) & \cdots & a_n(0) \end{bmatrix} . \quad (7.30)$$

In the frequency domain, the cancellation error can be expressed as

$$\boldsymbol{Z}_n = \mathcal{F}_M \left(\boldsymbol{x}_n - \chi_{n,n-1} \boldsymbol{c}_n \right) - diag(\boldsymbol{A}_n) \boldsymbol{C}_n \ , \quad (7.31)$$

where $\chi_{n,n-1} = \Psi_{n,n-1} - \Gamma_n$ is a $M \times M$ upper triangular Toeplitz matrix. Equation (7.31) suggests a computationally efficient, two-part echo cancellation technique. First, in the time domain, a short convolution is performed and the result subtracted from the received signals to compensate for the insufficient length of the cyclic extension. Second, in the frequency domain, cancellation of the residual echo is performed over a set of M independent echo subchannels. Observing that Eq. (7.31) is equivalent to $\boldsymbol{Z}_n = \boldsymbol{X}_n - \tilde{\Psi}_{n,n-1} \boldsymbol{C}_n$, where $\tilde{\Psi}_{n,n-1} = \mathcal{F}_M \Psi_{n,n-1} \mathcal{F}_M^{-1}$, the echo canceller adaptation by the LMS algorithm in the frequency domain takes the form

$$\boldsymbol{C}_{n+1} = \boldsymbol{C}_n + \alpha \tilde{\Psi}^*_{n,n-1} \boldsymbol{Z}_n \ , \quad (7.32)$$

where α is the adaptation gain, and $\tilde{\Psi}^*_{n,n-1}$ denotes the transpose conjugate of $\tilde{\Psi}_{n,n-1}$. We note that, alternatively, echo canceller adaptation may also be performed by the algorithm $\boldsymbol{C}_{n+1} = \boldsymbol{C}_n + \alpha \, diag(\boldsymbol{A}^*_n) \boldsymbol{Z}_n$, which entails a substantially lower computational complexity than the LMS algorithm, at the price of a slower rate of convergence.

In DMT systems it is essential that the length of the channel impulse response be much less than the number of subchannels, so that the reduction in data rate due to the cyclic extension may be considered negligible. Therefore, equalization is adopted in practice to shorten the length of the channel impulse response. From Eq. (7.31), however, we observe that transceiver complexity depends on the relative lengths of the echo and of the channel impulse responses. To reduce the length of the cyclic extension as well as the computational complexity of the echo canceller, various methods have been proposed to shorten both the channel and the echo impulse responses jointly [8].

7.5 Summary and Conclusions

Digital signal processing techniques for echo cancellation provide large echo attenuation, and eliminate the need for additional line interfaces and digital-to-analog and analog-to-digital converters that are required by echo cancellation in the analog signal domain.

The realization of digital echo cancellers in transceivers for high-speed full-duplex data transmission today is possible at a low cost thanks to the advances in VLSI technology. Digital techniques for echo cancellation are also appropriate for near-end crosstalk cancellation in transceivers for transmission over voice-grade cables at rates of several megabits per second for local-area network applications.

In voiceband modems for data transmission over the telephone network, digital techniques for echo cancellation also allow a precise tracking of the carrier phase and frequency shift of far-end echos.

References

[1] Cherubini, G., Analysis of the convergence behavior of adaptive distributed-arithmetic echo cancellers. *IEEE Trans. Commun.*, 41(11), 1703–1714, 1993.

Ford Motor Co.
Exhibit 1030
Page 223

[2] Cherubini, G., Ölçer, S., and Ungerboeck, G., A quaternaty partial-response class-IV transceiver for 125 Mbit/s data transmission over unshielded twisted-pair cables: Principles of operation and VLSI realization. *IEEE J. Sel. Areas Commun.*, 13(9), 1656–1669, 1995.

[3] Cherubini, G., Creigh, J., Ölçer, S., Rao, S.K., and Ungerboeck, G., 100BASE-T2: A new standard for 100 Mb/s Ethernet transmission over voice-grade cables. *IEEE Commun. Mag.*, 35(11), 115–122, 1997.

[4] Duttweiler, D.L., Adaptive filter performance with nonlinearities in the correlation multiplier. *IEEE Trans. Acoust., Speech, Signal Processing*, 30(8), 578–586, 1982.

[5] Falconer, D.D., Adaptive reference echo-cancellation. *IEEE Trans. Commun.*, 30(9), 2083–2094, 1982.

[6] Ho, M., Cioffi, J.M. and Bingham, J.A.C., Discrete multitone echo cancellation. *IEEE Trans. Commun.*, 44(7), 817–825, 1996.

[7] Lee, E.A. and Messerschmitt, D.G., *Digital Communication*, 2nd ed., Kluwer Academic Publishers, Boston MA, 1994.

[8] Melsa, P.J.W., Younce, R.C., and Rohrs, C.E., Impulse response shortening for discrete multitone transceivers. *IEEE Trans. Commun.*, 44(12), 1662–1672, 1996.

[9] Messerschmitt, D.G., Echo cancellation in speech and data transmission. *IEEE J. Sel. Areas Commun.*, 2(2), 283–297, 1984.

[10] Messerschmitt, D.G., Design issues for the ISDN U-Interface transceiver. *IEEE J. Sel. Areas Commun.*, 4(8), 1281–1293, 1986.

[11] Ruiz, A., Cioffi, J.M., and Kasturia, S., Discrete multiple tone modulation with coset coding for the spectrally shaped channel. *IEEE Trans. Commun.*, 40(6), 1012–1029, 1992.

[12] Smith, M.J., Cowan, C.F.N., and Adams, P.F., Nonlinear echo cancellers based on transpose distributed arithmetic. *IEEE Trans. Circuits and Systems*, 35(1), 6–18, 1988.

[13] Ungerboeck, G., Theory on the speed of convergence in adaptive equalizers for digital communication. *IBM J. Res. Develop.*, 16(6), 546–555, 1972.

[14] Weinstein, S.B., A passband data-driven echo-canceller for full-duplex transmission on two-wire circuits. *IEEE Trans. Commun.*, 25(7), 654–666, 1977.

Further Information

For further information on adaptive transversal filters with application to echo cancellation, see *Adaptive Filters: Structures, Algorithms, and Applications*, M.L. Honig and D.G. Messerschmitt, Kluwer, 1984.

Ford Motor Co.
Exhibit 1030
Page 224

8

Pseudonoise Sequences

8.1 Introduction 8-1
8.2 m Sequences 8-2
8.3 The q-ary Sequences with Low Autocorrelation 8-4
8.4 Families of Sequences with Low Crosscorrelation 8-5
Gold and Kasami Sequences • Quaternary Sequences with Low Crosscorrelation • Binary Kerdock Sequences
8.5 Aperiodic Correlation 8-8
Barker Sequences • Sequences with High Merit Factor • Sequences with Low Aperiodic Crosscorrelation
8.6 Other Correlation Measures 8-10
Partial-Period Correlation • Mean Square Correlation • Optical Orthogonal Codes
Defining Terms 8-12
References 8-12
Further Information 8-13

Tor Helleseth
University of Bergen

P. Vijay Kumar
University of Southern California

8.1 Introduction

Pseudonoise sequences (PN sequences), also referred to as pseudorandom sequences, are sequences that are deterministically generated and yet possess some properties that one would expect to find in randomly generated sequences. Applications of PN sequences include signal synchronization, navigation, radar ranging, random number generation, spread-spectrum communications, multipath resolution, cryptography, and signal identification in multiple-access communication systems. The *correlation* between two sequences $\{x(t)\}$ and $\{y(t)\}$ is the complex inner product of the first sequence with a shifted version of the second sequence. The correlation is called 1) an autocorrelation if the two sequences are the same, 2) a crosscorrelation if they are distinct, 3) a periodic correlation if the shift is a cyclic shift, 4) an aperiodic correlation if the shift is not cyclic, and 5) a partial-period correlation if the inner product involves only a partial segment of the two sequences. More precise definitions are given subsequently.

Binary $\boldsymbol{m}$ **sequences**, defined in the next section, are perhaps the best-known family of PN sequences. The balance, run-distribution, and autocorrelation properties of these sequences mimic those of random sequences. It is perhaps the random-like correlation properties of PN sequences that makes them most attractive in a communications system, and it is common to refer to any collection of low-correlation sequences as a family of PN sequences.

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 225

Section 8.2 begins by discussing m sequences. Thereafter, the discussion continues with a description of sequences satisfying various correlation constraints along the lines of the accompanying self-explanatory figure, Fig. 8.1. Expanded tutorial discussions on pseudorandom sequences may be found in [14], in [15, Chapter 5] and in [6].

8.2 m Sequences

A binary $\{0,1\}$ **shift-register sequence** $\{s(t)\}$ is a sequence that satisfies a linear recurrence relation of the form

$$\sum_{i=0}^{r} f_i s(t+i) = 0\ , \qquad \text{for all } t \geq 0 \tag{8.1}$$

where $r \geq 1$ is the *degree* of the recursion; the coefficients f_i belong to the finite field $GF(2) = \{0,1\}$ where the leading coefficient $f_r = 1$. Thus, both sequences $\{a(t)\}$ and $\{b(t)\}$ appearing in Fig. 8.2 are shift-register sequences. A sequence satisfying a recursion of the form in Eq. (8.1) is said to have *characteristic polynomial* $f(x) = \sum_{i=0}^{r} f_i x^i$. Thus, $\{a(t)\}$ and $\{b(t)\}$ have characteristic polynomials given by $f(x) = x^3+x+1$ and $f(x) = x^3+x^2+1$, respectively.

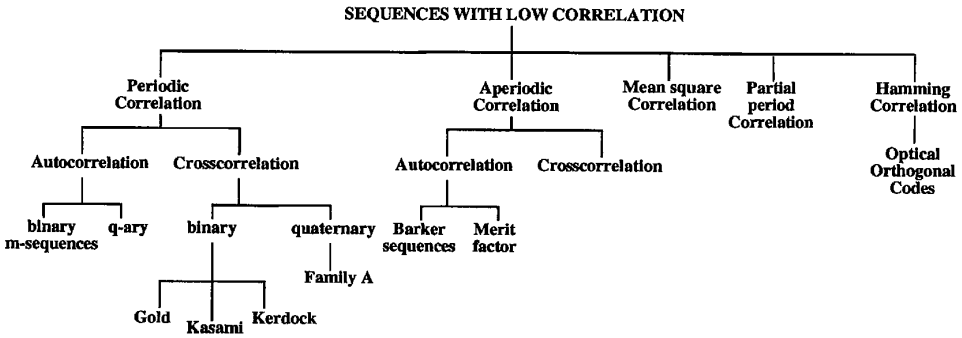


FIGURE 8.1 Overview of pseudonoise sequences.

Since an r-bit binary shift register can assume a maximum of 2^r different states, it follows that every shift-register sequence $\{s(t)\}$ is eventually periodic with period $n \leq 2^r$, i.e.,

$$s(t) = s(t+n), \qquad \text{for all } t \geq N$$

for some integer N. In fact, the maximum period of a shift-register sequence is $2^r - 1$, since a shift register that enters the all-zero state will remain forever in that state. The upper shift register in Fig. 8.2 when initialized with starting state 0 0 1 generates the periodic sequence $\{a(t)\}$ given by

$$0010111 \quad 0010111 \quad 0010111 \quad \cdots \tag{8.2}$$

of period $n = 7$. It follows then that this shift register generates sequences of maximal period starting from any nonzero initial state.

An m sequence is simply a binary shift-register sequence having maximal period. For every $r \geq 1$, m sequences are known to exist. The periodic **autocorrelation** function θ_s of

Ford Motor Co.
Exhibit 1030
Page 226

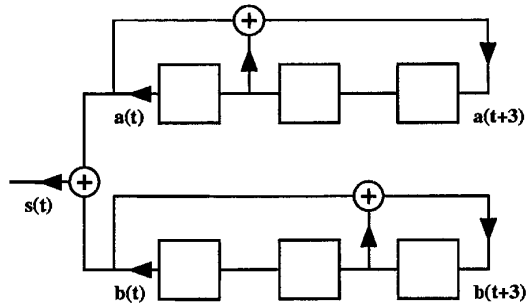


FIGURE 8.2 An example Gold sequence generator. Here $\{a(t)\}$ and $\{b(t)\}$ are m sequences of length 7.

a binary $\{0,1\}$ sequence $\{s(t)\}$ of period n is defined by

$$\theta_s(\tau) = \sum_{t=0}^{n-1} (-1)^{s(t+\tau)-s(t)} , \qquad 0 \le \tau \le n-1$$

An m sequence of length $2^r - 1$ has the following attributes. 1) *Balance property:* in each period of the m sequence there are 2^{r-1} ones and $2^{r-1} - 1$ zeros. 2) *Run property:* every nonzero binary s-tuple, $s \le r$ occurs 2^{r-s} times, the all-zero s-tuple occurs $2^{r-s} - 1$ times. 3) *Two-level autocorrelation function:*

$$\theta_s(\tau) = \begin{cases} n & \text{if } \tau = 0 \\ -1 & \text{if } \tau \neq 0 \end{cases} \tag{8.3}$$

The first two properties follow immediately from the observation that every nonzero r-tuple occurs precisely once in each period of the m sequence. For the third property, consider the difference sequence $\{s(t+\tau) - s(t)\}$ for $\tau \neq 0$. This sequence satisfies the same recursion as the m sequence $\{s(t)\}$ and is clearly not the all-zero sequence. It follows, therefore, that $\{s(t+\tau) - s(t)\} \equiv \{s(t+\tau')\}$ for some τ', $0 \le \tau' \le n-1$, i.e., is a different cyclic shift of the m sequence $\{s(t)\}$. The balance property of the sequence $\{s(t+\tau')\}$ then gives us attribute 3. The m sequence $\{a(t)\}$ in Eq. (8.2) can be seen to have the three listed properties.

If $\{s(t)\}$ is any sequence of period n and d is an integer, $1 \le d \le n$, then the mapping $\{s(t)\} \to \{s(dt)\}$ is referred to as a *decimation* of $\{s(t)\}$ by the integer d. If $\{s(t)\}$ is an m sequence of period $n = 2^r - 1$ and d is an integer relatively prime to $2^r - 1$, then the decimated sequence $\{s(dt)\}$ clearly also has period n. Interestingly, it turns out that the sequence $\{s(dt)\}$ is always also an m sequence of the same period. For example, when $\{a(t)\}$ is the sequence in Eq. (8.2), then

$$a(3t) = 0011101 \quad 0011101 \quad 0011101 \quad \cdots \tag{8.4}$$

and

$$a(2t) = 0111001 \quad 0111001 \quad 0111001 \quad \cdots \tag{8.5}$$

The sequence $\{a(3t)\}$ is also an m sequence of period 7, since it satisfies the recursion

$$s(t+3) + s(t+2) + s(t) = 0 \qquad \text{for all } t$$

of degree $r = 3$. In fact $\{a(3t)\}$ is precisely the sequence labeled $\{b(t)\}$ in Fig. 8.2. The sequence $\{a(2t)\}$ is simply a cyclically shifted version of $\{a(t)\}$ itself; this property holds in general. If $\{s(t)\}$ is any m sequence of period $2^r - 1$, then $\{s(2t)\}$ will always be a shifted

Ford Motor Co.
Exhibit 1030
Page 227

version of the same m sequence. Clearly, the same is true for decimations by any power of 2.

Starting from an m sequence of period $2^r - 1$, it turns out that one can generate all m sequences of the same period through decimations by integers d relatively prime to $2^r - 1$. The set of integers d, $1 \leq d \leq 2^r - 1$ satisfying $(d, 2^r - 1) = 1$ forms a group under multiplication modulo $2^r - 1$, with the powers $\{2^i | \, 0 \leq i \leq r - 1\}$ of 2 forming a subgroup of order r. Since decimation by a power of 2 yields a shifted version of the same m sequence, it follows that the number of distinct m sequences of period $2^r - 1$ is $[\phi(2^r - 1)/r]$ where $\phi(n)$ denotes the number of integers d, $1 \leq d \leq n$, relatively prime to n. For example, when $r = 3$, there are just two cyclically distinct m sequences of period 7, and these are precisely the sequences $\{a(t)\}$ and $\{b(t)\}$ discussed in the preceding paragraph. Tables provided in [12] can be used to determine the characteristic polynomial of the various m sequences obtainable through the decimation of a single given m sequence. The classical reference on m sequences is [4].

If one obtains a sequence of some large length n by repeatedly tossing an unbiased coin, then such a sequence will very likely satisfy the balance, run, and autocorrelation properties of an m sequence of comparable length. For this reason, it is customary to regard the extent to which a given sequence possesses these properties as a measure of randomness of the sequence. Quite apart from this, in many applications such as signal synchronization and radar ranging, it is desirable to have sequences $\{s(t)\}$ with low autocorrelation sidelobes i.e., $|\theta_s(\tau)|$ is small for $\tau \neq 0$. Whereas m sequences are a prime example, there exist other methods of constructing binary sequences with low out-of-phase autocorrelation.

Sequences $\{s(t)\}$ of period n having an autocorrelation function identical to that of an m sequence, i.e., having θ_s satisfying Eq. (8.3) correspond to well-studied combinatorial objects known as *cyclic Hadamard difference sets.* Known infinite families fall into three classes 1) Singer and Gordon, Mills and Welch, 2) quadratic residue, and 3) twin-prime difference sets. These correspond, respectively, to sequences of period n of the form $n = 2^r - 1, r \geq 1$; n prime; and $n = p(p+2)$ with both p and $p+2$ being prime in the last case. For a detailed treatment of cyclic difference sets, see [2]. A recent observation by Maschietti in [9] provides additional families of cyclic Hadamard difference sets that also correspond to sequences of period $n = 2^r - 1$.

8.3 The q-ary Sequences with Low Autocorrelation

As defined earlier, the autocorrelation of a binary $\{0, 1\}$ sequence $\{s(t)\}$ leads to the computation of the inner product of an $\{-1, +1\}$ sequence $\{(-1)^{s(t)}\}$ with a cyclically shifted version $\{(-1)^{s(t+\tau)}\}$ of itself. The $\{-1, +1\}$ sequence is transmitted as a phase shift by either 0° and 180° of a radio-frequency carrier, i.e., using binary phase-shift keying (PSK) modulation. If the modulation is q-ary PSK, then one is led to consider sequences $\{s(t)\}$ with symbols in the set Z_q, i.e., the set of integers modulo q. The relevant autocorrelation function $\theta_s(\tau)$ is now defined by

$$\theta_s(\tau) = \sum_{t=0}^{n-1} \omega^{s(t+\tau) - s(t)}$$

where n is the period of $\{s(t)\}$ and ω is a complex primitive qth root of unity. It is possible to construct sequences $\{s(t)\}$ over Z_q whose autocorrelation function satisfies

$$\theta_s(\tau) = \begin{cases} n & \text{if } \tau = 0 \\ 0 & \text{if } \tau \neq 0 \end{cases}$$

Ford Motor Co.
Exhibit 1030
Page 228

For obvious reasons, such sequences are said to have an *ideal autocorrelation function.*

We provide without proof two sample constructions. The sequences in the first construction are given by

$$s(t) = \begin{cases} t^2/2 \pmod{n} & \text{when } n \text{ is even} \\ t(t+1)/2 \pmod{n} & \text{when } n \text{ is odd} \end{cases}$$

Thus, this construction provides sequences with ideal autocorrelation for any period n. Note that the size q of the sequence symbol alphabet equals n when n is odd and $2n$ when n is even.

The second construction also provides sequences over Z_q of period n but requires that n be a perfect square. Let $n = r^2$ and let π be an arbitrary permutation of the elements in the subset $\{0, 1, 2, \ldots, (r-1)\}$ of Z_n: Let g be an arbitrary function defined on the subset $\{0, 1, 2, \ldots, r-1\}$ of Z_n. Then any sequence of the form

$$s(t) = rt_1\pi(t_2) + g(t_2) \pmod{n}$$

where $t = rt_1 + t_2$ with $0 \leq t_1, t_2 \leq r-1$ is the base-r decomposition of t, has an ideal autocorrelation function. When the alphabet size q equals or divides the period n of the sequence, ideal-autocorrelation sequences also go by the name *generalized bent functions.* For details, see [6].

8.4 Families of Sequences with Low Crosscorrelation

Given two sequences $\{s_1(t)\}$ and $\{s_2(t)\}$ over Z_q of period n, their **crosscorrelation** function $\theta_{1,2}(\tau)$ is defined by

$$\theta_{1,2}(\tau) = \sum_{t=0}^{n-1} \omega^{s_1(t+\tau)-s_2(t)}$$

where ω is a primitive qth root of unity. The crosscorrelation function is important in code-division multiple-access (CDMA) communication systems. Here, each user is assigned a distinct signature sequence and to minimize interference due to the other users, it is desirable that the signature sequences have pairwise, low values of crosscorrelation function. To provide the system in addition with a self-synchronizing capability, it is desirable that the signature sequences have low values of the autocorrelation function as well.

Let $\mathcal{F} = \{\{s_i(t)\} \mid 1 \leq i \leq M\}$ be a family of M sequences $\{s_i(t)\}$ over Z_q each of period n. Let $\theta_{i,j}(\tau)$ denote the crosscorrelation between the ith and jth sequence at shift τ, i.e.,

$$\theta_{i,j}(\tau) = \sum_{t=0}^{n-1} \omega^{s_i(t+\tau)-s_j(t)}, \qquad 0 \leq \tau \leq n-1$$

The classical goal in sequence design for CDMA systems has been minimization of the parameter

$$\theta_{\max} = \max\{|\theta_{i,j}(\tau)| \mid \text{ either } i \neq j \text{ or } \tau \neq 0\}$$

for fixed n and M. It should be noted though that, in practice, because of data modulation the correlations that one runs into are typically of an aperiodic rather than a periodic nature (see Section 8.5). The problem of designing for low aperiodic correlation, however, is a more difficult one. A typical approach, therefore, has been to design based on periodic correlation, and then to analyze the resulting design for its aperiodic correlation properties.

Ford Motor Co.
Exhibit 1030
Page 229

Again, in many practical systems, the mean square correlation properties are of greater interest than the worst-case correlation represented by a parameter such as $\theta_{\max}$. The mean square correlation is discussed in Section 8.6.

Bounds on the minimum possible value of $\theta_{\max}$ for given period n, family size M, and alphabet size q are available that can be used to judge the merits of a particular sequence design. The most efficient bounds are those due to Welch, Sidelnikov, and Levenshtein, see [6]. In CDMA systems, there is greatest interest in designs in which the parameter $\theta_{\max}$ is in the range $\sqrt{n} \leq \theta_{\max} \leq 2\sqrt{n}$. Accordingly, Table 8.1 uses the Welch, Sidelnikov, and Levenshtein bounds to provide an order-of-magnitude upper bound on the family size M for certain $\theta_{\max}$ in the cited range.

Practical considerations dictate that q be small. The bit-oriented nature of electronic hardware makes it preferable to have q a power of 2. With this in mind, a description of some efficient sequence families having low auto- and crosscorrelation values and alphabet sizes $q = 2$ and $q = 4$ are described next.

TABLE 8.1 Bounds on Family Size M for Given $n, \theta_{\max}$

$\theta_{\max}$	Upper bound on M $q = 2$	Upper Bound on M $q > 2$
$\sqrt{n}$	$n/2$	n
$\sqrt{2n}$	n	$n^2/2$
$2\sqrt{n}$	$3n^2/10$	$n^3/2$

8.4.1 Gold and Kasami Sequences

Given the low autocorrelation sidelobes of an m sequence, it is natural to attempt to construct families of low correlation sequences starting from m sequences. Two of the better known constructions of this type are the families of Gold and Kasami sequences.

Let r be odd and $d = 2^k + 1$ where $k, 1 \leq k \leq r-1$, is an integer satisfying $(k, r) = 1$. Let $\{s(t)\}$ be a cyclic shift of an m sequence of period $n = 2^r - 1$ that satisfies $S(dt) \not\equiv 0$ and let $\mathcal{G}$ be the *Gold* family of $2^r + 1$ sequences given by

$$\mathcal{G} = \{s(t)\} \cup \{s(dt)\} \cup \{\{s(t) + s(d[t+\tau])\} \mid 0 \leq \tau \leq n-1\}$$

Then each sequence in $\mathcal{G}$ has period $2^r - 1$ and the maximum-correlation parameter $\theta_{\max}$ of $\mathcal{G}$ satisfies

$$\theta_{\max} \leq \sqrt{2^{r+1}} + 1$$

An application of the Sidelnikov bound coupled with the information that $\theta_{\max}$ must be an odd integer yields that for the family $\mathcal{G}$, $\theta_{\max}$ is as small as it can possibly be. In this sense the family $\mathcal{G}$ is an optimal family. We remark that these comments remain true even when d is replaced by the integer $d = 2^{2k} - 2^k + 1$ with the conditions on k remaining unchanged.

The Gold family remains the best-known family of m sequences having low crosscorrelation. Applications include the Navstar Global Positioning System whose signals are based on Gold sequences.

The family of Kasami sequences has a similar description. Let $r = 2v$ and $d = 2^v + 1$. Let $\{s(t)\}$ be a cyclic shift of an m sequence of period $n = 2^r - 1$ that satisfies $s(dt) \not\equiv 0$, and consider the family of Kasami sequences given by

$$\mathcal{K} = \{s(t)\} \cup \{\{s(t) + s(d[t+\tau])\} \mid 0 \leq \tau \leq 2^v - 2\}$$

Ford Motor Co.
Exhibit 1030
Page 230

Then the Kasami family $\mathcal{K}$ contains 2^v sequences of period $2^r - 1$. It can be shown that in this case

$$\theta_{\max} = 1 + 2^v$$

This time an application of the Welch bound and the fact that $\theta_{\max}$ is an integer shows that the Kasami family is optimal in terms of having the smallest possible value of $\theta_{\max}$ for given n and M.

8.4.2 Quaternary Sequences with Low Crosscorrelation

The entries in Table 8.1 suggest that nonbinary (i.e., $q > 2$) designs may be used for improved performance. A family of quaternary sequences that outperform the Gold and Kasami sequences is now discussed below.

Let $f(x)$ be the characteristic polynomial of a binary m sequence of length $2^r - 1$ for some integer r. The coefficients of $f(x)$ are either 0 or 1. Now, regard $f(x)$ as a polynomial over Z_4 and form the product $(-1)^r f(x)f(-x)$. This can be seen to be a polynomial in x^2. Define the polynomial $g(x)$ of degree r by setting $g(x^2) = (-1)^r f(x)f(-x)$. Let $g(x) = \sum_{i=0}^{r} g_i x^i$ and consider the set of all quaternary sequences $\{a(t)\}$ satisfying the recursion $\sum_{i=0}^{r} g_i a(t+i) = 0$ for all t.

It turns out that with the exception of the all-zero sequence, all of the sequences generated in this way have period $2^r - 1$. Thus, the recursion generates a family $\mathcal{A}$ of $2^r + 1$ cyclically distinct quaternary sequences. Closer study reveals that the maximum correlation parameter $\theta_{\max}$ of this family satisfies $\theta_{\max} \leq 1 + \sqrt{2^r}$. Thus, in comparison to the family of Gold sequences, the family $\mathcal{A}$ offers a lower value of $\theta_{\max}$ (by a factor of $\sqrt{2}$) for the same family size. In comparison to the set of Kasami sequences, it offers a much larger family size for the same bound on $\theta_{\max}$. Family $\mathcal{A}$ sequences may be found discussed in [16, 3].

We illustrate with an example. Let $f(x) = x^3 + x + 1$ be the characteristic polynomial of the m sequence $\{a(t)\}$ in Eq. (8.1). Then over Z_4

$$g\left(x^2\right) = (-1)^3 f(x)f(-x) = x^6 + 2x^4 + x^2 + 3$$

so that $g(x) = x^3 + 2x^2 + x + 3$. Thus, the sequences in family $\mathcal{A}$ are generated by the recursion $s(t+3) + 2s(t+2) + s(t+1) + 3s(t) = 0 \bmod 4$. The corresponding shift register is shown in Fig. 8.3. By varying initial conditions, this shift register can be made to generate nine cyclically distinct sequences, each of length 7. In this case $\theta_{\max} \leq 1 + \sqrt{8}$.

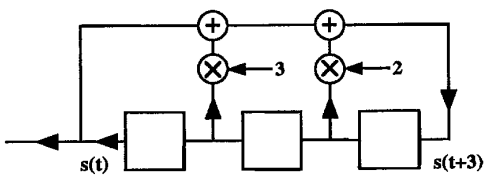


FIGURE 8.3 Shift register that generates family A quaternary sequences $\{s(t)\}$ of period 7.

8.4.3 Binary Kerdock Sequences

The Gold and Kasami families of sequences are closely related to binary linear cyclic codes. It is well known in coding theory that there exists nonlinear binary codes whose performance

Ford Motor Co.
Exhibit 1030
Page 231

exceeds that of the best possible linear code. Surprisingly, some of these examples come from binary codes, which are images of *linear quaternary* ($q = 4$) *codes* under the Gray map: $0 \to 00$, $1 \to 01$, $2 \to 11$, $3 \to 10$. A prime example of this is the Kerdock code, which recently has been shown to be the Gray image of a quaternary linear code. Thus, it is not surprising that the Kerdock code yields binary sequences that significantly outperform the family of Kasami sequences.

The Kerdock sequences may be constructed as follows: let $f(x)$ be the characteristic polynomial of an m sequence of period $2^r - 1, r$ odd. As before, regarding $f(x)$ as a polynomial over Z_4 (which happens to have $\{0,1\}$ coefficients), let the polynomial $g(x)$ over Z_4 be defined via $g(x^2) = -f(x)f(-x)$. [Thus, $g(x)$ is the characteristic polynomial of a family $\mathcal{A}$ sequence set of period $2^r - 1$.] Set $h(x) = -g(-x) = \sum_{i=0}^{r} h_i x^i$, and let S be the set of all Z_4 sequences satisfying the recursion $\sum_{i=0}^{r} h_i s(t+i) = 0$. Then S contain 4^r-distinct sequences corresponding to all possible distinct initializations of the shift register.

Let T denote the subset S of size 2^r-consisting of those sequences corresponding to initializations of the shift register only using the symbols 0 and 2 in Z_4. Then the set $S - T$ of size $4^r - 2^r$ contains a set $\mathcal{U}$ of 2^{r-1} cyclically distinct sequences each of period $2(2^r - 1)$. Given $x = a + 2b \in Z_4$ with $a, b \in \{0,1\}$, let μ denote the most significant bit (MSB) map $\mu(x) = b$. Let $\mathcal{K}_E$ denote the family of 2^{r-1} binary sequences obtained by applying the map μ to each sequence in $\mathcal{U}$. It turns out that each sequence in $\mathcal{U}$ also has period $2(2^r - 1)$ and that, furthermore, for the family $\mathcal{K}_E$, $\theta_{\max} \leq 2 + \sqrt{2^{r+1}}$. Thus, $\mathcal{K}_E$ is a much larger family than the Kasami family, while having almost exactly the same value of $\theta_{\max}$.

For example, taking $r = 3$ and $f(x) = x^3 + x + 1$, we have from the previous family $\mathcal{A}$ example that $g(x) = x^3 + 2x^2 + x + 3$, so that $h(x) = -g(-x) = x^3 + 2x^2 + x + 1$. Applying the MSB map to the head of the shift register, and discarding initializations of the shift register involving only 0's and 2's yields a family of four cyclically distinct binary sequences of period 14. Kerdock sequences are discussed in [6, 11, 1, 17].

8.5 Aperiodic Correlation

Let $\{x(t)\}$ and $\{y(t)\}$ be complex-valued sequences of length (or period) n, not necessarily distinct. Their *aperiodic correlation* values $\{\rho_{x,y}(\tau) | \ -(n-1) \leq \tau \leq n-1\}$ are given by

$$\rho_{x,y}(\tau) = \sum_{t=\max\{0,-\tau\}}^{\min\{n-1,n-1-\tau\}} x(t+\tau)y^*(t)$$

where $y^*(t)$ denotes the complex conjugate of $y(t)$. When $x \equiv y$, we will abbreviate and write ρ_x in place of $\rho_{x,y}$. The sequences described next are perhaps the most famous example of sequences with low-aperiodic autocorrelation values.

8.5.1 Barker Sequences

A binary $\{-1, +1\}$ sequence $\{s(t)\}$ of length n is said to be a *Barker sequence* if the aperiodic autocorrelation values $\rho_s(\tau)$ satisfy $|\rho_s(\tau)| \leq 1$ for all $\tau, -(n-1) \leq \tau \leq n-1$. The Barker property is preserved under the following transformations:

$$s(t) \to -s(t), \qquad s(t) \to (-1)^t s(t) \quad \text{and} \quad s(t) \to s(n-1-t)$$

Ford Motor Co.
Exhibit 1030
Page 232

as well as under compositions of the preceding transformations. Only the following Barker sequences are known:

$$\begin{aligned} n &= 2 \quad ++ \\ n &= 3 \quad ++- \\ n &= 4 \quad +++- \\ n &= 5 \quad +++-+ \\ n &= 7 \quad +++--+- \\ n &= 11 \quad +++---+--+- \\ n &= 13 \quad +++++--++-+-+ \end{aligned}$$

where $+$ denotes $+1$ and $-$ denotes -1 and sequences are generated from these via the transformations already discussed. It is known that if any other Barker sequence exists, it must have length $n > 1{,}898{,}884$, that is a multiple of 4.

For an upper bound to the maximum out-of-phase aperiodic autocorrelation of an m sequence, see [13].

8.5.2 Sequences with High Merit Factor

The *merit factor* F of a $\{-1, +1\}$ sequence $\{s(t)\}$ is defined by

$$F = \frac{n^2}{2\sum_{\tau=1}^{n-1} \rho_s^2(\tau)}$$

Since $\rho_s(\tau) = \rho_s(-\tau)$ for $1 \leq |\tau| \leq n-1$ and $\rho_s(0) = n$, factor F may be regarded as the ratio of the square of the in-phase autocorrelation, to the sum of the squares of the out-of-phase aperiodic autocorrelation values. Thus, the merit factor is one measure of the aperiodic autocorrelation properties of a binary $\{-1, +1\}$ sequence. It is also closely connected with the signal to self-generated noise ratio of a communication system in which coded pulses are transmitted and received.

Let F_n denote the largest merit factor of any binary $\{-1, +1\}$ sequence of length n. For example, at length $n = 13$, the Barker sequence of length 13 has a merit factor $F = F_{13} = 14.08$. Assuming a certain ergodicity postulate it was established by Golay that $\lim_{n\to\infty} F_n = 12.32$. Exhaustive computer searches carried out for $n \leq 40$ have revealed the following.

1. For $1 \leq n \leq 40$, $n \neq 11, 13$,

$$3.3 \leq F_n \leq 9.85 \ ,$$

2. $F_{11} = 12.1$, $F_{13} = 14.08$.

The value F_{11} is also achieved by a Barker sequence. From partial searches, for lengths up to 117, the highest known merit factor is between 8 and 9.56; for lengths from 118 to 200, the best-known factor is close to 6. For lengths > 200, statistical search methods have failed to yield a sequence having merit factor exceeding 5.

An *offset sequence* is one in which a fraction θ of the elements of a sequence of length n are chopped off at one end and appended to the other end, i.e., an offset sequence is a cyclic shift of the original sequence by $n\theta$ symbols. It turns out that the asymptotic merit factor of m sequences is equal to 3 and is independent of the particular offset of the m sequence. There exist offsets of sequences associated with quadratic-residue and twin-prime difference sets that achieve a larger merit factor of 6. Details may be found in [7].

8.5.3 Sequences with Low Aperiodic Crosscorrelation

If $\{u(t)\}$ and $\{v(t)\}$ are sequences of length $2n-1$ defined by

$$u(t) = \begin{cases} x(t) & \text{if } 0 \le t \le n-1 \\ 0 & \text{if } n \le t \le 2n-2 \end{cases}$$

and

$$v(t) = \begin{cases} y(t) & \text{if } 0 \le t \le n-1 \\ 0 & \text{if } n \le t \le 2n-2 \end{cases}$$

then

$$\{\rho_{x,y}(\tau) \mid -(n-1) \le \tau \le n-1\} = \{\theta_{u,v}(\tau) \mid 0 \le \tau \le 2n-2\} \tag{8.6}$$

Given a collection

$$U = \{\{x_i(t)\} \mid 1 \le i \le M\}$$

of sequences of length n over Z_q, let us define

$$\rho_{\max} = \max\{|\rho_{a,b}(\tau)| \mid a, b \in U \ , \ \text{either } a \ne b \text{ or } \tau \ne 0\}$$

It is clear from Eq. (8.6) how bounds on the *periodic* correlation parameter $\theta_{\max}$ can be adapted to give bounds on $\rho_{\max}$. Translation of the Welch bound gives that for every integer $k \ge 1$,

$$\rho_{\max}^{2k} \ge \left(\frac{n^{2k}}{M(2n-1)-1}\right)\left\{\frac{M(2n-1)}{\binom{2n+k-2}{k}} - 1\right\}$$

Setting $k = 1$ in the preceding bound gives

$$\rho_{\max} \ge n\sqrt{\frac{M-1}{M(2n-1)-1}}$$

Thus, for fixed M and large n, Welch's bound gives

$$\rho_{\max} \ge \mathcal{O}\left(n^{1/2}\right)$$

There exist sequence families which asymptotically achieve $\rho_{\max} \approx \mathcal{O}(n^{1/2})$, [10].

8.6 Other Correlation Measures

8.6.1 Partial-Period Correlation

The *partial-period (p-p) correlation* between the sequences $\{u(t)\}$ and $\{v(t)\}$ is the collection $\{\Delta_{u,v}(l,\tau,t_0) \mid 1 \le l \le n, 0 \le \tau \le n-1, 0 \le t_0 \le n-1\}$ of inner products

$$\Delta_{u,v}(l,\tau,t_0) = \sum_{t=t_0}^{t=t_0+l-1} u(t+\tau)v^*(t)$$

where l is the length of the partial period and the sum $t+\tau$ is again computed modulo n.

In direct-sequence CDMA systems, the pseudorandom signature sequences used by the various users are often very long for reasons of data security. In such situations, to minimize receiver hardware complexity, correlation over a partial period of the signature sequence is

Ford Motor Co.
Exhibit 1030
Page 234

often used to demodulate data, as well as to achieve synchronization. For this reason, the p-p correlation properties of a sequence are of interest.

Researchers have attempted to determine the moments of the p-p correlation. Here the main tool is the application of the Pless power-moment identities of coding theory [8]. The identities often allow the first and second p-p correlation moments to be completely determined. For example, this is true in the case of m sequences (the remaining moments turn out to depend upon the specific characteristic polynomial of the m sequence). Further details may be found in [15].

8.6.2 Mean Square Correlation

Frequently in practice, there is a greater interest in the mean-square correlation distribution of a sequence family than in the parameter $\theta_{\max}$. Quite often in sequence design, the sequence family is derived from a linear, binary cyclic code of length n by picking a set of cyclically distinct sequences of period n. The families of Gold and Kasami sequences are so constructed. In this case, as pointed out by Massey, the mean square correlation of the family can be shown to be either optimum or close to optimum, under certain easily satisfied conditions, imposed on the minimum distance of the dual code. A similar situation holds even when the sequence family does not come from a linear cyclic code. In this sense, mean square correlation is not a very discriminating measure of the correlation properties of a family of sequences. An expanded discussion of this issue may be found in [5].

8.6.3 Optical Orthogonal Codes

Given a pair of $\{0,1\}$ sequences $\{s_1(t)\}$ and $\{s_2(t)\}$ each having period n, we define the *Hamming correlation* function $\theta_{12}(\tau)$, $0 \leq \tau \leq n-1$, by

$$\theta_{12}(\tau) = \sum_{t=0}^{n-1} s_1(t+\tau)s_2(t)$$

Such correlations are of interest, for instance, in optical communication systems where the 1's and 0's in a sequence correspond to the presence or absence of pulses of transmitted light.

An (n, w, λ) optical orthogonal code (OOC) is a family $\mathcal{F} = \{\{s_i(t)\} \mid i = 1, 2, \ldots, M\}$, of M $\{0,1\}$ sequences of period n, constant Hamming weight w, where w is an integer lying between 1 and $n-1$ satisfying $\theta_{ij}(\tau) \leq \lambda$ whenever either $i \neq j$ or $\tau \neq 0$.

Note that the Hamming distance $d_{a,b}$ between a period of the corresponding codewords $\{a(t)\}$, $\{b(t)\}$, $0 \leq t \leq n-1$ in an (n, w, λ) OOC having Hamming correlation $\rho, 0 \leq \rho \leq \lambda$, is given by $d_{a,b} = 2(w - \rho)$, and, thus, OOCs are closely related to constant-weight error correcting codes. Given an (n, w, λ) OOC, by enlarging the OOC to include every cyclic shift of each sequence in the code, one obtains a constant-weight, minimum distance $d_{\min} \geq 2(w - \lambda)$ code. Conversely, given a constant-weight cyclic code of length n, weight w and minimum distance $d_{\min}$, one can derive an (n, w, λ) OOC code with $\lambda \leq w - d_{\min}/2$ by partitioning the code into cyclic equivalence classes and then picking precisely one representative from each equivalence class of size n.

By making use of this connection, one can derive bounds on the size of an OOC from known bounds on the size of constant-weight codes. The bound given next follows directly from the Johnson bound for constant weight codes [8]. The number $M(n, w, \lambda)$ of codewords

Ford Motor Co.
Exhibit 1030
Page 235

in a (n, w, λ) OOC satisfies

$$M(n,w,\lambda) \leq \frac{1}{w}\left\lfloor \frac{n-1}{w-1} \cdots \left\lfloor \frac{n-\lambda+1}{w-\lambda+1}\left\lfloor \frac{n-\lambda}{w-\lambda}\right\rfloor\right\rfloor \cdots \right\rfloor$$

An OOC code that achieves the Johnson bound is said to be optimal. A family $\{\mathcal{F}_n\}$ of OOCs indexed by the parameter n and arising from a common construction is said to be asymptotically optimum if

$$\lim_{n\to\infty} \frac{|\mathcal{F}_n|}{M(n,w,\lambda)} = 1$$

Constructions for optical orthogonal codes are available for the cases when $\lambda = 1$ and $\lambda = 2$. For larger values of λ, there exist constructions which are asymptotically optimum. Further details may be found in [6].

Defining Terms

Autocorrelation of a sequence: The complex inner product of the sequence with a shifted version itself.

Crosscorrelation of two sequences: The complex inner product of the first sequence with a shifted version of the second sequence.

***m* Sequence:** A periodic binary $\{0, 1\}$ sequence that is generated by a shift register with linear feedback and which has maximal possible period given the number of stages in the shift register.

Pseudonoise sequences: Also referred to as pseudorandom sequences (PN), these are sequences that are deterministically generated and yet possess some properties that one would expect to find in randomly generated sequences.

Shift-register sequence: A sequence with symbols drawn from a field, which satisfies a linear-recurrence relation and which can be implemented using a shift register.

References

[1] Barg, A. On small families of sequences with low periodic correlation, *Lecture Notes in Computer Science,* 781, 154–158, Berlin, Springer-Verlag, 1994.

[2] Baumert, L.D. *Cyclic Difference Sets,* Lecture Notes in Mathematics 182, Springer–Verlag, New York, 1971.

[3] Boztaş, S., Hammons, R., and Kumar, P.V. 4-phase sequences with near-optimum correlation properties, *IEEE Trans. Inform. Theory,* IT-38, 1101–1113, 1992.

[4] Golomb, S.W. *Shift Register Sequences,* Aegean Park Press, San Francisco, CA, 1982.

[5] Hammons, A.R., Jr. and Kumar, P.V. On a recent 4-phase sequence design for CDMA. *IEICE Trans. Commun.,* E76-B(8), 1993.

[6] Helleseth, T. and Kumar, P.V. (planned). Sequences with low correlation. In *Handbook of Coding Theory,* ed., V.S. Pless and W.C. Huffman, Elsevier Science Publishers, Amsterdam, 1998.

[7] Jensen, J.M., Jensen, H.E., and Høholdt, T. The merit factor of binary sequences related to difference sets. *IEEE Trans. Inform. Theory,* IT-37(May), 617–626, 1991.

[8] MacWilliams, F.J. and Sloane, N.J.A. *The Theory of Error-Correcting Codes,* North-Holland, Amsterdam, 1977.

[9] Maschietti, A. Difference sets and hyperovals, *Designs, Codes and Cryptography,* 14, 89–98, 1998.

Ford Motor Co.
Exhibit 1030
Page 236

[10] Mow, W.H. On McEliece's open problem on minimax aperiodic correlation. In *Proc. IEEE Intern. Symp. Inform. Theory,* 75, 1994.
[11] Nechaev, A. The Kerdock code in a cyclic form, *Discrete Math. Appl.,* 1, 365–384, 1991.
[12] Peterson, W.W. and Weldon, E.J., Jr. *Error-Correcting Codes,* 2nd ed. MIT Press, Cambridge, MA, 1972.
[13] Sarwate, D.V. An upper bound on the aperiodic autocorrelation function for a maximal-length sequence. *IEEE Trans. Inform. Theory,* IT-30(July), 685–687, 1984.
[14] Sarwate, D.V. and Pursley, M.B. Crosscorrelation properties of pseudorandom and related sequences. *Proc. IEEE,* 68(May), 593–619, 1980.
[15] Simon, M.K., Omura, J.K., Scholtz, R.A., and Levitt, B.K. *Spread Spectrum Communications Handbook,* revised ed., McGraw Hill, New York, 1994.
[16] Solé, P. A quaternary cyclic code and a family of quadriphase sequences with low correlation properties, *Coding Theory and Applications, Lecture Notes in Computer Science,* 388, 193–201, Berlin, Springer-Verlag, 1989.
[17] Udaya, P. and Siddiqi, M. Optimal biphase sequences with large linear complexity derived from sequences over Z_4, *IEEE Trans. Inform. Theory,* IT-42 (Jan), 206–216, 1996.

Further Information

A more in-depth treatment of pseudonoise sequences, may be found in the following.

[1] Golomb, S.W. *Shift Register Sequences,* Aegean Park Press, San Francisco, 1982.
[2] Helleseth, T. and Kumar, P.V. Sequences with Low Correlation, in *Handbook of Coding Theory,* edited by V.S. Pless and W.C. Huffman, Elsevier Science Publishers, Amsterdam, 1998 (planned).
[3] Sarwate, D.V. and Pursley, M.B. Crosscorrelation Properties of Pseudorandom and Related Sequences, *Proc. IEEE,* 68, May, 593–619, 1980.
[4] Simon, M.K., Omura, J.K., Scholtz, R.A., and Levitt, B.K. *Spread Spectrum Communications Handbook,* revised ed., McGraw Hill, New York, 1994.

Ford Motor Co.
Exhibit 1030
Page 237

9

Optimum Receivers

9.1 Introduction 9-1
9.2 Preliminaries 9-2
9.3 Karhunen–Loève Expansion 9-2
9.4 Detection Theory 9-4
9.5 Performance 9-5
9.6 Signal Space 9-6
9.7 Standard Binary Signalling Schemes 9-7
9.8 M-ary Optimal Receivers 9-8
9.9 More Realistic Channels 9-9
Random Phase Channels • Rayleigh Channel
9.10 Dispersive Channels 9-13
Defining Terms 9-14
References 9-14
Further Information 9-15

Geoffrey C. Orsak
Southern Methodist University

9.1 Introduction

Every engineer strives for optimality in design. This is particularly true for communications engineers since in many cases implementing suboptimal receivers and sources can result in dramatic losses in performance. As such, this chapter focuses on design principles leading to the implementation of optimum receivers for the most common communication environments.

The main objective in digital communications is to transmit a sequence of bits to a remote location with the highest degree of accuracy. This is accomplished by first representing bits (or more generally short bit sequences) by distinct waveforms of finite time duration. These time-limited waveforms are then transmitted (broadcasted) to the remote sites in accordance with the data sequence.

Unfortunately, because of the nature of the **communication channel,** the remote location receives a corrupted version of the concatenated signal waveforms. The most widely accepted model for the communication channel is the so-called **additive white Gaussian noise**[1] **channel (AWGN channel).** Mathematical arguments based upon the central

[1]For those unfamiliar with AWGN, a random process (waveform) is formally said to be white Gaussian noise if all collections of instantaneous observations of the process are jointly Gaussian and mutually independent. An important consequence of this property is that the power spectral density of the process

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 238

limit theorem [7], together with supporting empirical evidence, demonstrate that many common communication channels are accurately modeled by this abstraction. Moreover, from the design perspective, this is quite fortuitous since design and analysis with respect to this channel model is relatively straightforward.

9.2 Preliminaries

To better describe the digital communications process, we shall first elaborate on so-called binary communications. In this case, when the source wishes to transmit a bit value of 0, the transmitter broadcasts a specified waveform $s_0(t)$ over the **bit interval** $t \in [0, T]$. Conversely, if the source seeks to transmit the bit value of 1, the transmitter alternatively broadcasts the signal $s_1(t)$ over the same bit interval. The received waveform $R(t)$ corresponding to the first bit is then appropriately described by the following hypotheses testing problem:

$$\begin{aligned} H_0 &: R(t) = s_0(t) + \eta(t) \qquad 0 \le t \le T \\ H_1 &: R(t) = s_1(t) + \eta(t) \end{aligned} \tag{9.1}$$

where, as stated previously, $\eta(t)$ corresponds to AWGN with spectral height nominally given by $N_0/2$. It is the objective of the receiver to determine the bit value, i.e., the most accurate hypothesis from the received waveform $R(t)$.

The optimality criterion of choice in digital communication applications is the **total probability of error** normally denoted as P_e. This scalar quantity is expressed as

$$\begin{aligned} P_e \;=\; & Pr(\text{ declaring } 1 \,|\, 0 \text{ transmitted}) Pr(0 \text{ transmitted}) \\ & + Pr(\text{ declaring } 0 \,|\, 1 \text{ transmitted}) Pr(1 \text{ transmitted}) \end{aligned} \tag{9.2}$$

The problem of determining the optimal binary receiver with respect to the probability of error is solved by applying stochastic representation theory [10] to detection theory [5, 9]. The specific waveform representation of relevance in this application is the **Karhunen–Loève (KL) expansion.**

9.3 Karhunen–Loève Expansion

The Karhunen–Loève expansion is a generalization of the Fourier series designed to represent a random process in terms of deterministic basis functions and uncorrelated random variables derived from the process. Whereas the Fourier series allows one to model or represent deterministic time-limited energy signals in terms of linear combinations of complex exponential waveforms, the Karhunen–Loève expansion allows us to represent a second-order random process in terms of a set of **orthonormal** basis functions scaled by a sequence of random variables. The objective in this representation is to choose the basis of time functions so that the coefficients in the expansion are mutually uncorrelated random variables.

To be more precise, if $R(t)$ is a zero mean second-order random process defined over $[0, T]$ with covariance function $K_R(t, s)$, then so long as the basis of deterministic functions satisfy

is a constant with respect to frequency variation (spectrally flat). For more on AWGN, see Papoulis [4].

Ford Motor Co.
Exhibit 1030
Page 239

certain integral constraints [9], one may write $R(t)$ as

$$R(t) = \sum_{i=1}^{\infty} R_i \phi_i(t) \qquad 0 \le t \le T \tag{9.3}$$

where

$$R_i = \int_0^T R(t)\phi_i(t)\ \mathrm{d}t$$

In this case the R_i will be mutually uncorrelated random variables with the ϕ_i being deterministic basis functions that are complete in the space of square integrable time functions over $[0, T]$. Importantly, in this case, equality is to be interpreted as **mean-square equivalence,** i.e.,

$$\lim_{N \to \infty} E\left[\left(R(t) - \sum_{i=1}^{N} R_i \phi_i(t)\right)^2\right] = 0$$

for all $0 \le t \le T$.

FACT 9.1 *If $R(t)$ is AWGN, then any basis of the vector space of square integrable signals over $[0, T]$ results in uncorrelated and therefore independent Gaussian random variables.*

The use of Fact 9.1 allows for a conversion of a continuous time detection problem into a finite-dimensional detection problem. Proceeding, to derive the optimal binary receiver, we first construct our set of basis functions as the set of functions defined over $t \in [0, T]$ beginning with the signals of interest $s_0(t)$ and $s_1(t)$. That is,

$$\{s_0(t), s_1(t),\ \text{plus a countable number of functions which complete the basis}\}$$

In order to insure that the basis is orthonormal, we must apply the Gramm–Schmidt procedure[2] [6] to the full set of functions beginning with $s_0(t)$ and $s_1(t)$ to arrive at our final choice of basis $\{\phi_i(t)\}$.

FACT 9.2 *Let $\{\phi_i(t)\}$ be the resultant set of basis functions.*
Then for all $i > 2$, the $\phi_i(t)$ are orthogonal to $s_0(t)$ and $s_1(t)$. That is,

$$\int_0^T \phi_i(t) s_j(t)\, \mathrm{d}t = 0$$

for all $i > 2$ and $j = 0, 1$.

Using this fact in conjunction with Eq. (9.3), one may recognize that only the coefficients R_1 and R_2 are functions of our signals of interest. Moreover, since the R_i are mutually independent, the optimal receiver will, therefore, only be a function of these two values.

[2]The Gramm-Schmidt procedure is a deterministic algorithm that simply converts an arbitrary set of basis functions (vectors) into an equivalent set of orthonormal basis functions (vectors).

Ford Motor Co.
Exhibit 1030
Page 240

Thus, through the application of the KL expansion, we arrive at an equivalent hypothesis testing problem to that given in Eq. (9.1),

$$\begin{aligned} H_0 : \boldsymbol{R} &= \begin{bmatrix} \int_0^T \phi_1(t)s_0(t)\ \mathrm{d}t \\ \int_0^T \phi_2(t)s_0(t)\ \mathrm{d}t \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \\ H_1 : \boldsymbol{R} &= \begin{bmatrix} \int_0^T \phi_1(t)s_1(t)\ \mathrm{d}t \\ \int_0^T \phi_2(t)s_1(t)\ \mathrm{d}t \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \end{aligned} \tag{9.4}$$

where it is easily shown that η_1 and η_2 are mutually independent, zero-mean, Gaussian random variables with variance given by $N_0/2$, and where ϕ_1 and ϕ_2 are the first two functions from our orthonormal set of basis functions. Thus, the design of the optimal binary receiver reduces to a simple two-dimensional detection problem that is readily solved through the application of detection theory.

9.4 Detection Theory

It is well known from detection theory [5] that under the minimum P_e criterion, the optimal detector is given by the *maximum a posteriori rule* (*MAP*),

$$\text{choose}_i \ \text{ largest } p_{H_i|\boldsymbol{R}}(H_i \mid \boldsymbol{R} = \boldsymbol{r}) \tag{9.5}$$

i.e., determine the hypothesis that is most likely, given that our observation vector is $\boldsymbol{r}$. By a simple application of Bayes theorem [4], we immediately arrive at the central result in detection theory: the optimal binary detector is given by the likelihood ratio test (LRT),

$$L(\boldsymbol{R}) = \frac{p_{\boldsymbol{R}|H_1}(\boldsymbol{R})}{p_{\boldsymbol{R}|H_0}(\boldsymbol{R})} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\pi_0}{\pi_1} \tag{9.6}$$

where the π_i are the a priori probabilities of the hypotheses H_i being true. Since in this case we have assumed that the noise is white and Gaussian, the LRT can be written as

$$L(\boldsymbol{R}) = \frac{\prod_1^2 \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{1}{2} \frac{(R_i - s_{1,i})^2}{N_0/2} \right)}{\prod_1^2 \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{1}{2} \frac{(R_i - s_{0,i})^2}{N_0/2} \right)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\pi_0}{\pi_1} \tag{9.7}$$

where

$$s_{j,i} = \int_0^T \phi_i(t)s_j(t)\,\mathrm{d}t$$

By taking the logarithm and cancelling common terms, it is easily shown that the optimum binary receiver can be written as

$$\frac{2}{N_0} \sum_1^2 \boldsymbol{R}_i \left(s_{1,i} - s_{0,i}\right) - \frac{1}{N_0} \sum_1^2 \left(s_{1,i}^2 - s_{0,i}^2\right) \underset{H_0}{\overset{H_1}{\gtrless}} \ln \frac{\pi_0}{\pi_1} \tag{9.8}$$

Ford Motor Co.
Exhibit 1030
Page 241

This finite-dimensional version of the optimal receiver can be converted back into a continuous time receiver by the direct application of Parseval's theorem [4] where it is easily shown that

$$\begin{aligned} \sum_{i=1}^{2} R_i s_{k,i} &= \int_0^T R(t) s_k(t)\, \mathrm{d}t \\ \sum_{i=1}^{2} s_{k,i}^2 &= \int_0^T s_k^2(t)\, \mathrm{d}t \end{aligned} \tag{9.9}$$

By applying Eq. (9.9) to Eq. (9.8) the final receiver structure is then given by

$$\int_0^T R(t)\left[s_1(t) - s_0(t)\right] \mathrm{d}t - \frac{1}{2}\left(E_1 - E_0\right) \underset{H_0}{\overset{H_1}{\gtrless}} \frac{N_0}{2} \ln \frac{\pi_0}{\pi_1} \tag{9.10}$$

where E_1 and E_0 are the energies of signals $s_1(t)$ and $s_0(t)$, respectively. (See Fig. 9.1 for a block diagram.) Importantly, if the signals are equally likely ($\pi_0 = \pi_1$), the optimal receiver is independent of the typically unknown spectral height of the background noise.

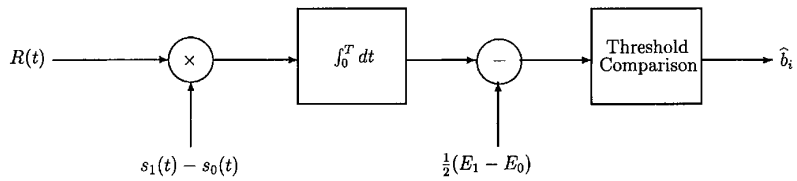


FIGURE 9.1 Optimal correlation receiver structure for binary communications.

One can readily observe that the optimal binary communication receiver correlates the received waveform with the difference signal $s_1(t) - s_0(t)$ and then compares the statistic to a threshold. This operation can be interpreted as identifying the signal waveform $s_i(t)$ that best correlates with the received signal $R(t)$. Based on this interpretation, the receiver is often referred to as the **correlation receiver.**

As an alternate means of implementing the correlation receiver, we may reformulate the computation of the left-hand side of Eq. (9.10) in terms of standard concepts in filtering. Let $h(t)$ be the impulse response of a linear, time-invariant (LTI) system. By letting $h(t) = s_1(T-t) - s_0(T-t)$, then it is easily verified that the output of $R(t)$ to a LTI system with impulse response given by $h(t)$ and then sampled at time $t = T$ gives the desired result. (See Fig. 9.2 for a block diagram.) Since the impulse response is matched to the signal waveforms, this implementation is often referred to as the **matched filter receiver.**

9.5 Performance

Because of the nature of the statistics of the channel and the relative simplicity of the receiver, performance analysis of the optimal binary receiver in AWGN is a straightforward task. Since the conditional statistics of the log likelihood ratio are Gaussian random vari-

Ford Motor Co.
Exhibit 1030
Page 242

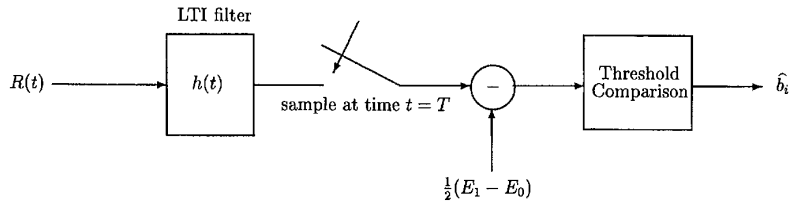


FIGURE 9.2 Optimal matched filter receiver structure for binary communications. In this case $h(t) = s_1(T - t) - s_0(t - t)$.

ables, the probability of error can be computed directly in terms of Marcum Q functions[3] as

$$P_e = Q\left(\frac{\|\boldsymbol{s}_0 - \boldsymbol{s}_1\|}{\sqrt{2N_0}}\right)$$

where the $\boldsymbol{s}_i$ are the two-dimensional signal vectors obtained from Eq. (9.4), and where $\|\boldsymbol{x}\|$ denotes the Euclidean length of the vector $\boldsymbol{x}$. Thus, $\|\boldsymbol{s}_0 - \boldsymbol{s}_1\|$ is best interpreted as the distance between the respective signal representations. Since the Q function is monotonically decreasing with an increasing argument, one may recognize that the probability of error for the optimal receiver decreases with an increasing separation between the signal representations, i.e., the more dissimilar the signals, the lower the P_e.

9.6 Signal Space

The concept of a **signal space** allows one to view the signal classification problem (receiver design) within a geometrical framework. This offers two primary benefits: first it supplies an often more intuitive perspective on the receiver characteristics (e.g., performance) and second it allows for a straightforward generalization to standard M-ary signalling schemes.

To demonstrate this, in Fig. 9.3, we have plotted an arbitrary signal space for the binary signal classification problem. The axes are given in terms of the basis functions $\phi_1(t)$ and $\phi_2(t)$. Thus, every point in the signal space is a time function constructed as a linear combination of the two basis functions. By Fact 9.2, we recall that both signals $s_0(t)$ and $s_1(t)$ can be constructed as a linear combination of $\phi_1(t)$ and $\phi_2(t)$ and as such we may identify these two signals in this figure as two points.

Since the decision statistic given in Eq. (9.8) is a linear function of the observed vector $\boldsymbol{R}$ which is also located in the signal space, it is easily shown that the set of vectors under which the receiver declares hypothesis H_i is bounded by a line in the signal space. This so-called **decision boundary** is obtained by solving the equation $\ln[L(\boldsymbol{R})] = 0$. (Here again we have assumed equally likely hypotheses.) In the case under current discussion, this decision boundary is simply the hyperplane separating the two signals in signal space. Because of the generality of this formulation, many problems in communication system design are best cast in terms of the signal space, that is, signal locations and decision boundaries.

[3]The Q function is the probability that a standard normal random variable exceeds a specified constant, i.e., $Q(x) = \int_x^\infty 1/\sqrt{2\pi}\exp(-z^2/2)\,\mathrm{d}z$.

Ford Motor Co.
Exhibit 1030
Page 243

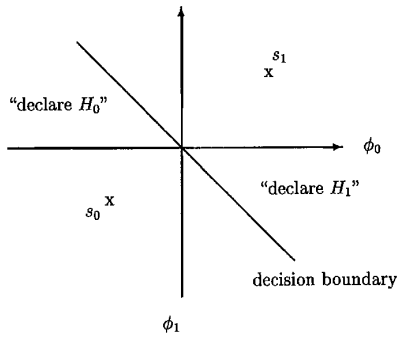


FIGURE 9.3 Signal space and decision boundary for optimal binary receiver.

9.7 Standard Binary Signalling Schemes

The framework just described allows us to readily analyze the most popular signalling schemes in binary communications: amplitude-shift keying (ASK), frequency-shift keying (FSK), and phase-shift keying (PSK). Each of these examples simply constitute a different selection for signals $s_0(t)$ and $s_1(t)$.

In the case of ASK, $s_0(t) = 0$, while $s_1(t) = \sqrt{2E/T}\sin(2\pi f_c t)$, where E denotes the energy of the waveform and f_c denotes the frequency of the carrier wave with $f_c T$ being an integer. Because $s_0(t)$ is the null signal, the signal space is a one-dimensional vector space with $\phi_1(t) = \sqrt{2/T}\sin(2\pi f_c t)$. This, in turn, implies that $\|s_0 - s_1\| = \sqrt{E}$. Thus, the corresponding probability of error for ASK is

$$P_e(\text{ ASK}) = Q\left(\sqrt{\frac{E}{2N_0}}\right)$$

For FSK, the signals are given by equal amplitude sinusoids with distinct center frequencies, that is, $s_i(t) = \sqrt{2E/T}\sin(2\pi f_i t)$ with $f_i T$ being two distinct integers. In this case, it is easily verified that the signal space is a two-dimensional vector space with $\phi_i(t) = \sqrt{2/T}\sin(2\pi f_i t)$ resulting in $\|\boldsymbol{s}_0 - \boldsymbol{s}_1\| = \sqrt{2\boldsymbol{E}}$. The corresponding error rate is given to be

$$P_e(\text{FSK}) = Q\left(\sqrt{\frac{E}{N_0}}\right)$$

Finally, with regard to PSK signalling, the most frequently utilized binary PSK signal set is an example of an antipodal signal set. Specifically, the antipodal signal set results in the greatest separation between the signals in the signal space subject to an energy constraint on both signals. This, in turn, translates into the energy constrained signal set with the minimum P_e. In this case, the $s_i(t)$ are typically given by $\sqrt{2E/T}\sin[2\pi f_c t + \theta(i)]$, where $\theta(0) = 0$ and $\theta(1) = \pi$. As in the ASK case, this results in a one-dimensional signal space, however, in this case $\|\boldsymbol{s}_0 - \boldsymbol{s}_1\| = 2\sqrt{\boldsymbol{E}}$ resulting in probability of error given by

$$P_e(\text{PSK}) = Q\left(\sqrt{\frac{2E}{N_0}}\right)$$

In all three of the described cases, one can readily observe that the resulting performance is a function of only the signal-to-noise ratio E/N_0. In the more general case, the performance

Ford Motor Co.
Exhibit 1030
Page 244

will be a function of the intersignal energy to noise ratio. To gauge the relative difference in performance of the three signalling schemes, in Fig. 9.4, we have plotted the P_e as a function of the SNR. Please note the large variation in performance between the three schemes for even moderate values of SNR.

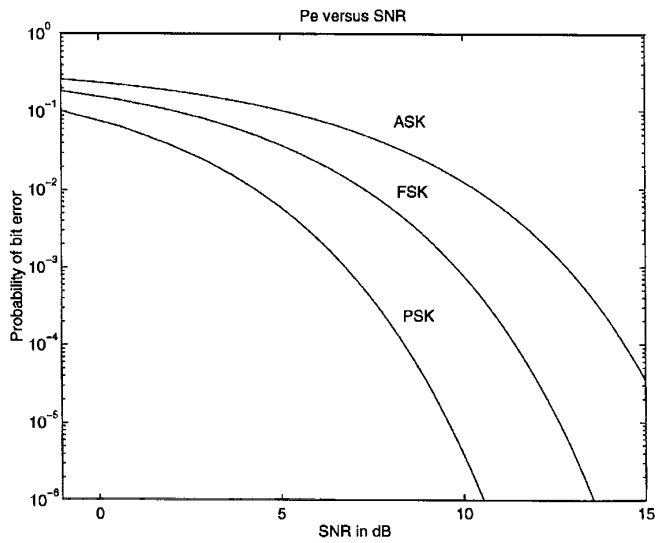


FIGURE 9.4 P_e vs. the signal to noise ratio in decibels [dB $= 10\log(E/N_0)$] for amplitude-shift keying, frequency-shift keying, and phase-shift keying; note that there is a 3-dB difference in performance from ASK to FSK to PSK.

9.8 *M*-ary Optimal Receivers

In binary signalling schemes, one seeks to transmit a single bit over the bit interval $[0, T]$. This is to be contrasted with M-ary signalling schemes where one transmits multiple bits simultaneously over the so-called symbol interval $[0, T]$. For example, using a signal set with 16 separate waveforms will allow one to transmit a length four-bit sequence per symbol (waveform). Examples of M-ary waveforms are quadrature phase-shift keying (QPSK) and quadrature amplitude modulation (QAM).

The derivation of the optimum receiver structure for M-ary signalling requires the straightforward application of fundamental results in detection theory. As with binary signalling, the Karhunen–Loève expansion is the mechanism utilized to convert a hypotheses testing problem based on continuous waveforms into a vector classification problem. Depending on the complexity of the M waveforms, the signal space can be as large as an M-dimensional vector space.

By extending results from the binary signalling case, it is easily shown that the optimum M-ary receiver computes

$$\xi_i[R(t)] = \int_0^T s_i(t)R(t)\ \mathrm{d}t - \frac{E_i}{2} + \frac{N_0}{2}\ln \pi_i \qquad i = 1, \ldots, M$$

Ford Motor Co.
Exhibit 1030
Page 245

where, as before, the $s_i(t)$ constitute the signal set with the π_i being the corresponding a priori probabilities. After computing M separate values of ξ_i, the minimum probability of error receiver simply chooses the largest amongst this set. Thus, the M-ary receiver is implemented with a bank of correlation or matched filters followed by choose-largest decision logic.

In many cases of practical importance, the signal sets are selected so that the resulting signal space is a two-dimensional vector space irrespective of the number of signals. This simplifies the receiver structure in that the sufficient statistics are obtained by implementing only two matched filters. Both QPSK and QAM signal sets fit into this category. As an example, in Fig. 9.5, we have depicted the signal locations for standard 16-QAM signalling with the associated decision boundaries. In this case we have assumed an equally likely signal set. As can be seen, the optimal decision rule selects the signal representation that is closest to the received signal representation in this two-dimensional signal space.

9.9 More Realistic Channels

As is unfortunately often the case, many channels of practical interest are not accurately modeled as simply an AWGN channel. It is often that these channels impose nonlinear effects on the transmitted signals. The best example of this are channels that impose a random phase and random amplitude onto the signal. This typically occurs in applications such as in mobile communications, where one often experiences rapidly changing path lengths from source to receiver.

Fortunately, by the judicious choice of signal waveforms, it can be shown that the selection of the ϕ_i in the Karhunen–Loève transformation is often independent of these unwanted parameters. In these situations, the random amplitude serves only to scale the signals in signal space, whereas the random phase simply imposes a rotation on the signals in signal space.

Since the Karhunen–Loève basis functions typically do not depend on the unknown parameters, we may again convert the continuous time classification problem to a vector channel problem where the received vector $\boldsymbol{R}$ is computed as in Eq. (9.3). Since this vector is a function of both the unknown parameters (i.e., in this case amplitude A and phase ν), to obtain a likelihood ratio test independent of A and ν, we simply apply Bayes theorem to obtain the following form for the LRT:

$$L(\boldsymbol{R}) = \frac{E\left[p_{\boldsymbol{R}|H_1,A,\nu}\left(\boldsymbol{R} \mid H_1, A, \nu\right)\right]}{E\left[p_{\boldsymbol{R}|H_0,A,\nu}\left(\boldsymbol{R} \mid H_0, A, \nu\right)\right]} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \frac{\pi_0}{\pi_1}$$

where the expectations are taken with respect to A and ν, and where $p_{\boldsymbol{R}|H_i,A,\nu}$ are the conditional probability density functions of the signal representations. Assuming that the background noise is AWGN, it can be shown that the LRT simplifies to choosing the largest amongst

$$\xi_i[R(t)] = \pi_i \int_{A,\nu} \exp\left\{ \frac{2}{N_0} \int_0^T R(t) s_i(t \mid A, \nu)\ \mathrm{d}t - \frac{E_i(A,\nu)}{N_0} \right\} p_{A,\nu}(A,\nu)\ \mathrm{d}A\,\mathrm{d}\nu$$

$$i = 1, \ldots, M \qquad (9.11)$$

It should be noted that in the Eq. (9.11) we have explicitly shown the dependence of the transmitted signals s_i on the parameters A and ν. The final receiver structures, together

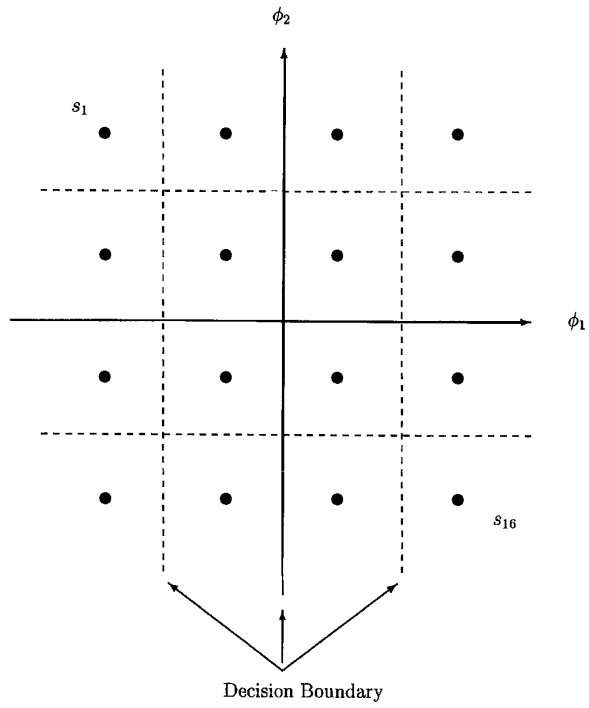


FIGURE 9.5 Signal space representation of 16-QAM signal set. Optimal decision regions for equally likely signals are also noted.

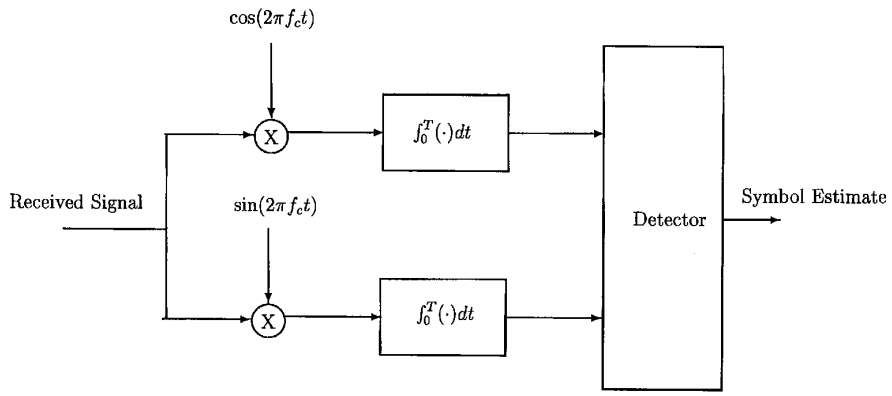


FIGURE 9.6 Optimum receiver structure for noncoherent (random or unknown phase) ASK demodulation.

Ford Motor Co.
Exhibit 1030
Page 247

with their corresponding performance are, thus, a function of both the choice of signal sets and the probability density functions of the random amplitude and random phase.

9.9.1 Random Phase Channels

If we consider first the special case where the channel simply imposes a uniform random phase on the signal, then it can be easily shown that the so-called in-phase and quadrature statistics obtained from the received signal $R(t)$ (denoted by R_I and R_Q, respectively), are sufficient statistics for the signal classification problem. These quantities are computed as

$$R_I(i) = \int_0^T R(t) \cos\left[2\pi f_c(i)t\right] \; \mathrm{d}t$$

and

$$R_Q(i) = \int_0^T R(t) \sin\left[2\pi f_c(i)t\right] \; \mathrm{d}t$$

where in this case the index i corresponds to the center frequencies of hypotheses H_i, (e.g., FSK signalling). The optimum binary receiver selects the largest from amongst

$$\xi_i[R(t)] = \pi_i \exp\left(-\frac{E_i}{N_0}\right) I_0\left[\frac{2}{N_0}\sqrt{R_I^2(i) + R_Q^2(i)}\right] \qquad i = 1, \ldots, M$$

where I_0 is a zeroth-order, modified Bessel function of the first kind. If the signals have equal energy and are equally likely (e.g., FSK signalling), then the optimum receiver is given by

$$R_I^2(1) + R_Q^2(1) \underset{H_0}{\overset{H_1}{\gtrless}} R_I^2(0) + R_Q^2(0)$$

One may readily observe that the optimum receiver bases its decision on the values of the two envelopes of the received signal $\sqrt{R_I^2(i) + R_Q^2(i)}$ and, as a consequence, is often referred to as an envelope or square-law detector. Moreover, it should be observed that the computation of the envelope is independent of the underlying phase of the signal and is as such known as a noncoherent receiver.

The computation of the error rate for this detector is a relatively straightforward exercise resulting in

$$P_e(\text{ noncoherent}) = \frac{1}{2} \exp\left(-\frac{E}{2N_0}\right)$$

As before, note that the error rate for the noncoherent receiver is simply a function of the SNR.

9.9.2 Rayleigh Channel

As an important generalization of the described random phase channel, many communication systems are designed under the assumption that the channel introduces both a random amplitude and a random phase on the signal. Specifically, if the original signal sets are of the form $s_i(t) = m_i(t)\cos(2\pi f_c t)$ where $m_i(t)$ is the baseband version of the message (i.e., what distinguishes one signal from another), then the so-called **Rayleigh channel** introduces random distortion in the received signal of the following form:

$$s_i(t) = A m_i(t) \cos\left(2\pi f_c t + \nu\right)$$

where the amplitude A is a Rayleigh random variable[4] and where the random phase ν is a uniformly distributed between zero and 2π.

To determine the optimal receiver under this distortion, we must first construct an alternate statistical model for $s_i(t)$. To begin, it can be shown from the theory of random variables [4] that if X_I and X_Q are statistically independent, zero mean, Gaussian random variables with variance given by σ^2, then

$$Am_i(t)\cos(2\pi f_c t + \nu) = m_i(t)X_I\cos(2\pi f_c t) + m_i(t)X_Q\sin(2\pi f_c t)$$

Equality here is to be interpreted as implying that both A and ν will be the appropriate random variables. From this, we deduce that the combined uncertainty in the amplitude and phase of the signal is incorporated into the Gaussian random variables X_I and X_Q. The in-phase and quadrature components of the signal $s_i(t)$ are given by $s_{Ii}(t) = m_i(t)\cos(2\pi f_c t)$ and $s_{Q_i}(t) = m_i(t)\sin(2\pi f_c t)$, respectively. By appealing to Eq. (9.11), it can be shown that the optimum receiver selects the largest from

$$\xi_i[R(t)] = \frac{\pi_i}{1+\frac{2E_i}{N_0}\sigma^2}\exp\left[\frac{\sigma^2}{\frac{1}{2}+\frac{E_i}{N_0}\sigma^2}\left(\langle R(t), s_{Ii}(t)\rangle^2 + \langle R(t), s_{Q_i}(t)\rangle^2\right)\right]$$

where the inner product

$$\langle R(t), S_i(t)\rangle = \int_0^T R(t)s_i(t)\ \mathrm{d}t$$

Further, if we impose the conditions that the signals be equally likely with equal energy over the symbol interval, then optimum receiver selects the largest amongst

$$\xi_i[R(t)] = \sqrt{\langle R(t), s_{Ii}(t)\rangle^2 + \langle R(t), s_{Q_i}(t)\rangle^2}$$

Thus, much like for the random phase channel, the optimum receiver for the Rayleigh channel computes the projection of the received waveform onto the in-phase and quadrature components of the hypothetical signals. From a signal space perspective, this is akin to computing the length of the received vector in the subspace spanned by the hypothetical signal. The optimum receiver then chooses the largest amongst these lengths.

As with the random phase channel, computing the performance is a straightforward task resulting in (for the equally likely, equal energy case)

$$P_e(\text{ Rayleigh}) = \frac{\frac{1}{2}}{\left(1+\frac{E\sigma^2}{N_0}\right)}$$

Interestingly, in this case the performance depends not only on the SNR, but also on the variance (spread) of the Rayleigh amplitude A. Thus, if the amplitude spread is large, we expect to often experience what is known as deep fades in the amplitude of the received waveform and as such expect a commensurate loss in performance.

[4]The density of a Rayleigh random variable is given by $p_A(a) = a/\sigma^2\exp(-a^2/2\sigma^2)$ for $a \geq 0$.

Ford Motor Co.
Exhibit 1030
Page 249

9.10 Dispersive Channels

The **dispersive channel** model assumes that the channel not only introduces AWGN but also distorts the signal through a filtering process. This model incorporates physical realities such as multipath effects and frequency selective fading. In particular, the standard model adopted is depicted in the block diagram given in Fig. 9.7. As can be seen, the receiver observes a filtered version of the signal plus AWGN. If the impulse response of the channel is known, then we arrive at the optimum receiver design by applying the previously presented theory. Unfortunately, the duration of the filtered signal can be a complicating factor. More often than not, the channel will increase the duration of the transmitted signals, hence, leading to the description, dispersive channel.

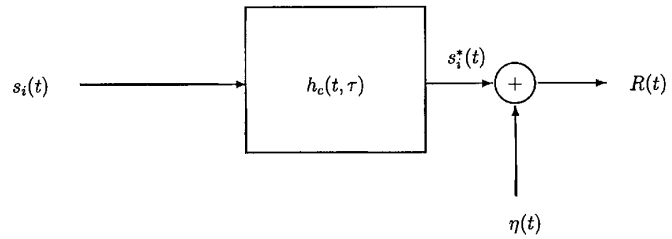


FIGURE 9.7 Standard model for dispersive channel. The time varying impulse response of the channel is denoted by $h_c(t,\tau)$.

However, if the designers take this into account by shortening the duration of $s_i(t)$ so that the duration of $s_i^*(t)$ is less than T, then the optimum receiver chooses the largest amongst

$$\xi_i(R(t)) = \frac{N_0}{2}\ln\pi_i + \langle R(t), s_i^*(t)\rangle - \frac{1}{2}E_i^*$$

If we limit our consideration to equally likely binary signal sets, then the minimum P_e matches the received waveform to the filtered versions of the signal waveforms. The resulting error rate is given by

$$P_e(\text{ dispersive}) = Q\left(\frac{\|\boldsymbol{s}_0^* - \boldsymbol{s}_0^*\|}{\sqrt{2\boldsymbol{N}_0}}\right)$$

Thus, in this case the minimum P_e is a function of the separation of the filtered version of the signals in the signal space.

The problem becomes substantially more complex if we cannot insure that the filtered signal durations are less than the symbol lengths. In this case we experience what is known as **intersymbol interference (ISI)**. That is, observations over one symbol interval contain not only the symbol information of interest but also information from previous symbols. In this case we must appeal to optimum sequence estimation [5] to take full advantage of the information in the waveform. The basis for this procedure is the maximization of the joint likelihood function conditioned on the sequence of symbols. This procedure not only defines the structure of the optimum receiver under ISI but also is critical in the decoding of convolutional codes and coded modulation. Alternate adaptive techniques to solve this problem involve the use of channel equalization.

Ford Motor Co.
Exhibit 1030
Page 250

Defining Terms

Additive white Gaussian noise (AWGN) channel: The channel whose model is that of corrupting a transmitted waveform by the addition of white (i.e., spectrally flat) Gaussian noise.

Bit (symbol) interval: The period of time over which a single symbol is transmitted.

Communication channel: The medium over which communication signals are transmitted. Examples are fiber optic cables, free space, or telephone lines.

Correlation or matched filter receiver: The optimal receiver structure for digital communications in AWGN.

Decision boundary: The boundary in signal space between the various regions where the receiver declares H_i. Typically a hyperplane when dealing with AWGN channels.

Dispersive channel: A channel that elongates and distorts the transmitted signal. Normally modeled as a time-varying linear system.

Intersymbol interference: The ill-effect of one symbol smearing into adjacent symbols thus interfering with the detection process. This is a consequence of the channel filtering the transmitted signals and therefore elongating their duration, see dispersive channel.

Karhunen–Loève expansion: A representation for second-order random processes. Allows one to express a random process in terms of a superposition of deterministic waveforms. The scale values are uncorrelated random variables obtained from the waveform.

Mean-square equivalence: Two random vectors or time-limited waveforms are mean-square equivalent if and only if the expected value of their mean-square error is zero.

Orthonormal: The property of two or more vectors or time-limited waveforms being mutually orthogonal and individually having unit length. Orthogonality and length are typically measured by the standard Euclidean inner product.

Rayleigh channel: A channel that randomly scales the transmitted waveform by a Rayleigh random variable while adding an independent uniform phase to the carrier.

Signal space: An abstraction for representing a time limited waveform in a low-dimensional vector space. Usually arrived at through the application of the Karhunen–Loève transformation.

Total probability of error: The probability of classifying the received waveform into any of the symbols that were not transmitted over a particular bit interval.

References

[1] Gibson, J.D., *Principles of Digital and Analog Communications,* 2nd ed., MacMillan, New York, 1993.

[2] Haykin, S., *Communication Systems,* 3rd ed., John Wiley & Sons, New York, 1994.

[3] Lee, E.A. and Messerschmitt, D.G., *Digital Communication,* Kluwer Academic Publishers, Norwell, MA, 1988.

Ford Motor Co.
Exhibit 1030
Page 251

[4] Papoulis, A., *Probability, Random Variables, and Stochastic Processes,* 3rd ed., McGraw-Hill, New York, 1991.
[5] Poor, H.V., *An Introduction to Signal Detection and Estimation,* Springer-Verlag, New York, 1988.
[6] Proakis, J.G., *Digital Communications,* 2nd ed., McGraw-Hill, New York, 1989.
[7] Shiryayev, A.N., *Probability,* Springer-Verlag, New York, 1984.
[8] Sklar, B., *Digital Communications, Fundamentals and Applications,* Prentice Hall, Englewood Cliffs, NJ, 1988.
[9] Van Trees, H.L., *Detection, Estimation, and Modulation Theory, Part I,* John Wiley & Sons, New York, 1968.
[10] Wong, E. and Hajek, B., *Stochastic Processes in Engineering Systems,* Springer-Verlag, New York, 1985.
[11] Wozencraft, J.M. and Jacobs, I., *Principles of Communication Engineering,* reissue, Waveland Press, Prospect Heights, Illinois, 1990.
[12] Ziemer, R.E. and Peterson, R.L., *Introduction to Digital Communication,* Macmillan, New York, 1992.

Further Information

The fundamentals of receiver design were put in place by Wozencraft and Jacobs in their seminal book. Since that time, there have been many outstanding textbooks in this area. For a sampling see [1, 2, 3, 8, 12]. For a complete treatment on the use and application of detection theory in communications see [5, 9]. For deeper insights into the Karhunen–Loève expansion and its use in communications and signal processing see [10].

Ford Motor Co.
Exhibit 1030
Page 252

10

Forward Error Correction Coding

10.1 Introduction 10-1
10.2 Fundamentals of Block Coding 10-3
10.3 Structure and Decoding of Block Codes 10-3
10.4 Important Classes of Block Codes 10-5
10.5 Principles of Convolutional Coding 10-6
10.6 Decoding of Convolutional Codes 10-10
10.7 Trellis-Coded Modulation 10-10
10.8 Additional Measures 10-13
10.9 Turbo Codes 10-13
10.10 Applications 10-14
Defining Terms 10-15
References 10-17
Further Information 10-17

V.K. Bhargava
University of Victoria

I.J. Fair
University of Alberta

10.1 Introduction

In 1948, Claude Shannon issued a challenge to communications engineers by proving that communication systems could be made arbitrarily reliable as long as a fixed percentage of the transmitted signal was redundant [9]. He showed that limits exist only on the rate of communication and not its accuracy, and went on to prove that errorless transmission could be achieved in an additive white Gaussian noise (AWGN) environment with infinite bandwidth if the ratio of energy per data bit to noise power spectral density exceeds the **Shannon Limit**. He did not, however, indicate how this could be achieved. Subsequent research has led to a number of techniques that introduce redundancy to allow for correction of errors without retransmission. These techniques, collectively known as forward error correction (FEC) coding techniques, are used in systems where a reverse channel is not available for requesting retransmission, the delay with retransmission would be excessive, the expected number of errors would require a large number of retransmissions, or retransmission would be awkward to implement [10].

A simplified model of a digital communication system which incorporates FEC coding is shown in Fig. 10.1. The FEC code acts on a **discrete data channel** comprising all system elements between the encoder output and decoder input. The encoder maps the source data to q-ary code symbols which are modulated and transmitted. During transmission, this signal can be corrupted, causing errors to arise in the demodulated symbol sequence. The FEC decoder attempts to correct these errors and restore the original source data.

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 253

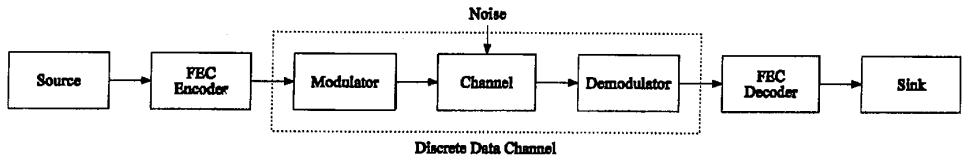


FIGURE 10.1 Block diagram of a digital communication system with forward error correction.

A demodulator which outputs only a value for the q-ary symbol received during each symbol interval is said to make **hard decisions.** In the **binary symmetric channel** (BSC), hard decisions are made on binary symbols and the probability of error is independent of the value of the symbol. One example of a BSC is the coherently demodulated binary phase-shift-keyed (BPSK) signal corrupted by AWGN. The conditional probability density functions which result with this system are depicted in Fig. 10.2. The probability of error is given by the area under the density functions that lies across the decision threshold, and is a function of the symbol energy E_s and the one-sided noise power spectral density N_0.

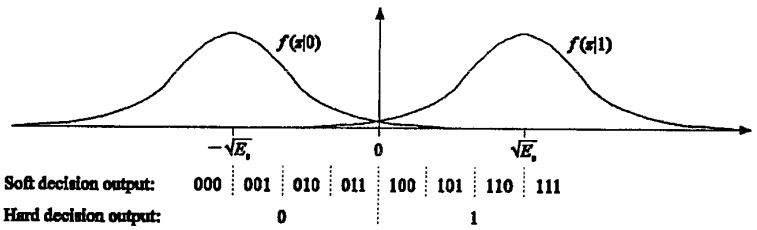


FIGURE 10.2 Hard and soft decision demodulation of a coherently demodulated BPSK signal corrupted by AWGN. $f(z \mid 1)$ and $f(z \mid 0)$ are the Gaussianly distributed conditional probability density functions at the threshold device.

Alternatively, the demodulator can make **soft decisions** or output an estimate of the symbol value along with an indication of its confidence in this estimate. For example, if the BPSK demodulator uses three-bit quantization, the two least significant bits can be taken as a confidence measure. Possible soft-decision thresholds for the BPSK signal are depicted in Fig. 10.2. In practice, there is little to be gained by using many soft-decision quantization levels.

Block and convolutional codes introduce redundancy by adding parity symbols to the message data. They map k source symbols to n code symbols and are said to have **code rate** $R = k/n$. With fixed information rates, this redundancy results in increased bandwidth and lower energy per transmitted symbol. At low signal-to-noise ratios, these codes cannot compensate for these impairments, and performance is degraded. At higher ratios of information symbol energy E_b to noise spectral density N_0, however, there is **coding gain** since the performance improvement offered by coding more than compensates for these impairments. Coding gain is usually defined as the reduction in required E_b/N_0 to achieve a specific error rate in an error-control coded system over one without coding. In contrast to block and convolutional codes, trellis-coded modulation introduces redundancy by expanding the size of the signal set rather than increasing the number of symbols transmitted, and so offers the advantages of coding to band-limited systems.

Each of these coding techniques is considered in turn. Following a discussion of **inter-**

Ford Motor Co.
Exhibit 1030
Page 254

leaving and concatenated coding, this chapter gives an overview of a recent and significant advance in coding, the development of Turbo codes, and concludes with a brief overview of FEC applications.

10.2 Fundamentals of Block Coding

In block codes there is a one-to-one mapping between k-symbol source words and n-symbol codewords. With q-ary signalling, q^k out of the q^n possible n-tuples are valid code vectors. The set of all n-tuples forms a **vector space** in which the q^k code vectors are distributed. The **Hamming distance** between any two code vectors is the number of symbols in which they differ; the **minimum distance** $d_{\min}$ of the code is the smallest Hamming distance between any two codewords.

There are two contradictory objectives of block codes. The first is to distribute the code vectors in the vector space such that the distance between them is maximized. Then, if the decoder receives a corrupted vector, by evaluating the nearest valid code vector it will decode the correct word with high probability. The second is to pack the vector space with as many code vectors as possible to reduce the redundancy in transmission.

When code vectors differ in at least $d_{\min}$ positions, a decoder which evaluates the nearest code vector to each received word is guaranteed to correct up to t random symbol errors per word if

$$d_{\min} \geq 2t + 1 \tag{10.1}$$

Alternatively, all $q^n - q^k$ illegal words can be detected, including all error patterns with $d_{\min} - 1$ or fewer errors. In general, a block code can correct all patterns of t or fewer errors and detect all patterns of u or fewer errors provided that $u \geq t$ and

$$d_{\min} \geq t + u + 1 \tag{10.2}$$

If $q = 2$, knowledge of the positions of the errors is sufficient for their correction; if $q > 2$, the decoder must determine both the positions and values of the errors. If the demodulator indicates positions in which the symbol values are unreliable, the decoder can assume their value unknown and has only to solve for the value of these symbols. These positions are called **erasures.** A block code can correct up to t errors and v erasures in each word if

$$d_{\min} \geq 2t + v + 1 \tag{10.3}$$

10.3 Structure and Decoding of Block Codes

Shannon showed that the performance limit of codes with fixed code rate improves as the block length increases. As n and k increase, however, practical implementation requires that the mapping from message to code vector not be arbitrary but that an underlying structure to the code exist. The structures developed to date limit the error correcting capability of these codes to below what Shannon proved possible, on average, for a code with random codeword assignments. Although Turbo codes have made significant strides towards approaching the Shannon Limit, the search for good constructive codes continues.

A property which simplifies implementation of the coding operations is that of code linearity. A code is **linear** if the addition of any two code vectors forms another code vector, which implies that the code vectors form a subspace of the vector space of n-tuples. This subspace, which contains the all-zero vector, is spanned by any set of k linearly independent

code vectors. Encoding can be described as the multiplication of the information k-tuple by a **generator matrix** $\boldsymbol{G}$, of dimension $k \times n$, which contains these basis vectors as rows. That is, a message vector $\boldsymbol{m}_i$ is mapped to a code vector $\boldsymbol{c}_i$ according to

$$\boldsymbol{c}_i = \boldsymbol{m}_i \boldsymbol{G}, \qquad i = 0, 1, \ldots, q^k - 1 \tag{10.4}$$

where elementwise arithmetic is defined in the **finite field** GF(q). In general, this encoding procedure results in code vectors with nonsystematic form in that the values of the message symbols cannot be determined by inspection of the code vector. However, if $\boldsymbol{G}$ has the form $[\boldsymbol{I}_k, \boldsymbol{P}]$ where $\boldsymbol{I}_k$ is the $k \times k$ identity matrix and $\boldsymbol{P}$ is a $k \times (n-k)$ matrix of parity checks, then the k most significant symbols of each code vector are identical to the message vector and the code has **systematic** form. This notation assumes that vectors are written with their most significant or first symbols in time on the left, a convention used throughout this chapter.

For each generator matrix there is an $(n-k) \times k$ **parity check matrix** $\boldsymbol{H}$ whose rows are orthogonal to the rows in $\boldsymbol{G}$, i.e., $\boldsymbol{G}\boldsymbol{H}^T = \boldsymbol{0}$. If the code is systematic, $\boldsymbol{H} = [-\boldsymbol{P}^T, \boldsymbol{I}_{n-k}]$. Since all codewords are linear sums of the rows in $\boldsymbol{G}$, it follows that $\boldsymbol{c}_i \boldsymbol{H}^T = \boldsymbol{0}$ for all $i, i = 0, 1, \ldots, q^k - 1$, and that the validity of the demodulated vectors can be checked by performing this multiplication. If a codeword $\boldsymbol{c}$ is corrupted during transmission so that the hard-decision demodulator outputs the vector $\hat{\boldsymbol{c}} = \boldsymbol{c} + \boldsymbol{e}$, where $\boldsymbol{e}$ is a nonzero error pattern, the result of this multiplication is an $(n-k)$-tuple that is indicative of the validity of the sequence. This result, called the **syndrome** $\boldsymbol{s}$, is dependent only on the error pattern since

$$\boldsymbol{s} = \hat{\boldsymbol{c}}\boldsymbol{H}^T = (\boldsymbol{c} + \boldsymbol{e})\boldsymbol{H}^T = \boldsymbol{c}\boldsymbol{H}^T + \boldsymbol{e}\boldsymbol{H}^T = \boldsymbol{e}\boldsymbol{H}^T \tag{10.5}$$

If the error pattern is a code vector, the errors go undetected. For all other error patterns, however, the syndrome is nonzero. Since there are $q^{n-k} - 1$ nonzero syndromes, $q^{n-k} - 1$ error patterns can be corrected. When these patterns include all those with t or fewer errors and no others, the code is said to be a **perfect code.** Few codes are perfect; most codes are capable of correcting some patterns with more than t errors. **Standard array decoders** use lookup tables to associate each syndrome with an error pattern but become impractical as the block length and number of parity symbols increases. Algebraic decoding algorithms have been developed for codes with stronger structure. These algorithms are simplified with imperfect codes if the patterns corrected are limited to those with t or fewer errors, a simplification called **bounded distance decoding.**

Cyclic codes are a subclass of linear block codes with an algebraic structure that enables encoding to be implemented with a linear feedback shift register and decoding to be implemented without a lookup table. As a result, most block codes in use today are cyclic or are closely related to cyclic codes. These codes are best described if vectors are interpreted as polynomials and the arithmetic follows the rules for polynomials where the elementwise operations are defined in GF(q). In a cyclic code, all codeword polynomials are multiples of a **generator polynomial** $g(x)$ of degree $n-k$. This polynomial is chosen to be a divisor of $x^n - 1$ so that a cyclic shift of a code vector yields another code vector, giving this class of codes its name. A message polynomial $m_i(x)$ can be mapped to a codeword polynomial $c_i(x)$ in nonsystematic form as

$$c_i(x) = m_i(x)g(x), \qquad i = 0, 1, \ldots, q^k - 1 \tag{10.6}$$

In systematic form, codeword polynomials have the form

$$c_i(x) = m_i(x)x^{n-k} - r_i(x), \qquad i = 0, 1, \ldots, q^k - 1 \tag{10.7}$$

Ford Motor Co.
Exhibit 1030
Page 256

where $r_i(x)$ is the remainder of $m_i(x)x^{n-k}$ divided by $g(x)$. Polynomial multiplication and division can be easily implemented with shift registers [5].

The first step in decoding the demodulated word is to determine if the word is a multiple of $g(x)$. This is done by dividing it by $g(x)$ and examining the remainder. Since polynomial division is a linear operation, the resulting syndrome $s(x)$ depends only on the error pattern. If $s(x)$ is the all-zero polynomial, transmission is errorless or an undetectable error pattern has occurred. If $s(x)$ is nonzero, at least one error has occurred. This is the principle of the **cyclic redundancy check** (CRC). It remains to determine the most likely error pattern that could have generated this syndrome.

Single error correcting binary codes can use the syndrome to immediately locate the bit in error. More powerful codes use this information to determine the locations and values of multiple errors. The most prominent approach of doing so is with the iterative technique developed by Berlekamp. This technique, which involves computing an error-locator polynomial and solving for its roots, was subsequently interpreted by Massey in terms of the design of a minimum-length shift register. Once the location and values of the errors are known, Chien's search algorithm efficiently corrects them. The implementation complexity of these decoders increases only as the square of the number of errors to be corrected [4] but does not generalize easily to accommodate soft-decision information. Other decoding techniques, including Chase's algorithm and threshold decoding, are easier to implement with soft-decision input [6]. Berlekamp's algorithm can be used in conjunction with transform-domain decoding, which involves transforming the received block with a finite field Fourier-like transform and solving for errors in the transform domain. Since the implementation complexity of these decoders depends on the block length rather than the number of symbols corrected, this approach results in simpler circuitry for codes with high redundancy [13].

Other block codes have also been constructed, including codes that are based on transform-domain spectral properties, codes that are designed specifically for correction of burst errors, and codes that are decodable with straightforward threshold or majority logic decoders [5, 6, 7].

10.4 Important Classes of Block Codes

When errors occur independently, Bose–Chaudhuri–Hocquenghem (BCH) codes provide one of the best performances of known codes for a given block length and code rate. They are cyclic codes with $n = q^m - 1$, where m is any integer greater than 2. They are designed to correct up to t errors per word and so have **designed distance** $d = 2t + 1$; the minimum distance may be greater. Generator polynomials for these codes are listed in many texts, including [6]. These polynomials are of degree less than or equal to mt, and so $k \geq n - mt$. BCH codes can be shortened to accommodate system requirements by deleting positions for information symbols.

Some subclasses of these codes are of special interest. Hamming codes are perfect single error correcting binary BCH codes. Full length codes have $n = 2^m - 1$ and $k = n - m$ for any m greater than 2. The duals of these codes are maximal-length codes, with $n = 2^m - 1$, $k = m$, and $d_{\min} = 2^{m-1}$. All $2^m - 1$ nonzero code vectors in these codes are cyclic shifts of a single nonzero code vector. Reed–Solomon (RS) codes are nonbinary BCH codes defined over GF(q), where q is often taken as a power of two so that symbols can be represented by a sequence of bits. In these cases, correction of even a single symbol allows for correction of a burst of bit errors. The block length is $n = q - 1$, and the minimum distance $d_{\min} = 2t + 1$ is achieved using only $2t$ parity symbols. Since RS codes meet the Singleton bound of

Ford Motor Co.
Exhibit 1030
Page 257

$d_{\min} \leq n-k+1$, they have the largest possible minimum distance for these values of n and k and are called **maximum distance separable** codes.

The Golay codes are the only nontrivial perfect codes that can correct more than one error. The (11, 6) ternary Golay code has minimum distance 5. The (23, 12) binary code is a triple error correcting BCH code with $d_{\min} = 7$. To simplify implementation, it is often extended to a (24, 12) code through the addition of an extra parity bit. The extended code has $d_{\min} = 8$.

The (23, 12) Golay code is also a binary quadratic residue code. These cyclic codes have prime length of the form $n = 8m \pm 1$, with $k = (n+1)/2$ and $d_{\min} \geq \sqrt{n}$. Some of these codes are as good as the best codes known with these values of n and k, but it is unknown if there are good quadratic residue codes with large n [5].

Reed-Muller codes are equivalent to binary cyclic codes with an additional overall parity bit. For any m, the rth-order Reed-Muller code has $n = 2^m$, $k = \Sigma_{i=0}^{r} \binom{m}{i}$, and $d_{\min} = 2^{m-r}$. The rth-order and $(m-r-1)$th-order codes are duals, and the first-order codes are similar to maximal-length codes. These codes, and the closely related Euclidean geometry and projective geometry codes, can be decoded with threshold decoding.

The performance of several of these block codes is shown in Fig. 10.3 in terms of decoded bit error probability vs. E_b/N_0 for systems using coherent, hard-decision demodulated BPSK signalling. Many other block codes have also been developed, including Goppa codes, quasicyclic codes, burst error correcting Fire codes, and other lesser known codes.

10.5 Principles of Convolutional Coding

Convolutional codes map successive information k-tuples to a series of n-tuples such that the sequence of n-tuples has distance properties that allow for detection and correction of errors. Although these codes can be defined over any alphabet, their implementation has largely been restricted to binary signals, and only binary convolutional codes are considered here.

In addition to the code rate $R = k/n$, the **constraint length** K is an important parameter for these codes. Definitions vary; we will use the definition that K equals the number of k-tuples that affect formation of each n-tuple during encoding. That is, the value of an n-tuple depends on the k-tuple that arrives at the encoder during that encoding interval as well as the $K-1$ previous information k-tuples.

Binary convolutional encoders can be implemented with kK-stage shift registers and n modulo-2 adders, an example of which is given in Fig. 10.4(a) for a rate 1/2, constraint length 3 code. The encoder shifts in a new k-tuple during each encoding interval and samples the outputs of the adders sequentially to form the coded output.

Although connection diagrams similar to that of Fig. 10.4(a) completely describe the code, a more concise description can be given by stating the values of n, k, and K and giving the adder connections in the form of vectors or polynomials. For instance, the rate 1/2 code has the generator vectors $\boldsymbol{g}_1 = 111$ and $\boldsymbol{g}_2 = 101$, or equivalently, the generator polynomials $g_1(x) = x^2 + x + 1$ and $g_2(x) = x^2 + 1$. Alternatively, a convolutional code can be characterized by its impulse response, the coded sequence generated due to input of a single logic-1. It is straightforward to verify that the circuit in Fig. 10.4(a) has the impulse response 111011. Since modulo-2 addition is a linear operation, convolutional codes are linear, and the coded output can be viewed as the convolution of the input sequence with the impulse response, hence the name of this coding technique. Shifted versions of the impulse response or generator vectors can be combined to form an infinite-order generator matrix which also describes the code.

Ford Motor Co.
Exhibit 1030
Page 258

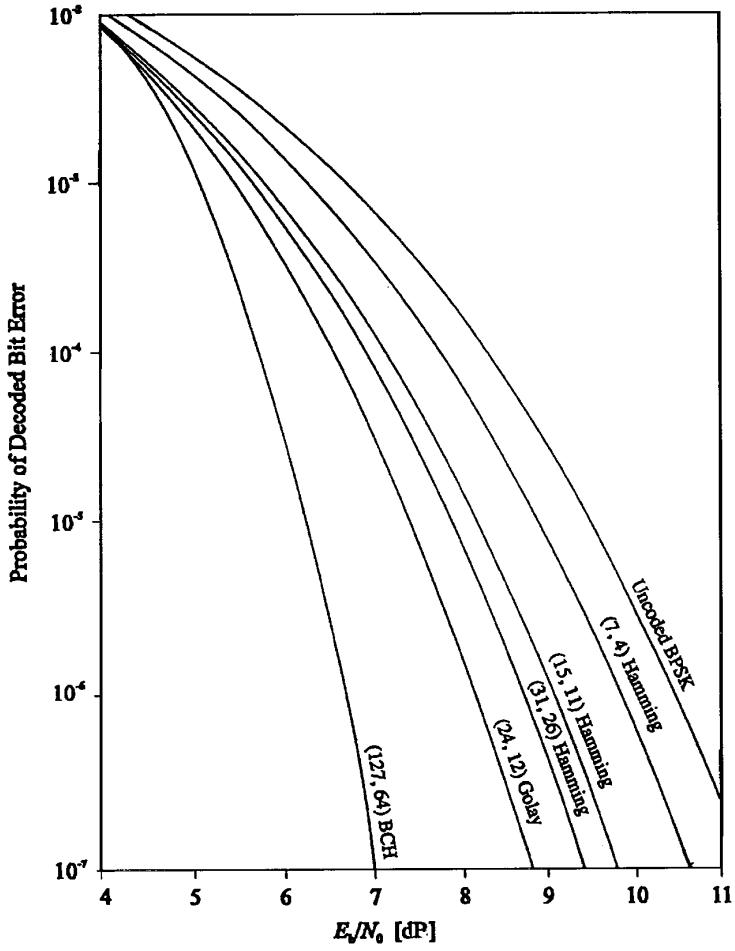
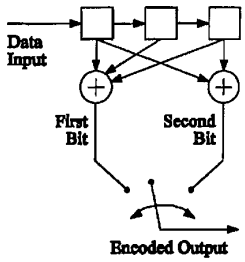


FIGURE 10.3 Block code performance. *Source:* Sklar, B., 1988, *Digital Communications: Fundamentals and Applications,* © 1988, p. 300. Reprinted by permission of Prentice-Hall, Inc., Englewood Cliffs, NJ.

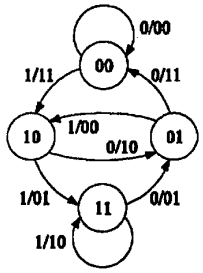
Shift register circuits can be modeled as finite state machines. A Mealy machine description of a convolutional encoder requires $2^{k(K-1)}$ states, each describing a different value of the $K-1$ k-tuples which have most recently entered the shift register. Each state has 2^k exit paths which correspond to the value of the incoming k-tuple. A state machine description for the rate 1/2 encoder depicted in Fig. 10.4(a) is given in Fig. 10.4(b). States are labeled with the contents of the two leftmost register stages; edges are labeled with information bit values and their corresponding coded output.

The dimension of time is added to the description of the encoder with tree and trellis diagrams. The tree diagram for the rate 1/2 convolutional code is given in Fig. 10.4(c), assuming the shift register is initially clear. Each node represents an encoding interval, from which the upper branch is taken if the input bit is a 0 and the lower branch is taken if the input bit is a 1. Each branch is labeled with the corresponding output bit sequence. A drawback of the tree representation is that it grows without bound as the length of the

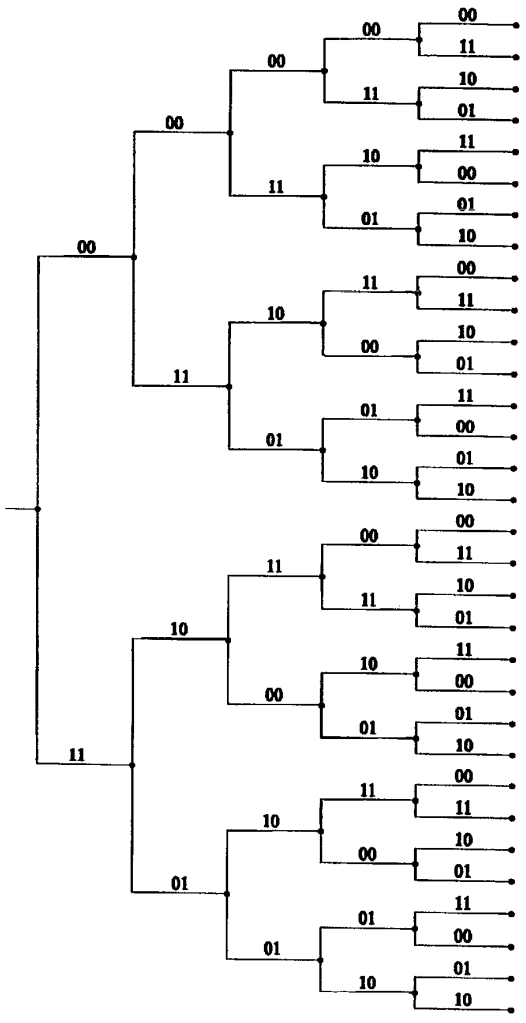
Ford Motor Co.
Exhibit 1030
Page 259



(a) Connection diagram



(b) State machine model



(c) Tree diagram

FIGURE 10.4 A rate 1/2, constraint length 3 convolutional code.

Ford Motor Co.
Exhibit 1030
Page 260

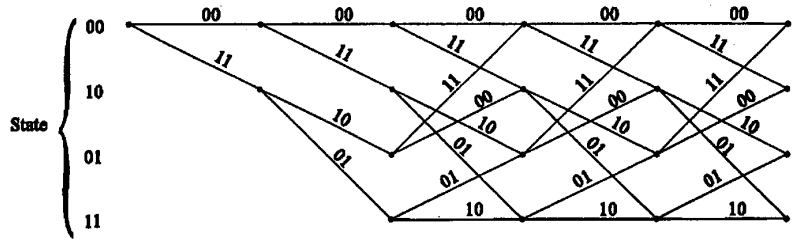


FIGURE 10.4 *(Continued).*

input sequence increases. This is overcome with the trellis diagram depicted in Fig. 10.4(d), Again, encoding results in left-to-right movement, where the upper of the two branches is taken whenever the input is a 0, the lower branch is taken when the input is a 1, and the output is the bit sequence which weights the branch taken. Each level of nodes corresponds to a state of the encoder as shown on the left-hand side of the diagram.

If the received sequence contains errors, it may no longer depict a valid path through the tree or trellis. It is the job of the decoder to determine the original path. In doing so, the decoder does not so much correct errors as find the closest valid path to the received sequence. As a result, the error correcting capability of a convolutional code is more difficult to quantify than that of a block code; it depends on how valid paths differ. One measure of this difference is the **column distance** $d_c(i)$, the minimum Hamming distance between all coded sequences generated over i encoding intervals which differ in the first interval. The nondecreasing sequence of column distance values is the **distance profile** of the code. The column distance after K intervals is the minimum distance of the code and is important for evaluating the performance of a code that uses threshold decoding. As i increases, $d_c(i)$ approaches the **free distance** of the code, d_{free}, which is the minimum Hamming distance in the set of arbitrarily long paths that diverge and then remerge in the trellis.

With maximum likelihood decoding, convolutional codes can generally correct up to t errors within three to five constraint lengths, depending on how the errors are distributed, where

$$d_{\text{free}} \geq 2t + 1 \tag{10.8}$$

The free distance can be calculated by exhaustively searching for the minimum-weight path that returns to the all-zero state, or evaluating the term of lowest degree in the generating function of the code.

The objective of a convolutional code is to maximize these distance properties. They generally improve as the constraint length of the code increases, and nonsystematic codes generally have better properties than systematic ones. Good codes have been found by computer search and are tabulated in many texts, including [6]. Convolutional codes with high code rate can be constructed by **puncturing** or periodically deleting coded symbols from a low rate code. A list of low rate codes and perforation matrices that result in good high rate codes can be found in many sources, including [13]. The performance of good punctured codes approaches that of the best convolutional codes known with similar rate, and decoder implementation is significantly less complex.

Convolutional codes can be **catastrophic,** having the potential to generate an unlimited number of decoded bit errors in response to a finite number of errors in the demodulated bit sequence. Catastrophic error propagation is avoided if the code has generator polynomials with a greatest common divisor of the form x^a for any a or, equivalently, if there are no

Ford Motor Co.
Exhibit 1030
Page 261

closed-loop paths in the state diagram with all-zero output other than the one taken with all-zero input. **Systematic codes** are not catastrophic.

10.6 Decoding of Convolutional Codes

In 1967, Viterbi developed a maximum likelihood decoding algorithm that takes advantage of the trellis structure to reduce the complexity of the evaluation. This algorithm has become known as the **Viterbi algorithm.** With each received n-tuple, the decoder computes a **metric** or measure of likelihood for all paths that could have been taken during that interval and discards all but the most likely to terminate on each node. An arbitrary decision is made if path metrics are equal. The metrics can be formed using either hard or soft decision information with little difference in implementation complexity.

If the message has finite length and the encoder is subsequently flushed with zeros, a single decoded path remains. With a BSC, this path corresponds to the valid code sequence with minimum Hamming distance from the demodulated sequence. Full-length decoding becomes impractical as the length of the message sequence increases. The most likely paths tend to have a common stem, however, and selecting the trace value four or five times the constraint length prior to the present decoding depth results in near-optimum performance. Since the number of paths examined during each interval increases exponentially with the constraint length, the Viterbi algorithm also becomes impractical for codes with large constraint length. To date, Viterbi decoding has been implemented for codes with constraint lengths up to ten. Other decoding techniques, such as sequential and threshold decoding, can be used with larger constraint lengths.

Sequential decoding was proposed by Wozencraft, and the most widely used algorithm was developed by Fano. Rather than tracking multiple paths through the trellis, the sequential decoder operates on a single path while searching the code tree for a path with high probability. It makes tentative decisions regarding the transmitted sequence, computes a metric between its proposed path and the demodulated sequence, and moves forward through the tree as long as the metric indicates that the path is likely. If the likelihood of the path becomes low, the decoder moves backward, searching other paths until it finds one with high probability. The number of computations involved in this procedure is almost independent of the constraint length and is typically quite small, but it can be highly variable, depending on the channel. Buffers must be provided to store incoming sequences as the decoder searches the tree. Their overflow is a significant limiting factor in the performance of these decoders.

Figure 10.5 compares the performance of the Viterbi and sequential decoding algorithms for several convolutional codes operating on coherently demodulated BPSK signals corrupted by AWGN. Other decoding algorithms have also been developed, including syndrome decoding methods such as table look-up feedback decoding and threshold decoding [6]. These algorithms are easily implemented but offer suboptimal performance. Techniques such as the one discussed by [1] have been developed to support both soft input and soft output, but these decoding techniques typically increase decoder complexity.

10.7 Trellis-Coded Modulation

Trellis-coded modulation (TCM) has received considerable attention since its development by Ungerboeck in the late 1970s [11]. Unlike block and convolutional codes, TCM schemes achieve coding gain by increasing the size of the signal alphabet and using multilevel/phase signalling. Like convolutional codes, sequences of coded symbols are restricted to certain

Ford Motor Co.
Exhibit 1030
Page 262

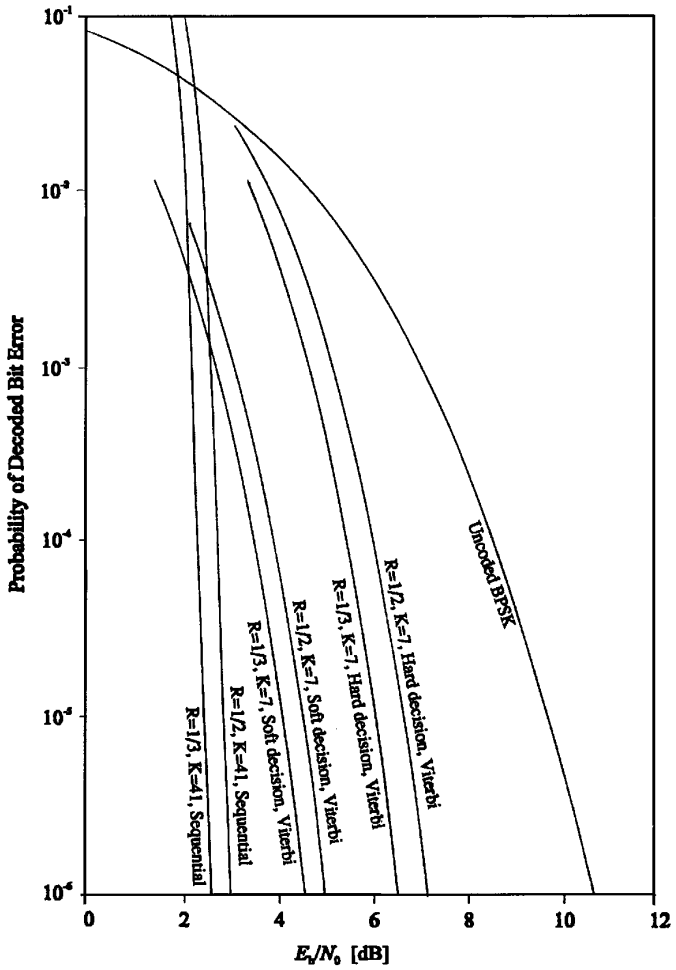


FIGURE 10.5 Convolutional code performance. *Source:* Omura, J.K. and Levitt, B.K., © 1982 IEEE, "Coded Error Probability Evaluation for Antijam Communication Systems," *IEEE Trans. Commun.*, vol. COM-30, no. 5, pp. 896–903. Reprinted by permission of IEEE.

valid patterns. In TCM, these patterns are chosen to have large Euclidean distance from one another so that a large number of corrupted sequences can be corrected. The Viterbi algorithm is often used to decode these sequences. Since the symbol transmission rate does not increase, coded and uncoded signals require the same transmission bandwidth. If transmission power is held constant, the signal constellation of the coded signal is denser. The loss in symbol separation, however, is more than overcome by the error correction capability of the code.

Ungerboeck investigated the increase in channel capacity that can be obtained by increasing the size of the signal set and restricting the pattern of transmitted symbols, and concluded that almost all of the additional capacity can be gained by doubling the number of points in the signal constellation. This is accomplished by encoding the binary data

Ford Motor Co.
Exhibit 1030
Page 263

with a rate $R = k/(k+1)$ code and mapping sequences of $k+1$ coded bits to points in a constellation of 2^{k+1} symbols. For example, the rate 2/3 encoder of Fig. 10.6(a) encodes pairs of source bits to three coded bits. Figure 10.6(b) depicts one stage in the trellis of the coded output where, with the convolutional code, the state of the encoder is defined by the values of the two most recent bits to enter the shift register. Note that unlike the trellis for the convolutional code, this trellis contains parallel paths between nodes.

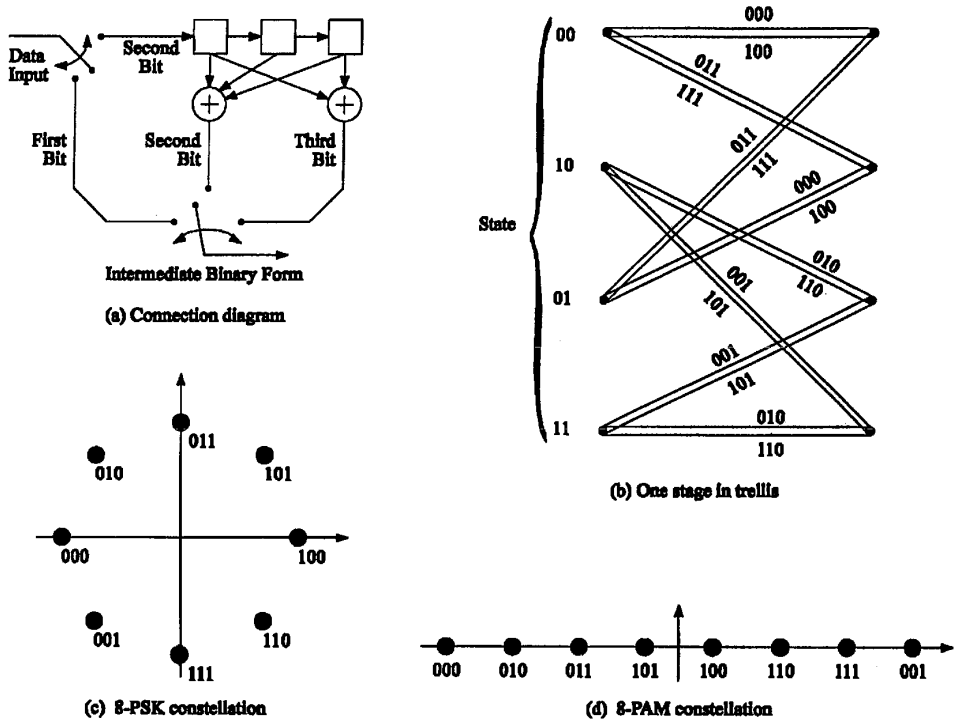


FIGURE 10.6 Rate 2/3 trellis-coded modulation.

The key to improving performance with TCM is to map the coded bits to points in the signal space such that the Euclidean distance between transmitted sequences is maximized. A method that ensures improved Euclidean distance is the method of **set partitioning.** This involves separating all parallel paths on the trellis with maximum distance and assigning the next greatest distance to paths that diverge from or merge onto the same node. Figures 10.6(c) and 10.6(d) give examples of mappings for the rate 2/3 code with 8-PSK and 8-PAM signal constellations, respectively.

As with convolutional codes, the free distance of a TCM code is defined as the minimum distance between paths through the trellis, where the distance of concern is now Euclidean distance rather than Hamming distance. The free distance of an uncoded signal is defined as the distance between the closest signal points. When coded and uncoded signals have the same average power, the coding gain of the TCM system is defined as

$$\text{coding gain} = 20\log_{10}\left(\frac{d_{\text{free, coded}}}{d_{\text{free, uncoded}}}\right) \tag{10.9}$$

Ford Motor Co.
Exhibit 1030
Page 264

It can be shown that the simple, rate 2/3 8 phase-shift keying (PSK) and 8 pulse-amplitude modulation (PAM) TCM systems provide gains of 3 dB and 3.3 dB, respectively, [6]. More complex TCM systems yield gains up to 6 dB. Tables of good codes are given in [11].

10.8 Additional Measures

When the demodulated sequence contains bursts of errors, the performance of codes designed to correct independent errors improves if coded sequences are **interleaved** prior to transmission and deinterleaved prior to decoding. Deinterleaving separates the burst errors, making them appear more random and increasing the likelihood of accurate decoding. It is generally sufficient to interleave several block lengths of a block coded signal or several constraint lengths of a convolutionally encoded signal. Block interleaving is the most straightforward approach, but delay and memory requirements are halved with convolutional and helical interleaving techniques. Periodicity in the way sequences are combined is avoided with pseudorandom interleaving.

Serially **concatenated codes,** first investigated by Forney, use two levels of coding to achieve a level of performance with less complexity than a single coding stage would require. The inner code interfaces with the modulator and demodulator and corrects the majority of the errors; the outer code corrects errors that appear at the output of the inner-code decoder. A convolutional code with Viterbi decoding is usually chosen as the inner code, and an RS code is often chosen as the outer code due to its ability to correct the bursts of bit errors which can result with incorrect decoding of trellis-coded sequences. Interleaving and deinterleaving outer-code symbols between coding stages offers further protection against the burst error output of the inner code.

Product codes effectively place the data in a two dimensional array and use FEC techniques over both the rows and columns of this array. Not only do these codes result in error protection in two dimensions, but the manner in which the array is constructed can offer advantages similar to those achieved through interleaving.

10.9 Turbo Codes

The most recent significant achievement in FEC coding is the development of **Turbo codes** [3]. The principle of this coding technique is to encode the data with two or more **constituent codes** concatenated in parallel form. The received sequence is decoded in an iterative, serial approach using soft-input, soft-output decoders. This iterative decoding approach involves feedback of information in a manner similar to processes within the turbo engine, giving this coding technique its name.

Turbo codes effectively result in the construction of relatively long codewords with few codewords being close in terms of Hamming distance, while at the same time constraining the implementation complexity of the decoder to practical limits. The first Turbo codes developed used recursive systematic convolutional codes as the constituent codes, and punctured them to improve the code rate. The use of other constituent codes has since been considered. Two or more of these codes are concatenated in parallel,where code concatenation is combined with interleaving in order to increase the independence of the data sequences encoded by the constituent encoders. This apparent increase in randomness, implemented with simple interleavers, is an important contributing factor to the excellent performance of the decoders.

As in other multi-stage coding techniques, the complexity of the decoder is limited through use of separate decoding stages for each constituent code. The input to the first

Ford Motor Co.
Exhibit 1030
Page 265

stage is the soft output of the demodulator for a finite-length received symbol sequence. Subsequent stages use both the demodulator output and an output of the previous decoding stage which is indicative of the reliability of the symbols. This information, gleaned from soft-output decoders, is called **extrinsic information**.

Decoding proceeds by iterating through constituent decoders, each forwarding updated extrinsic information to the next decoder, until a predefined number of iterations has been completed or the extrinsic information indicates that high reliability has been achieved. This approach results in very good performance at low values of E_b/N_0. Simulations have demonstrated error rates of 10^{-5} at signal-to-noise ratios appreciably less than 1 dB. At higher values of E_b/N_0, however, the performance curves can exhibit flattening if constituent codes are chosen in a manner that results in an overall small Hamming distance for the code.

Although this coding technique has shown great promise, there remains considerable work with regard to optimizing code parameters. Great strides have been made over the last few years in understanding the structure of these codes and relating them to serially concatenated and product codes, but many researchers are still examining these codes in order to advance their development. With this research will come optimization of the Turbo code process and application of these codes in various communication systems.

10.10 Applications

FEC coding remained of theoretical interest until advances in digital technology and improvements in decoding algorithms made their implementation possible. It has since become an attractive alternative to improving other system components or boosting transmission power. FEC codes are commonly used in digital storage systems, deep-space and satellite communication systems, terrestrial radio and band limited wireline systems, and have also been proposed for fiber optic transmission. Accordingly, the theory and practice of error correcting codes now occupies a prominent position in the field of communications engineering.

Deep-space systems began using forward error correction in the early 1970s to reduce transmission power requirements, and used multiple error correcting RS codes for the first time in 1977 to protect against corruption of compressed image data in the Voyager missions [12]. The Consultative Committee for Space Data Systems (CCSDS) has since recommended use of a concatenated coding system which uses a rate 1/2, constraint length 7 convolutional inner code and a (255, 223) RS outer code.

Coding is now commonly used in satellite systems to reduce power requirements and overall hardware costs and to allow closer orbital spacing of geosynchronous satellites [2]. FEC codes play integral roles in the VSAT, MSAT, INTELSAT, and INMARSAT systems [13]. Further, a (31, 15) RS code is used in the joint tactical information distribution system (JTIDS), a (7, 2) RS code is used in the air force satellite communication system (AFSATCOM), and a (204, 192) RS code has been designed specifically for satellite time division multiple access (TDMA) systems. Another code designed for military applications involves concatenation of a Golay and RS code with interleaving to ensure an imbalance of 1's and 0's in the transmitted symbol sequence and enhance signal recovery under severe noise and interference [2].

TCM has become commonplace in transmission of data over voiceband telephone channels. Modems developed since 1984 use trellis coded QAM modulation to provide robust communication at rates above 9.6 kb/s. Various coding techniques are used in the new digital cellular and personal communication standards, with an emphasis on convolutional and cyclic redundancy check codes [8].

Ford Motor Co.
Exhibit 1030
Page 266

FEC codes have also been widely used in digital recording systems, most prominently in the compact disc digital audio system. This system uses two levels of coding and interleaving in the cross-interleaved RS coding (CIRC) system to correct errors that result from disc imperfections and dirt and scratches which accumulate during use. Steps are also taken to mute uncorrectable sequences [12].

Defining Terms

Binary symmetric channel: A memoryless discrete data channel with binary signalling, hard-decision demodulation, and channel impairments that do not depend on the value of the symbol transmitted.

Bounded distance decoding: Limiting the error patterns which are corrected in an imperfect code to those with t or fewer errors.

Catastrophic code: A convolutional code in which a finite number of code symbol errors can cause an unlimited number of decoded bit errors.

Code rate: The ratio of source word length to codeword length, indicative of the amount of information transmitted per encoded symbol.

Coding gain: The reduction in signal-to-noise ratio required for specified error performance in a block or convolutional coded system over an uncoded system with the same information rate, channel impairments, and modulation and demodulation techniques. In TCM, the ratio of the squared free distance in the coded system to that of the uncoded system.

Column distance: The minimum Hamming distance between convolutionally encoded sequences of a specified length with different leading n-tuples.

Constituent codes: Two or more FEC codes that are combined in concatenated coding techniques.

Cyclic code: A block code in which cyclic shifts of code vectors are also code vectors.

Cyclic redundancy check: When the syndrome of a cyclic block code is used to detect errors.

Designed distance: The guaranteed minimum distance of a BCH code designed to correct up to t errors.

Discrete data channel: The concatenation of all system elements between FEC encoder output and decoder input.

Distance profile: The minimum Hamming distance after each encoding interval of convolutionally encoded sequences which differ in the first interval.

Erasure: A position in the demodulated sequence where the symbol value is unknown.

Extrinsic information: The output of a constituent soft decision decoder that is forwarded as input to the next decoding stage in iterative decoding of Turbo codes.

Finite field: A finite set of elements and operations of addition and multiplication that satisfy specific properties. Often called Galois fields and denoted GF(q), where q is the number of elements in the field. Finite fields exist for all q which are prime or the power of a prime.

Free distance: The minimum Hamming weight of convolutionally encoded sequences that diverge and remerge in the trellis. Equals the maximum column distance and the limiting value of the distance profile.

Ford Motor Co.
Exhibit 1030
Page 267

Generator matrix: A matrix used to describe a linear code. Code vectors equal the information vectors multiplied by this matrix.

Generator polynomial: The polynomial that is a divisor of all codeword polynomials in a cyclic block code; a polynomial that describes circuit connections in a convolutional encoder.

Hamming distance: The number of symbols in which codewords differ.

Hard decision: Demodulation that outputs only a value for each received symbol.

Interleaving: Shuffling the coded bit sequence prior to modulation and reversing this operation following demodulation. Used to separate and redistribute burst errors over several codewords (block codes) or constraint lengths (trellis codes) for higher probability of correct decoding by codes designed to correct random errors.

Linear code: A code whose code vectors form a vector space. Equivalently, a code where the addition of any two code vectors forms another code vector.

Maximum distance separable: A code with the largest possible minimum distance given the block length and code rate. These codes meet the Singleton bound of $d_{\min} \leq n - k + 1$.

Metric: A measure of goodness against which items are judged. In the Viterbi algorithm, an indication of the probability of a path being taken given the demodulated symbol sequence.

Minimum distance: In a block code, the smallest Hamming distance between any two codewords. In a convolutional code, the column distance after K intervals.

Parity check matrix: A matrix whose rows are orthogonal to the rows in the generator matrix of a linear code. Errors can be detected by multiplying the received vector by this matrix.

Perfect code: A t error correcting (n, k) block code in which $q^{n-k} - 1 = \Sigma_{i=1}^{t} \binom{n}{i}$.

Puncturing: Periodic deletion of code symbols from the sequence generated by a convolutional encoder for purposes of constructing a higher rate code. Also, deletion of parity bits in a block code.

Set partitioning: Rules for mapping coded sequences to points in the signal constellation that always result in a larger Euclidean distance for a TCM system than an uncoded system, given appropriate construction of the trellis.

Shannon Limit: The ratio of energy per data bit E_b to one-sided noise power spectral density N_0 in an AWGN channel above which errorless transmission is possible when bandwidth limitations are not placed on the signal and transmission is at channel capacity. This limit has the value $\ln 2 = 0.693 = -1.6$ dB.

Soft decision: Demodulation that outputs an estimate of the received symbol value along with an indication of the reliability of this value. Usually implemented by quantizing the received signal to more levels than there are symbol values.

Standard array decoding: Association of an error pattern with each syndrome by way of a lookup table.

Syndrome: An indication of whether or not errors are present in the demodulated symbol sequence.

Systematic code: A code in which the values of the message symbols can be identified by inspection of the code vector.

Vector space: An algebraic structure comprised of a set of elements in which operations of vector addition and scalar multiplication are defined. For our purposes, a

Ford Motor Co.
Exhibit 1030
Page 268

set of n-tuples consisting of symbols from GF(q) with addition and multiplication defined in terms of elementwise operations from this finite field.

Viterbi algorithm: A maximum-likelihood decoding algorithm for trellis codes that discards low-probability paths at each stage of the trellis, thereby reducing the total number of paths that must be considered.

References

[1] Bahl, L.R., Cocke, J., Jelinek, F., and Raviv, J., Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate. *IEEE Transactions on Information Theory,* 20, 248–287, 1974.

[2] Berlekamp, E.R., Peile, R.E., and Pope, S.P., The application of error control to communications. *IEEE Commun. Mag.,* 25(4), 44–57, 1987.

[3] Berrou, C., Glavieux, A., and Thitimajshima, P., Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes. *Proceedings of ICC'93,* Geneva, Switzerland, 1064–1070, 1993. Later expanded and published as: Berrou, C., Glavieux, A., 1996. Near Optimum Error Correcting Coding and Decoding. *IEEE Transactions on Communications,* 44(10), 1261–1271, 1996.

[4] Bhargava, V.K., Forward error correction schemes for digital communications. *IEEE Commun. Mag.,* 21(1), 11–19, 1983.

[5] Blahut, R.E., *Theory and Practice of Error Control Codes,* Addison-Wesley, Reading, MA, 1983.

[6] Clark, G.C. Jr. and Cain, J.B., *Error Correction Coding for Digital Communications,* Plenum Press, New York, 1981.

[7] Lin, S. and Costello, D.J. Jr., *Error Control Coding: Fundamentals and Applications,* Prentice-Hall, Englewood Cliffs, NJ, 1983.

[8] Rappaport, T.S., *Wireless Communications, Principles and Practice,* Prentice-Hall and IEEE Press, NJ, 1996.

[9] Shannon, C.E., A mathematical theory of communication. *Bell Syst. Tech. J.,* 27(3), 379–423 and 623–656, 1948.

[10] Sklar, B., *Digital Communications: Fundamentals and Applications,* Prentice-Hall, Englewood Cliffs, NJ, 1988.

[11] Ungerboeck, G., Trellis-coded modulation with redundant signal sets. *IEEE Commun. Mag.,* 25(2), 5–11 and 12–21, 1987.

[12] Wicker, S.B. and Bhargava, V.K., *Reed-Solomon Codes and Their Applications,* IEEE Press, NJ, 1994.

[13] Wu, W.W., Haccoun, D., Peile, R., and Hirata, Y., Coding for satellite communication. *IEEE J. Selected Areas in Commun.,* SAC-5(4), 724–748, 1987.

Further Information

There is now a large amount of literature on the subject of FEC coding. An introduction to the philosophy and limitations of these codes can be found in the second chapter of Lucky's book *Silicon Dreams: Information, Man, and Machine,* St. Martin's Press, New York, 1989. More practical introductions can be found in overview chapters of many communications texts. The number of texts devoted entirely to this subject also continues to grow. Although these texts summarize the algebra underlying block codes, more in-depth treatments can be found in mathematical texts. Survey papers appear occasionally in the literature, but the interested reader is directed to the seminal papers by Shannon, Ham-

Ford Motor Co.
Exhibit 1030
Page 269

ming, Reed and Solomon, Bose and Chaudhuri, Hocquenghem, Wozencraft, Fano, Forney, Berlekamp, Massey, Viterbi, Ungerboeck, Berrou and Glavieux, among others. The most recent advances in the theory and implementation of error control codes are published in *IEEE Transactions on Information Theory, IEEE Transactions on Communications,* and special issues of *IEEE Journal on Selected Areas in Communications.*

Ford Motor Co.
Exhibit 1030
Page 270

11

Spread Spectrum Communications

Laurence B. Milstein
University of California

Marvin K. Simon
Jet Propulsion Laboratory

11.1 A Brief History 11-1
11.2 Why Spread Spectrum? 11-1
11.3 Basic Concepts and Terminology 11-2
11.4 Spread Spectrum Techniques 11-2
Direct Sequence Modulation • Frequency Hopping Modulation • Time Hopping Modulation • Hybrid Modulations
11.5 Applications of Spread Spectrum 11-8
Military • Commercial
Defining Terms 11-13
References 11-14

11.1 A Brief History

Spread spectrum (SS) has its origin in the military arena where the friendly communicator is 1) susceptible to detection/interception by the enemy and 2) vulnerable to intentionally introduced unfriendly interference (jamming). Communication systems that employ spread spectrum to reduce the communicator's detectability and combat the enemy-introduced interference are respectively referred to as **low probability of intercept (LPI)** and **antijam** (AJ) **communication systems**. With the change in the current world political situation wherein the U.S. Department of Defense (DOD) has reduced its emphasis on the development and acquisition of new communication systems for the original purposes, a host of new commercial applications for SS has evolved, particularly in the area of cellular mobile communications. This shift from military to commercial applications of SS has demonstrated that the basic concepts that make SS techniques so useful in the military can also be put to practical peacetime use. In the next section, we give a simple description of these basic concepts using the original military application as the basis of explanation. The extension of these concepts to the mentioned commercial applications will be treated later on in the chapter.

11.2 Why Spread Spectrum?

Spread spectrum is a communication technique wherein the transmitted modulation is *spread* (increased) in bandwidth prior to transmission over the channel and then *despread* (decreased) in bandwidth by the same amount at the receiver. If it were not for the fact that the communication channel introduces some form of narrowband (relative to the spread bandwidth) interference, the receiver performance would be transparent to the spreading

0-8493-8597-0/99/$0.00+$.50
© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 271

and despreading operations (assuming that they are identical inverses of each other). That is, after **despreading** the received signal would be identical to the transmitted signal prior to **spreading**. In the presence of narrowband interference, however, there is a significant advantage to employing the spreading/despreading procedure described. The reason for this is as follows. Since the interference is introduced after the transmitted signal is spread, then, whereas the despreading operation at the receiver shrinks the desired signal back to its original bandwidth, at the same time it spreads the undesired signal (interference) in bandwidth by the same amount, thus reducing its power spectral density. This, in turn, serves to diminish the effect of the interference on the receiver performance, which depends on the amount of interference power in the despread bandwidth. It is indeed this very simple explanation, which is at the heart of all spread spectrum techniques.

11.3 Basic Concepts and Terminology

To describe this process analytically and at the same time introduce some terminology that is common in spread spectrum parlance, we proceed as follows. Consider a communicator that desires to send a message using a transmitted power S Watts (W) at an information rate R_b bits/s (bps). By introducing a SS modulation, the bandwidth of the transmitted signal is increased from R_b Hz to $W_{\rm SS}$ Hz where $W_{\rm SS} \gg R_b$ denotes the **spread spectrum bandwidth**. Assume that the channel introduces, in addition to the usual thermal noise (assumed to have a single-sided power spectral density (PSD) equal to N_0 W/Hz), an additive interference (jamming) having power J distributed over some bandwidth W_J. After despreading, the desired signal bandwidth is once again now equal to R_b Hz and the interference PSD is now $N_J = J/W_{\rm SS}$. Note that since the thermal noise is assumed to be white, i.e., it is uniformly distributed over all frequencies, its PSD is unchanged by the despreading operation and, thus, remains equal to N_0. Regardless of the signal and interferer waveforms, the equivalent bit energy-to-total noise spectral density ratio is, in terms of the given parameters,

$$\frac{E_b}{N_t} = \frac{E_b}{N_0 + N_J} = \frac{S/R_b}{N_0 + J/W_{\rm SS}} \tag{11.1}$$

For most practical scenarios, the jammer limits performance and, thus, the effects of receiver noise in the channel can be ignored. Thus, assuming $N_J \gg N_0$, we can rewrite Eq. (11.1) as

$$\frac{E_b}{N_t} \cong \frac{E_b}{N_J} = \frac{S/R_b}{J/W_{\rm SS}} = \frac{S}{J}\frac{W_{\rm SS}}{R_b} \tag{11.2}$$

where the ratio J/S is the *jammer-to-signal power ratio* and the ratio $W_{\rm SS}/R_b$ is the **spreading ratio** and is defined as the **processing gain** of the system. Since the ultimate error probability performance of the communication receiver depends on the ratio E_b/N_J, we see that from the communicator's viewpoint his goal should be to minimize J/S (by choice of S) and maximize the processing gain (by choice of $W_{\rm SS}$ for a given desired information rate). The possible strategies for the jammer will be discussed in the section on military applications dealing with AJ communications.

11.4 Spread Spectrum Techniques

By far the two most popular spreading techniques are **direct sequence (DS) modulation** and **frequency hopping (FH) modulation**. In the following subsections, we present a brief description of each.

Ford Motor Co.
Exhibit 1030
Page 272

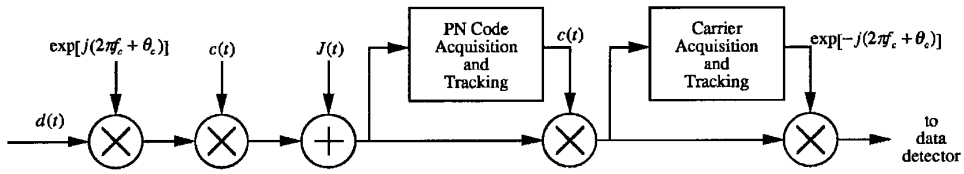


FIGURE 11.1 A DS-BPSK system (complex form).

11.4.1 Direct Sequence Modulation

A direct sequence modulation $c(t)$ is formed by linearly modulating the output sequence $\{c_n\}$ of a pseudorandom number generator onto a train of pulses, each having a duration T_c called the **chip time**. In mathematical form,

$$c(t) = \sum_{n=-\infty}^{\infty} c_n p\,(t - nT_c) \tag{11.3}$$

where $p(t)$ is the basic pulse shape and is assumed to be of rectangular form. This type of modulation is usually used with binary phase-shift-keyed (BPSK) information signals, which have the complex form $d(t)\exp\{j(2\pi f_c t + \theta_c)\}$, where $d(t)$ is a binary-valued data waveform of rate $1/T_b$ bits/s and f_c and θ_c are the frequency and phase of the data-modulated carrier, respectively. As such, a DS/BPSK signal is formed by multiplying the BPSK signal by $c(t)$ (see Fig. 11.1), resulting in the real transmitted signal

$$x(t) = \mathrm{Re}\,\{c(t)d(t)\exp\,[j\,(2\pi f_c t + \theta_c)]\} \tag{11.4}$$

Since T_c is chosen so that $T_b \gg T_c$, then relative to the bandwidth of the BPSK information signal, the bandwidth of the DS/BPSK signal[1] is effectively increased by the ratio $T_b/T_c = W_{SS}/2R_b$, which is one-half the spreading factor or processing gain of the system. At the receiver, the sum of the transmitted DS/BPSK signal and the channel interference $I(t)$ (as discussed before, we ignore the presence of the additive thermal noise) are ideally multiplied by the identical DS modulation (this operation is known as despreading), which returns the DS/BPSK signal to its original BPSK form whereas the real interference signal is now the real wideband signal $\mathrm{Re}\{I(t)c(t)\}$. In the previous sentence, we used the word ideally, which implies that the PN waveform used for despreading at the receiver is identical to that used for spreading at the transmitter. This simple implication covers up a multitude of tasks that a practical DS receiver must perform. In particular, the receiver must first acquire the PN waveform. That is, the local PN random generator that generates the PN waveform **at** the receiver used for despreading must be aligned (synchronized) to within one chip of the PN waveform of the received DS/BPSK signal. This is accomplished by employing some sort of **search algorithm** which typically steps the local PN waveform sequentially in time by a fraction of a chip (e.g., half a chip) and at each position searches for a high degree of correlation between the received and local PN reference waveforms. The search terminates when the correlation exceeds a given threshold, which is an indication that the alignment has been achieved. After bringing the two PN waveforms into **coarse alignment**, a **tracking algorithm** is employed to maintain **fine alignment**. The most

[1]For the usual case of a rectangular spreading pulse $p(t)$, the PSD of the DS/BPSK modulation will have $(\sin x/x)^2$ form with first zero crossing at $1/T_c$, which is nominally taken as one-half the spread spectrum bandwidth W_{SS}.

Ford Motor Co.
Exhibit 1030
Page 273

popular forms of tracking loops are the continuous time **delay-locked loop** and its time-multiplexed version the **tau–dither loop**. It is the difficulty in synchronizing the receiver PN generator to subnanosecond accuracy that limits PN chip rates to values on the order of hundreds of Mchips/s, which implies the same limitation on the DS spread spectrum bandwidth W_{SS}.

11.4.2 Frequency Hopping Modulation

A **frequency hopping (FH) modulation** $c(t)$ is formed by nonlinearly modulating a train of pulses with a sequence of pseudorandomly generated frequency shifts $\{f_n\}$. In mathematical terms, $c(t)$ has the complex form

$$c(t) = \sum_{n=-\infty}^{\infty} \exp\{j\,(2\pi f_n + \phi_n)\}\, p\,(t - nT_h) \tag{11.5}$$

where $p(t)$ is again the basic pulse shape having a duration T_h, called the **hop time** and $\{\phi_n\}$ is a sequence of random phases associated with the generation of the hops. FH modulation is traditionally used with multiple-frequency-shift-keyed (MFSK) information signals, which have the complex form $\exp\{j[2\pi(f_c + d(t))t]\}$, where $d(t)$ is an M-level digital waveform (M denotes the symbol alphabet size) representing the information frequency modulation at a rate $1/T_s$ symbols/s (sps). As such, an FH/MFSK signal is formed by complex multiplying the MFSK signal by $c(t)$ resulting in the real transmitted signal

$$x(t) = \mathrm{Re}\,\{c(t)\exp\{j\,[2\pi(f_c + d(t))t]\}\} \tag{11.6}$$

In reality, $c(t)$ is never generated in the transmitter. Rather, $x(t)$ is obtained by applying the sequence of pseudorandom frequency shifts $\{f_n\}$ directly to the frequency synthesizer that generates the carrier frequency f_c (see Fig. 11.2). In terms of the actual implementation,

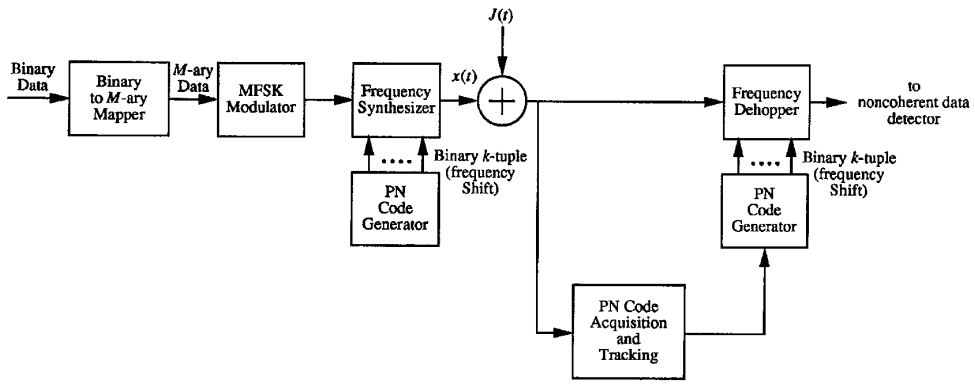


FIGURE 11.2 An FH-MFSK system.

successive (not necessarily disjoint) k-chip segments of a PN sequence drive a frequency synthesizer, which hops the carrier over 2^k frequencies. In view of the large bandwidths over which the frequency synthesizer must operate, it is difficult to maintain phase coherence from hop to hop, which explains the inclusion of the sequence $\{\phi_n\}$ in the Eq. (11.5) model for $c(t)$. On a short term basis, e.g., within a given hop, the signal bandwidth is identical

Ford Motor Co.
Exhibit 1030
Page 274

to that of the MFSK information modulation, which is typically much smaller than W_{SS}. On the other hand, when averaged over many hops, the signal bandwidth is equal to W_{SS}, which can be on the order of several GHz, i.e., an order of magnitude larger than that of implementable DS bandwidths. The exact relation between W_{SS}, T_h, T_s and the number of frequency shifts in the set $\{f_n\}$ will be discussed shortly.

At the receiver, the sum of the transmitted FH/MFSK signal and the channel interference $I(t)$ is ideally complex multiplied by the identical FH modulation (this operation is known as **dehopping**), which returns the FH/MFSK signal to its original MFSK form, whereas the real interference signal is now the wideband (in the average sense) signal $\mathrm{Re}\{I(t)c(t)\}$. Analogous to the DS case, the receiver must acquire and track the FH signal so that the dehopping waveform is as close to the hopping waveform $c(t)$ as possible.

FH systems are traditionally classified in accordance with the relationship between T_h and T_s. **Fast frequency-hopped (FFH)** systems are ones in which there exists one or more hops per data symbol, that is, $T_s = NT_h$ (N an integer) whereas **slow frequency-hopped (SFH)** systems are ones in which there exists more than one symbol per hop, that is, $T_h = NT_s$. It is customary in SS parlance to refer to the FH/MFSK tone of shortest duration as a "chip", despite the same usage for the PN chips associated with the code generator that drives the frequency synthesizer. Keeping this distinction in mind, in an FFH system where, as already stated, there are multiple hops per data symbol, a chip is equal to a hop. For SFH, where there are multiple data symbols per hop, a chip is equal to an MFSK symbol. Combining these two statements, the chip rate R_c in an FH system is given by the larger of $R_h = 1/T_h$ and $R_s = 1/T_s$ and, as such, is the highest system clock rate.

The frequency spacing between the FH/MFSK tones is governed by the chip rate R_c and is, thus, dependent on whether the FH modulation is FFH or SFH. In particular, for SFH where $R_c = R_s$, the spacing between FH/MFSK tones is equal to the spacing between the MFSK tones themselves. For noncoherent detection (the most commonly encountered in FH/MFSK systems), the separation of the MFSK symbols necessary to provide orthogonality[2] is an integer multiple of R_s. Assuming the minimum spacing, i.e., R_s, the entire spread spectrum band is then partitioned into a total of $N_t = W_{SS}/R_s = W_{SS}/R_c$ equally spaced FH tones. One arrangement, which is by far the most common, is to group these N_t tones into $N_b = N_t/M$ contiguous, nonoverlapping bands, each with bandwidth $MR_s = MR_c$; see Fig. 11.3a. Assuming symmetric MFSK modulation around the carrier frequency, then the center frequencies of the $N_b = 2^k$ bands represent the set of hop carriers, each of which is assigned to a given k-tuple of the PN code generator. In this fixed arrangement, each of the N_t FH/MFSK tones corresponds to the combination of a unique hop carrier (PN code k-tuple) and a unique MFSK symbol. Another arrangement, which provides more protection against the sophisticated interferer (jammer), is to overlap adjacent M-ary bands by an amount equal to R_c; see Fig. 11.3b. Assuming again that the center frequency of each band corresponds to a possible hop carrier, then since all but $M-1$ of the N_t tones are available as center frequencies, the number of hop carriers has been increased from N_t/M to $N_t-(M-1)$, which for $N_t \gg M$ is approximately an increase in randomness by a factor of M.

[2]An optimum noncoherent MFSK detector consists of a bank of energy detectors each matched to one of the M frequencies in the MFSK set. In terms of this structure, the notion of *orthogonality* implies that for a given transmitted frequency there will be no crosstalk (energy spillover) in any of the other $M-1$ energy detectors.

Ford Motor Co.
Exhibit 1030
Page 275

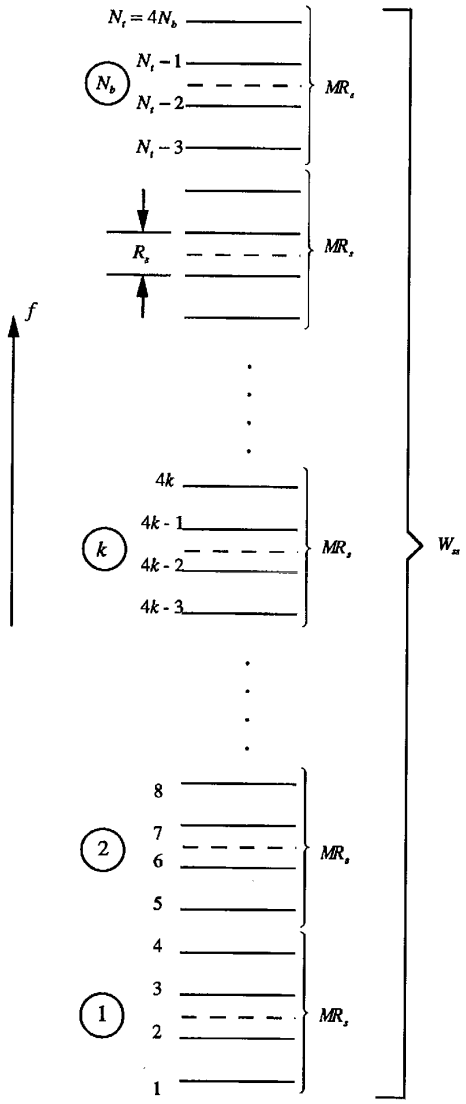


Figure 11.3a Frequency distribution for FH-4FSK—nonoverlapping bands. Dashed lines indicate location of hop frequencies.

For FFH, where $R_c = R_h$, the spacing between FH/MFSK tones is equal to the hop rate. Thus, the entire spread spectrum band is partitioned into a total of $N_t = W_{SS}/R_h = W_{SS}/R_c$ equally spaced FH tones, each of which is assigned to a unique k-tuple of the PN code generator that drives the frequency synthesizer. Since for FFH there are R_h/R_s hops per symbol, then the metric used to make a noncoherent decision on a particular symbol is obtained by summing up R_h/R_s detected chip (hop) energies, resulting in a so-called *noncoherent combining loss.*

Ford Motor Co.
Exhibit 1030
Page 276

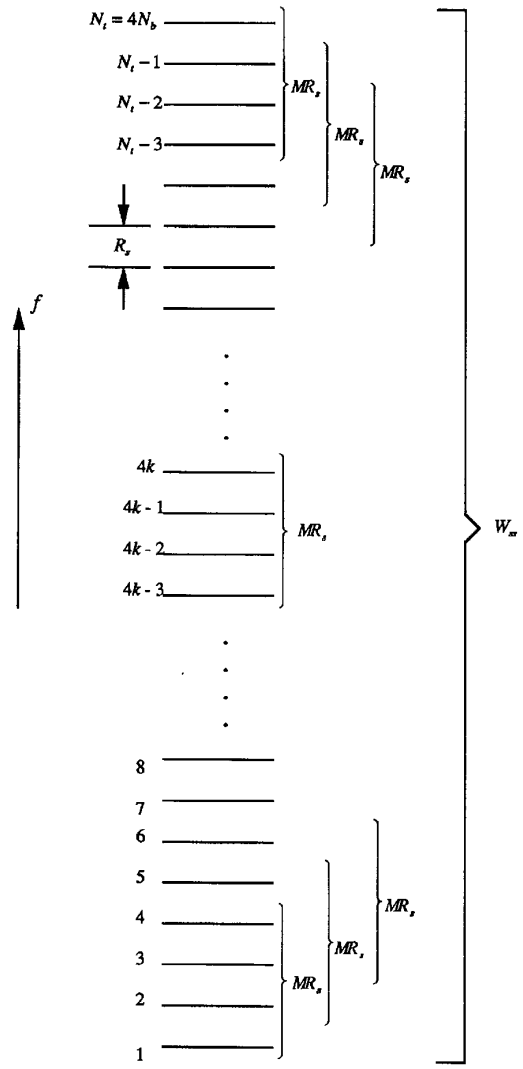


Figure 11.3b Frequency distribution for FH-4FSK—over-lapping bands.

11.4.3 Time Hopping Modulation

Time hopping (TH) is to spread spectrum modulation what pulse position modulation (PPM) is to information modulation. In particular, consider segmenting time into intervals of T_f seconds and further segment each T_f interval into M_T increments of width T_f/M_T. Assuming a pulse of maximum duration equal to T_f/M_T, then a **time hopping spread spectrum** modulation would take the form

$$c(t) = \sum_{n=-\infty}^{\infty} p\left[t - \left(n + \frac{a_n}{M_T}\right) T_f\right] \tag{11.7}$$

Ford Motor Co.
Exhibit 1030
Page 277

where a_n denotes the pseudorandom position (one of M_T uniformly spaced locations) of the pulse within the T_f-second interval.

For DS and FH, we saw that *multiplicative* modulation, that is the transmitted signal is the product of the SS and information signals, was the natural choice. For TH, *delay* modulation is the natural choice. In particular, a TH-SS modulation takes the form

$$x(t) = \mathrm{Re}\left\{c(t - d(t)) \exp\left[j\left(2\pi f_c + \phi_T\right)\right]\right\} \tag{11.8}$$

where $d(t)$ is a digital information modulation at a rate $1/T_s$ sps. Finally, the dehopping procedure at the receiver consists of removing the sequence of delays introduced by $c(t)$, which restores the information signal back to its original form and spreads the interferer.

11.4.4 Hybrid Modulations

By blending together several of the previous types of SS modulation, one can form **hybrid** modulations that, depending on the system design objectives, can achieve a better performance against the interferer than can any of the SS modulations acting alone. One possibility is to multiply several of the $c(t)$ wideband waveforms [now denoted by $c^{(i)}(t)$ to distinguish them from one another] resulting in a SS modulation of the form

$$c(t) = \prod_i c^{(i)}(t) \tag{11.9}$$

Such a modulation may embrace the advantages of the various $c^{(i)}(t)$, while at the same time mitigating their individual disadvantages.

11.5 Applications of Spread Spectrum

11.5.1 Military

Antijam (AJ) Communications

As already noted, one of the key applications of spread spectrum is for antijam communications in a hostile environment. The basic mechanism by which a **direct sequence spread spectrum** receiver attenuates a noise jammer was illustrated in Section 11.3. Therefore, in this section, we will concentrate on tone jamming.

Assume the received signal, denoted $r(t)$, is given by

$$r(t) = Ax(t) + I(t) + n_w(t) \tag{11.10}$$

where $x(t)$ is given in Eq. (11.4), A is a constant amplitude,

$$I(t) = \alpha \cos\left(2\pi f_c t + \theta\right) \tag{11.11}$$

and $n_w(t)$ is additive white Gaussian noise (AWGN) having two-sided spectral density $N_0/2$. In Eq. (11.11), α is the amplitude of the tone jammer and θ is a random phase uniformly distributed in $[0, 2\pi]$.

If we employ the standard correlation receiver of Fig. 11.4, it is straightforward to show that the final test statistic out of the receiver is given by

$$g(T_b) = AT_b + \alpha \cos\theta \int_0^{T_b} c(t)\,\mathrm{d}t + N\left(T_b\right) \tag{11.12}$$

Ford Motor Co.
Exhibit 1030
Page 278

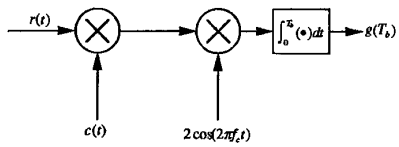


FIGURE 11.4 Standard correlation receiver.

where $N(T_b)$ is the contribution to the test statistic due to the AWGN. Noting that, for rectangular chips, we can express

$$\int_0^{T_b} c(t)\,\mathrm{d}t = T_c \sum_{i=1}^{M} c_i \tag{11.13}$$

where

$$M \triangleq \frac{T_b}{T_c} \tag{11.14}$$

is one-half of the processing gain, it is straightforward to show that, for a given value of θ, the signal-to-noise-plus-interference ratio, denoted by S/N_{total}, is given by

$$\frac{S}{N_{\text{total}}} = \frac{1}{\frac{N_0}{2E_b} + \left(\frac{J}{MS}\right)\cos^2\theta} \tag{11.15}$$

In Eq. (11.15), the jammer power is

$$J \triangleq \frac{\alpha^2}{2} \tag{11.16}$$

and the signal power is

$$S \triangleq \frac{A^2}{2} \tag{11.17}$$

If we look at the second term in the denominator of Eq. (11.15), we see that the ratio J/S is divided by M. Realizing that J/S is the ratio of the jammer power to the signal power before despreading, and J/MS is the ratio of the same quantity after despreading, we see that, as was the case for noise jamming, the benefit of employing direct sequence spread spectrum signalling in the presence of tone jamming is to reduce the effect of the jammer by an amount on the order of the processing gain.

Finally, one can show that an estimate of the average probability of error of a system of this type is given by

$$P_e = \frac{1}{2\pi}\int_0^{2\pi} \phi\left(-\sqrt{\frac{S}{N_{\text{total}}}}\right)\mathrm{d}\theta \tag{11.18}$$

where

$$\phi(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-y^2/2}\,\mathrm{d}y \tag{11.19}$$

If Eq. (11.18) is evaluated numerically and plotted, the results are as shown in Fig. 11.5. It is clear from this figure that a large initial power advantage of the jammer can be overcome by a sufficiently large value of the processing gain.

Ford Motor Co.
Exhibit 1030
Page 279

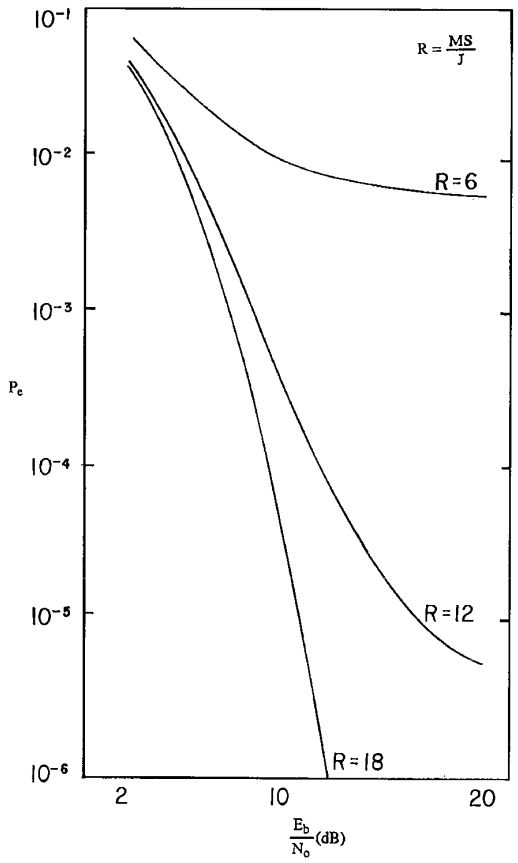


FIGURE 11.5 Plotted results of Eq. (11.18).

Low-Probability of Intercept (LPI)

The opposite side of the AJ problem is that of LPI, that is, the desire to hide your signal from detection by an intelligent adversary so that your transmissions will remain unnoticed and, thus, neither jammed nor exploited in any manner. This idea of designing an LPI system is achieved in a variety of ways, including transmitting at the smallest possible power level, and limiting the transmission time to as short an interval in time as is possible. The choice of signal design is also important, however, and it is here that spread spectrum techniques become relevant.

The basic mechanism is reasonably straightforward; if we start with a conventional narrowband signal, say a BPSK waveform having a spectrum as shown in Fig. 11.6a, and then spread it so that its new spectrum is as shown in Fig. 11.6b, the peak amplitude of the spectrum after spreading has been reduced by an amount on the order of the processing gain relative to what it was before spreading. Indeed, a sufficiently large processing gain will result in the spectrum of the signal after spreading falling below the ambient thermal noise level. Thus, there is no easy way for an unintended listener to determine that a transmission is taking place.

Ford Motor Co.
Exhibit 1030
Page 280

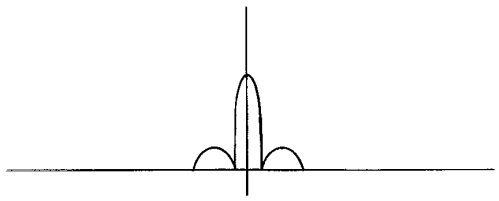


Figure 11.6a



Figure 11.6b

That is not to say the spread signal cannot be detected, however, merely that it is more difficult for an adversary to learn of the transmission. Indeed, there are many forms of so-called intercept receivers that are specifically designed to accomplish this very task. By way of example, probably the best known and simplest to implement is a **radiometer**, which is just a device that measures the total power present in the received signal. In the case of our intercept problem, even though we have lowered the power spectral density of the transmitted signal so that it falls below the noise floor, we have not lowered its power (i.e., we have merely spread its power over a wider frequency range). Thus, if the radiometer integrates over a sufficiently long period of time, it will eventually determine the presence of the transmitted signal buried in the noise. The key point, of course, is that the use of the spreading makes the interceptor's task much more difficult, since he has no knowledge of the spreading code and, thus, cannot despread the signal.

11.5.2 Commercial

Multiple Access Communications

From the perspective of commercial applications, probably the most important use of spread spectrum communications is as a multiple accessing technique. When used in this manner, it becomes an alternative to either frequency division multiple access (FDMA) or time division multiple access (TDMA) and is typically referred to as either code division multiple access (CDMA) or spread spectrum multiple access (SSMA). When using CDMA, each signal in the set is given its own spreading sequence. As opposed to either FDMA, wherein all users occupy disjoint frequency bands but are transmitted simultaneously in time, or TDMA, whereby all users occupy the same bandwidth but transmit in disjoint intervals of time, in CDMA, all signals occupy the same bandwidth and are transmitted simultaneously in time; the different waveforms in CDMA are distinguished from one another at the receiver by the specific spreading codes they employ.

Since most CDMA detectors are correlation receivers, it is important when deploying such a system to have a set of spreading sequences that have relatively low-pairwise cross-correlation between any two sequences in the set. Further, there are two fundamental types of operation in CDMA, synchronous and asynchronous. In the former case, the symbol transition times of all of the users are aligned; this allows for orthogonal sequences to be used as the spreading sequences and, thus, eliminates interference from one user to another.

Ford Motor Co.
Exhibit 1030
Page 281

Alternately, if no effort is made to align the sequences, the system operates asychronously; in this latter mode, multiple access interference limits the ultimate channel capacity, but the system design exhibits much more flexibility.

CDMA has been of particular interest recently for applications in wireless communications. These applications include cellular communications, personal communications services (PCS), and wireless local area networks. The reason for this popularity is primarily due to the performance that spread spectrum waveforms display when transmitted over a multipath fading channel.

To illustrate this idea, consider DS signalling. As long as the duration of a single chip of the spreading sequence is less than the multipath delay spread, the use of DS waveforms provides the system designer with one of two options. First, the multipath can be treated as a form of interference, which means the receiver should attempt to attenuate it as much as possible. Indeed, under this condition, all of the multipath returns that arrive at the receiver with a time delay greater than a chip duration from the multipath return to which the receiver is synchronized (usually the first return) will be attenuated because of the processing gain of the system.

Alternately, the multipath returns that are separated by more than a chip duration from the main path represent independent "looks" at the received signal and can be used constructively to enhance the overall performance of the receiver. That is, because all of the multipath returns contain information regarding the data that is being sent, that information can be extracted by an appropriately designed receiver. Such a receiver, typically referred to as a RAKE receiver, attempts to resolve as many individual multipath returns as possible and then to sum them coherently. This results in an *implicit* diversity gain, comparable to the use of *explicit* diversity, such as receiving the signal with multiple antennas.

The condition under which the two options are available can be stated in an alternate manner. If one envisions what is taking place in the frequency domain, it is straightforward to show that the condition of the chip duration being smaller than the multipath delay spread is equivalent to requiring that the spread bandwidth of the transmitted waveform exceed what is called the coherence bandwidth of the channel. This latter quantity is simply the inverse of the multipath delay spread and is a measure of the range of frequencies that fade in a highly correlated manner. Indeed, anytime the coherence bandwidth of the channel is less than the spread bandwidth of the signal, the channel is said to be *frequency selective* with respect to the signal. Thus, we see that to take advantage of DS signalling when used over a multipath fading channel, that signal should be designed such that it makes the channel appear frequency selective.

In addition to the desirable properties that spread spectrum signals display over multipath channels, there are two other reasons why such signals are of interest in cellular-type applications. The first has to do with a concept known as the reuse factor. In conventional cellular systems, either analog or digital, in order to avoid excessive interference from one cell to its neighbor cells, the frequencies used by a given cell are not used by its immediate neighbors (i.e., the system is designed so that there is a certain spatial separation between cells that use the same carrier frequencies). For CDMA, however, such spatial isolation is typically not needed, so that so-called *universal reuse* is possible.

Further, because CDMA systems tend to be interference limited, for those applications involving voice transmission, an additional gain in the capacity of the system can be achieved by the use of *voice activity detection*. That is, in any given two-way telephone conversation, each user is typically talking only about 50% of the time. During the time when a user is quiet, he is not contributing to the instantaneous interference. Thus, if a sufficiently large number of users can be supported by the system, statistically only about one-half of them will be active simultaneously, and the effective capacity can be doubled.

Ford Motor Co.
Exhibit 1030
Page 282

Interference Rejection

In addition to providing multiple accessing capability, spread spectrum techniques are of interest in the commercial sector for basically the same reasons they are in the military community, namely their AJ and LPI characteristics. However, the motivations for such interest differ. For example, whereas the military is interested in ensuring that systems they deploy are robust to interference generated by an intelligent adversary (i.e., exhibit jamming resistance), the interference of concern in commercial applications is unintentional. It is sometimes referred to as cochannel interference (CCI) and arises naturally as the result of many services using the same frequency band at the same time. And while such scenarios almost always allow for some type of spatial isolation between the interfering waveforms, such as the use of narrow-beam antenna patterns, at times the use of the inherent interference suppression property of a spread spectrum signal is also desired. Similarly, whereas the military is very much interested in the LPI property of a spread spectrum waveform, as indicated in Section 11.3, there are applications in the commercial segment where the same characteristic can be used to advantage.

To illustrate these two ideas, consider a scenario whereby a given band of frequencies is somewhat sparsely occupied by a set of conventional (i.e., nonspread) signals. To increase the overall spectral efficiency of the band, a set of spread spectrum waveforms can be overlaid on the same frequency band, thus forcing the two sets of users to share common spectrum. Clearly, this scheme is feasible only if the mutual interference that one set of users imposes on the other is within tolerable limits. Because of the interference suppression properties of spread spectrum waveforms, the despreading process at each spread spectrum receiver will attenuate the components of the final test statistic due to the overlaid narrowband signals. Similarly, because of the LPI characteristics of spread spectrum waveforms, the increase in the overall noise level as seen by any of the conventional signals, due to the overlay, can be kept relatively small.

Defining Terms

Antijam communication system: A communication system designed to resist intentional jamming by the enemy.

Chip time (interval): The duration of a single pulse in a direct sequence modulation; typically much smaller than the information symbol interval.

Coarse alignment: The process whereby the received signal and the despreading signal are aligned to within a single chip interval.

Dehopping: Despreading using a frequency-hopping modulation.

Delay-locked loop: A particular implementation of a closed-loop technique for maintaining fine alignment.

Despreading: The notion of decreasing the bandwidth of the received (spread) signal back to its information bandwidth.

Direct sequence modulation: A signal formed by linearly modulating the output sequence of a pseudorandom number generator onto a train of pulses.

Direct sequence spread spectrum: A spreading technique achieved by multiplying the information signal by a direct sequence modulation.

Fast frequency-hopping: A spread spectrum technique wherein the hop time is less than or equal to the information symbol interval, i.e., there exist one or more hops per data symbol.

Fine alignment: The state of the system wherein the received signal and the despreading signal are aligned to within a small fraction of a single chip interval.

Frequency-hopping modulation: A signal formed by nonlinearly modulating a train of pulses with a sequence of pseudorandomly generated frequency shifts.

Hop time (interval): The duration of a single pulse in a frequency-hopping modulation.

Hybrid spread spectrum: A spreading technique formed by blending together several spread spectrum techniques, e.g., direct sequence, frequency-hopping, etc.

Low-probability-of-intercept communication system: A communication system designed to operate in a hostile environment wherein the enemy tries to detect the presence and perhaps characteristics of the friendly communicator's transmission.

Processing gain (spreading ratio): The ratio of the spread spectrum bandwidth to the information data rate.

Radiometer: A device used to measure the total energy in the received signal.

Search algorithm: A means for coarse aligning (synchronizing) the despreading signal with the received spread spectrum signal.

Slow frequency-hopping: A spread spectrum technique wherein the hop time is greater than the information symbol interval, i.e., there exists more than one data symbol per hop.

Spread spectrum bandwidth: The bandwidth of the transmitted signal after spreading.

Spreading: The notion of increasing the bandwidth of the transmitted signal by a factor far in excess of its information bandwidth.

Tau–dither loop: A particular implementation of a closed-loop technique for maintaining fine alignment.

Time-hopping spread spectrum: A spreading technique that is analogous to pulse position modulation.

Tracking algorithm: An algorithm (typically closed loop) for maintaining fine alignment.

References

[1] Cook, C.F., Ellersick, F.W., Milstein, L.B., and Schilling, D.L., *Spread Spectrum Communications*, IEEE Press, 1983.

[2] Dixon, R.C., *Spread Spectrum Systems*, 3rd ed., John Wiley and Sons, Inc. 1994.

[3] Holmes, J.K., *Coherent Spread Spectrum Systems*, John Wiley and Sons, Inc. 1982.

[4] Simon, M.K., Omura, J.K., Scholtz, R.A., and Levitt, B.K., *Spread Spectrum Communications Handbook*, McGraw Hill, 1994 (previously published as *Spread Spectrum Communications*, Computer Science Press, 1985).

[5] Ziemer, R.E. and Peterson, R.L., *Digital Communications and Spread Spectrum Techniques*, Macmillan, 1985.

Ford Motor Co.
Exhibit 1030
Page 284

12

Diversity

12.1 Introduction 12-1
12.2 Diversity Schemes 12-3
Space Diversity • Polarization Diversity • Angle Diversity • Frequency Diversity • Path Diversity • Time Diversity • Transformed Diversity
12.3 Diversity Combining Techniques 12-6
Selection Combining • Maximal Ratio Combining • Equal Gain Combining • Loss of Diversity Gain Due to Branch Correlation and Unequal Branch Powers
12.4 Effect of Diversity Combining on Bit Error Rate 12-10
12.5 Concluding Remarks 12-11
Defining Terms 12-11
References 12-11

Arogyaswami J. Paulraj
Stanford University

12.1 Introduction

Diversity is a commonly used technique in mobile radio systems to combat signal **fading.** The basic principle of diversity is as follows. If several replicas of the same information-carrying signal are received over multiple channels with comparable strengths, which exhibit independent fading, then there is a good likelihood that at least one or more of these received signals will not be in a fade at any given instant in time, thus making it possible to deliver adequate signal level to the receiver. Without diversity techniques, in noise limited conditions, the transmitter would have to deliver a much higher power level to protect the link during the short intervals when the channel is severely faded. In mobile radio, the power available on the reverse link is severely limited by the battery capacity of hand-held subscriber units. Diversity methods play a crucial role in reducing transmit power needs. Also, cellular communication networks are mostly interference limited and, once again, mitigation of channel fading through use of diversity can translate into reduced variability of carrier-to-interference ratio (C/I), which in turn means lower C/I margin and hence better reuse factors and higher system capacity.

The basic principles of diversity have been known since 1927 when the first experiments in space diversity were reported. There are many techniques for obtaining independently fading branches, and these can be subdivided into two main classes. The first are explicit techniques where explicit redundant signal transmission is used to exploit diversity channels. Use of dual polarized signal transmission and reception in many point-to-point radios is an example of explicit diversity. Clearly such redundant signal transmission involves a penalty in frequency spectrum or additional power. In the second class are implicit diversity tech-

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 285

niques: the signal is transmitted only once, but the decorrelating effects in the propagation medium such as multipaths are exploited to receive signals over multiple diversity channels. A good example of implicit diversity is the **RAKE receiver** in code division multiple access (CDMA) systems, which uses independent fading of resolvable multipaths to achieve diversity gain. Figure 12.1 illustrates the principle of diversity where two independently fading signals are shown along with the selection diversity output signal which selects the stronger signal. The fades in the resulting signal have been substantially smoothed out while also yielding higher average power.

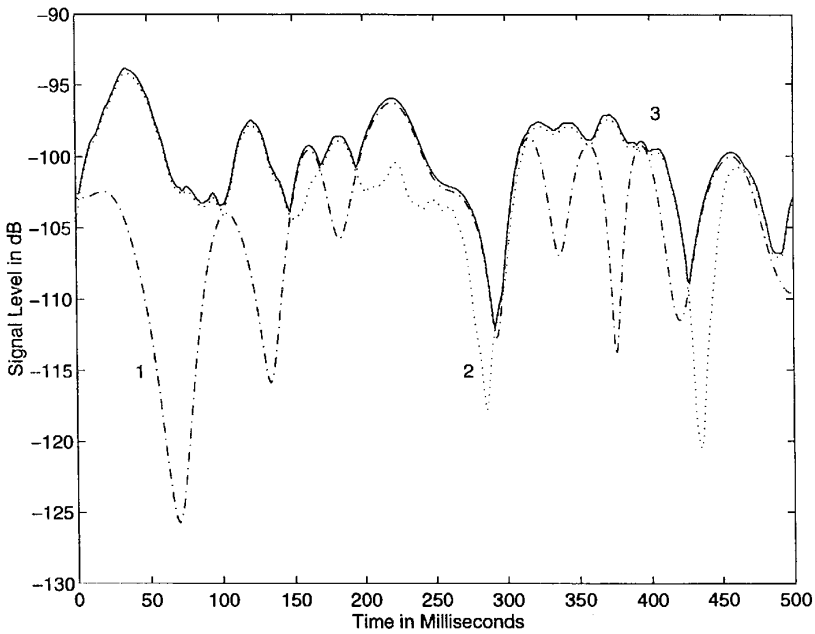


FIGURE 12.1 Example of diversity combining. Two independently fading signals 1 and 2. The signal 3 is the result of selecting the strongest signal.

If antennas are used in transmit, they can be exploited for diversity. If the transmit channel is known, the antennas can be driven with complex conjugate channel weighting to co-phase the signals at the receive antenna. If the forward channel is not known, we have several methods to convert space selective fading at the transmit antennas to other forms of diversity exploitable in the receiver.

Exploiting diversity needs careful design of the communication link. In explicit diversity, multiple copies of the same signal are transmitted in channels using either a frequency, time, or polarization dimension. At the receiver end we need arrangements to receive the different diversity branches (this is true for both explicit and implicit diversity). The different diversity branches are then combined to reduce signal **outage probability** or bit error rate.

In practice, the signals in the diversity branches may not show completely independent fading. The envelope cross correlation ρ between these signals is a measure of their inde-

Ford Motor Co.
Exhibit 1030
Page 286

pendence.

$$\rho = \frac{E\left[[r_1 - \bar{r}_1]\,[r_2 - \bar{r}_2]\right]}{\sqrt{E\,|r_1 - \bar{r}_1|^2\,E\,|r_2 - \bar{r}_2|^2}}$$

where r_1 and r_2 represent the instantaneous envelope levels of the normalized signals at the two receivers and $\bar{r_1}$ and $\bar{r_2}$ are their respective means. It has been shown that a cross correlation of 0.7 [3] between signal envelopes is sufficient to provide a reasonable degree of diversity gain. Depending on the type of diversity employed, these diversity channels must be sufficiently *separated* along the appropriate diversity dimension. For spatial diversity, the antennas should be separated by more than the *coherence distance* to ensure a cross correlation of less than 0.7. Likewise in frequency diversity, the frequency separation must be larger than the *coherence bandwidth,* and in time diversity the separation between channel reuse in time should be longer than the *coherence time.* These coherence factors in turn depend on the channel characteristics. The coherence distance, coherence bandwidth and coherence time vary inversely as the angle spread, delay spread, and Doppler spread, respectively.

If the receiver has a number of diversity branches, it has to combine these branches to maximize the signal level. Several techniques have been studied for diversity combining. We will describe three main techniques: selection combining, equal gain combining, and maximal ratio combining.

Finally, we should note that diversity is primarily used to combat fading and if the signal does not show significant fading in the first place, for example when there is a direct path component, diversity combining may not provide significant diversity gain. In the case of antenna diversity, array gain proportional to the number of antennas will still be available.

12.2 Diversity Schemes

There are several techniques for obtaining diversity branches, sometimes also known as diversity dimensions. The most important of these are discussed in the following sections.

12.2.1 Space Diversity

This has historically been the most common form of diversity in mobile radio base stations. It is easy to implement and does not require additional frequency spectrum resources. Space diversity is exploited on the reverse link at the base station receiver by spacing antennas apart so as to obtain sufficient decorrelation. The key for obtaining minimum uncorrelated fading of antenna outputs is adequate spacing of the antennas. The required spacing depends on the degree of multipath angle spread. For example if the multipath signals arrive from all directions in the azimuth, as is usually the case at the mobile, antenna spacing (coherence distance) of the order of 0.5λ to 0.8λ is quite adequate [5]. On the other hand if the multipath angle spread is small, as in the case of base stations, the coherence distance is much larger. Also empirical measurements show a strong coupling between antenna height and spatial correlation. Larger antenna heights imply larger coherence distances. Typically 10λ to 20λ separation is adequate to achieve $\rho = 0.7$ at base stations in suburban settings when the signals arrive from the broadside direction. The coherence distance can be 3 to 4 times larger for endfire arrivals. The endfire problem is averted in base stations with trisectored antennas as each sector needs to handle only signals arriving $\pm 60°$ off the broadside. The coherence distance depends strongly on the terrain. Large multipath angle spread means smaller coherence distance. Base stations normally use space

Ford Motor Co.
Exhibit 1030
Page 287

diversity in the horizontal plane only. Separation in the vertical plane can also be used, and the necessary spacing depends upon vertical multipath angle spread. This can be small for distant mobiles making vertical plane diversity less attractive in most applications.

Space diversity is also exploitable at the transmitter. If the forward channel is known, it works much like receive space diversity. If it is not known, then space diversity can be transformed to another form of diversity exploitable at the receiver. (See Section 12.2.7 below).

If antennas are used at transmit and receive, the M transmit and N receive antennas both contribute to diversity. It can be shown that if simple weighting is used without additional bandwidth or time/memory processing, then maximum diversity gain is obtained if the transmitter and receiver use the left and right singular vectors of the $M \times N$ channel matrix, respectively. However, to approach the maximum $M \times N$ order diversity order will require the use of additional bandwidth or time/memory-based methods.

12.2.2 Polarization Diversity

In mobile radio environments, signals transmitted on orthogonal polarizations exhibit low fade correlation, and therefore, offer potential for diversity combining. Polarization diversity can be obtained either by explicit or implicit techniques. Note that with polarization only two diversity branches are available as against space diversity where several branches can be obtained using multiple antennas. In explicit polarization diversity, the signal is transmitted and received in two orthogonal polarizations. For a fixed total transmit power, the power in each branch will be 3 dB lower than if single polarization is used. In the implicit polarization technique, the signal is launched in a single polarization, but is received with cross-polarized antennas. The propagation medium couples some energy into the cross-polarization plane. The observed cross-polarization coupling factor lies between 8 to 12 dB in mobile radio [8, 1]. The cross-polarization envelope decorrelation has been found to be adequate. However, the large branch imbalance reduces the available diversity gain.

With hand-held phones, the handset can be held at random orientations during a call. This results in energy being launched with varying polarization angles ranging from vertical to horizontal. This further increases the advantage of cross-polarized antennas at the base station since the two antennas can be combined to match the received signal polarization. This makes polarization diversity even more attractive. Recent work [4] has shown that with variable launch polarization, a cross-polarized antenna can give comparable overall (matching plus diversity) performance to a vertically polarized space diversity antenna.

Finally, we should note that cross-polarized antennas can be deployed in a compact antenna assembly and do not need large physical separation needed in space diversity antennas. This is an important advantage in the **PCS** base stations where low profile antennas are needed.

12.2.3 Angle Diversity

In situations where the angle spread is very high, such as indoors or at the mobile unit in urban locations, signals collected from multiple nonoverlapping beams offer low fade correlation with balanced power in the diversity branches. Clearly, since directional beams imply use of antenna aperture, angle diversity is closely related to space diversity. Angle diversity has been utilized in indoor wireless LANs, where its use allows substantial increase in LAN throughputs [2].

Ford Motor Co.
Exhibit 1030
Page 288

12.2.4 Frequency Diversity

Another technique to obtain decorrelated diversity branches is to transmit the same signal over different frequencies. The frequency separation between carriers should be larger than the coherence bandwidth. The coherence bandwidth, of course, depends on the multipath delay spread of the channel. The larger the delay spread, the smaller the coherence bandwidth and the more closely we can space the frequency diversity channels. Clearly, frequency diversity is an explicit diversity technique and needs additional frequency spectrum.

A common form of frequency diversity is multicarrier (also known as multitone) modulation. This technique involves sending redundant data over a number of closely spaced carriers to benefit from frequency diversity, which is then exploited by applying **interleaving** and **channel coding/forward error correction** across the carriers. Another technique is to use **frequency hopping** wherein the interleaved and channel coded data stream is transmitted with widely separated frequencies from burst to burst. The wide frequency separation is chosen to guarantee independent fading from burst to burst.

12.2.5 Path Diversity

This implicit diversity is available if the signal bandwidth is much larger than the channel coherence bandwidth. The basis for this method is that when the multipath arrivals can be resolved in the receiver and since the paths fade independently, diversity gain can be obtained. In CDMA systems, the multipath arrivals must be separated by more than one *chip* period and the RAKE receiver provides the diversity [9]. In TDMA systems, the multipath arrivals must be separated by more than one *symbol* period and the MLSE receiver provides the diversity.

12.2.6 Time Diversity

In mobile communications channels, the mobile motion together with scattering in the vicinity of the mobile causes time selective fading of the signal with Rayleigh fading statistics for the signal envelope. Signal fade levels separated by the *coherence time* show low correlation and can be used as diversity branches if the same signal can be transmitted at multiple instants separated by the coherence time. The coherence time depends on the Doppler spread of the signal, which in turn is a function of the mobile speed and the carrier frequency.

Time diversity is usually exploited via interleaving, forward-error correction (FEC) coding, and **automatic request for repeat** (ARQ). These are sophisticated techniques to exploit channel coding and time diversity. One fundamental drawback with time diversity approaches is the delay needed to collect the repeated or interleaved transmissions. If the coherence time is large, as for example when the vehicle is slow moving, the required delay becomes too large to be acceptable for interactive voice conversation.

The statistical properties of fading signals depend on the field component used by the antenna, the vehicular speed, and the carrier frequency. For an idealized case of a mobile surrounded by scatterers in all directions, the autocorrelation function of the received signal $x(t)$ (note this is not the envelope $r(t)$) can be shown to be

$$E\left[x(t)x(t+\tau)\right] = J_0\left(2\pi\tau v/\lambda\right)$$

where J_0 is a Bessel function of the 0th order and v is the mobile velocity.

Ford Motor Co.
Exhibit 1030
Page 289

12.2.7 Transformed Diversity

In transformed diversity, the space diversity branches at the transmitter are transformed into other forms of diversity branches exploitable at the receiver. This is used when the forward channel is not known and shifts the responsibility of diversity combining to the receiver which has the necessary channel knowledge.

Space to Frequency

- *Antenna-delay.* Here the signal is transmitted from two or more antennas with delays of the order of a chip or symbol period in CDMA or TDMA, respectively. The different transmissions simulate resolved path arrivals that can be used as diversity branches by the RAKE or MLSE equalizer.
- *Multicarrier modulation.* The data stream after interleaving and coding is modulated as a multicarrier output using an inverse DFT. The carriers are then mapped to the different antennas. The space selective fading at the antennas is now transformed to frequency selective fading and diversity is obtained during decoding.

Space to Time

- *Antenna hopping/phase rolling.* In this method the data stream after coding and interleaving is switched randomly from antenna to antenna. The space selective fading at the transmitter is converted into a time selective fading at the receiver. This is a form of "active" fading.
- *Space-time coding.* The approach in space-time coding is to split the encoded data into multiple data streams each of which is modulated and simultaneously transmitted from different antennas. The received signal is a superposition of the multiple transmitted signals. Channel decoding can be used to recover the data sequence. Since the encoded data arrive over uncorrelated fade branches, diversity gain can be realized.

12.3 Diversity Combining Techniques

Several diversity combining methods are known. We describe three main techniques: selection, maximal ratio, and equal gain. They can be used with each of the diversity schemes discussed above.

12.3.1 Selection Combining

This is the simplest and perhaps the most frequently used form of diversity combining. In this technique, one of the two diversity branches with the highest carrier-to-noise ratio (C/N) is connected to the output. See Fig. 12.2(a).

The performance improvement due to selection diversity can be seen as follows. Let the signal in each branch exhibit Rayleigh fading with mean power σ^2. The density function of the envelope is given by

$$p(r_i) = \frac{r_i}{\sigma^2} e^{\frac{-r_i^2}{2\sigma^2}} \tag{12.1}$$

Ford Motor Co.
Exhibit 1030
Page 290

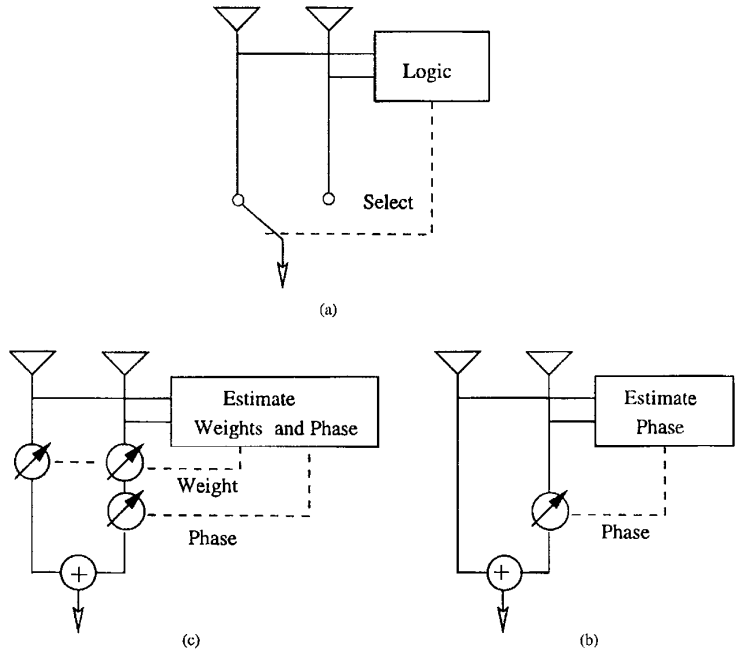


FIGURE 12.2 Diversity combining methods for two diversity branches.

where r_i is the signal envelope in each branch. If we define two new variables

$$\gamma_i = \frac{\text{Instantaneous signal power in each branch}}{\text{Mean noise power}}$$

$$\Gamma = \frac{\text{Mean signal power in each branch}}{\text{Mean noise power}}$$

then the probability that the C/N is less than or equal to some specified value γ_s is

$$\text{Prob}\left[\gamma_i \leq \gamma_s\right] = 1 - e^{-\gamma_s/\Gamma} \tag{12.2}$$

The probability that γ_i in all branches with independent fading will be simultaneously less than or equal to γ_s is then

$$\text{Prob}\left[\gamma_1, \gamma_2, \ldots \gamma_M \leq \gamma_s\right] = \left(1 - e^{-\gamma_s/\Gamma}\right)^M \tag{12.3}$$

This is the distribution of the best signal envelope from the two diversity branches. Figure 12.3 shows the distribution of the combiner output C/N for M = 1,2,3, and 4 branches. The improvement in signal quality is significant. For example at 99% reliability level, the improvement in C/N is 10 dB for two branches and 16 dB for four branches.

Selection combining also increases the mean C/N of the combiner output and can be shown to be [3]

$$\text{Mean}\left(\gamma_s\right) = \Gamma \sum_{k=1}^{M} \frac{1}{k} \tag{12.4}$$

This indicates that with 4 branches, for example, the mean C/N of the selected branch is 2.08 better than the mean C/N in any one branch.

Ford Motor Co.
Exhibit 1030
Page 291

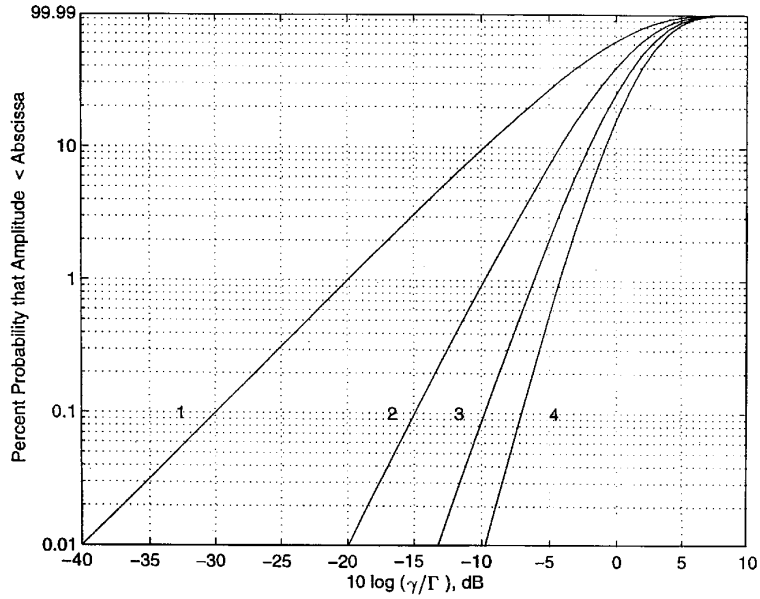


FIGURE 12.3 Probability distribution of signal envelope for selection combining.

12.3.2 Maximal Ratio Combining

In this technique the M diversity branches are first co-phased and then weighted proportionally to their signal level before summing. See Fig. 12.2(b). The distribution of the maximal ratio combiner has been shown to be [5]

$$\text{Prob}\,[\gamma \leq \gamma_m] = 1 - e^{(-\gamma_m/\Gamma)} \sum_{k=1}^{M} \frac{(\gamma_m/\Gamma)^{k-1}}{(k-1)!} \tag{12.5}$$

The distribution of output of a maximal ratio combiner is shown in Fig. 12.4. Maximal ratio combining is known to be optimal in the sense that it yields the best statistical reduction of fading of any linear diversity combiner. In comparison to the selection combiner, at 99% reliability level, the maximal ratio combiner provides a 11.5 dB gain for two branches and a 19 dB gain for four branches, an improvement of 1.5 and 3 dB, respectively, over the selection diversity combiner.

The mean C/N of the combined signal may be easily shown to be

$$\text{Mean}\,(\gamma_m) = M\Gamma \tag{12.6}$$

Therefore, combiner output mean varies linearly with M. This confirms the intuitive result that the output C/N averaged over fades should provide gain proportional to the number of diversity branches. This is a situation similar to conventional beamforming.

12.3.3 Equal Gain Combining

In some applications, it may be difficult to estimate the amplitude accurately, the combining gains may all be set to unity, and the diversity branches merely summed after co-phasing. [See Fig. 12.2(c)].

Ford Motor Co.
Exhibit 1030
Page 292

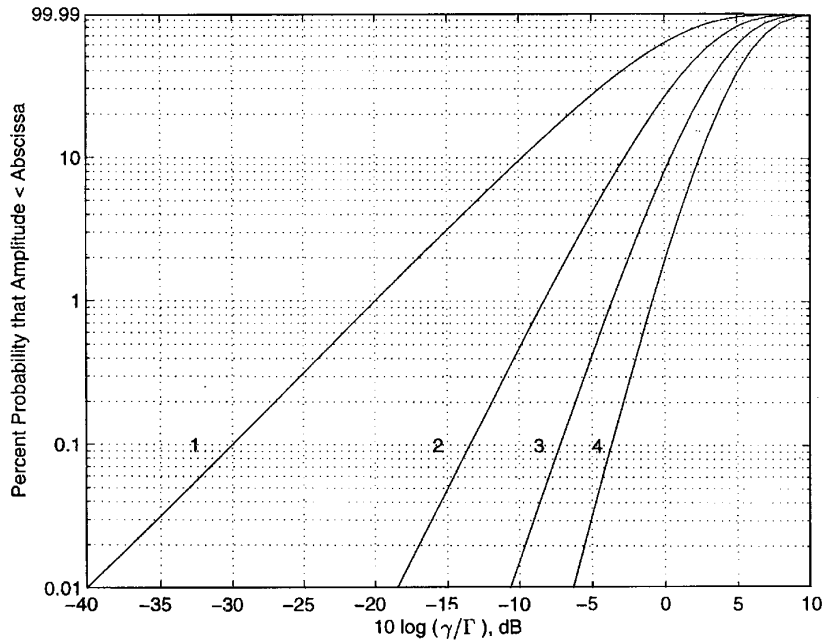


FIGURE 12.4 Probability distribution for signal envelope for maximal ratio combining.

The distribution of equal gain combiner does not have a neat expression and has been computed by numerical evaluation. Its performance has been shown to be very close to within a decibel to maximal ratio combining. The mean C/N can be shown to be [3]

$$\text{Mean}\,(\gamma_e) = \Gamma\left[1 + \frac{\pi}{4}(M-1)\right] \tag{12.7}$$

Like maximal ratio combining, the mean C/N for equal gain combining grows almost linearly with M and is approximately only one decibel poorer than maximal ratio combiner even with an infinite number of branches.

12.3.4 Loss of Diversity Gain Due to Branch Correlation and Unequal Branch Powers

The above analysis assumed that the fading signals in the diversity branches were all uncorrelated and of equal power. In practice, this may be difficult to achieve and as we saw earlier, the branch cross-correlation coefficient $\rho = 0.7$ is considered to be acceptable. Also, equal mean powers in diversity branches are rarely available. In such cases we can expect a certain loss of diversity gain. However, since most of the damage in fading is due to deep fades, and also since the chance of coincidental deep fades is small even for moderate branch correlation, one can expect a reasonable tolerance to branch correlation.

The distribution of the output signal envelope of maximal ratio combiner has been shown to be [6]:

$$\text{Prob}\,[\gamma_m] = \sum_{n=1}^{M} \frac{A_n}{2\lambda_n} e^{-\gamma_m/2\lambda_n} \tag{12.8}$$

where λ_n are the eigenvalues of the M $\times$ M branch envelope covariance matrix whose elements are defined by

$$\mathbf{R}_{ij} = E\left[r_i r_j^*\right] \tag{12.9}$$

and A_n is defined by

$$A_n = \prod_{\substack{k=1 \\ k \neq n}}^{M} \frac{1}{1 - \lambda_k/\lambda_n} \tag{12.10}$$

12.4 Effect of Diversity Combining on Bit Error Rate

So far we have studied the distribution of the instantaneous envelope or C/N after diversity combining. We will now briefly survey how diversity combining affects BER performance in digital radio links; we assume maximal ratio combining.

To begin let us first examine the effect of Rayleigh fading on the BER performance of digital transmission links. This has been studied by several authors and is summarized in [7]. Table 12.1 gives the BER expressions in the large E_b/N_0 case for coherent binary PSK and coherent binary orthogonal FSK for unfaded and Rayleigh faded AWGN (additive white Gaussian noise channels) channels. $\bar{E}_b/N_0$ represents the average E_b/N_0 for the fading channel.

TABLE 12.1 Comparison of BER Performance for Unfaded and Rayleigh Faded Signals

Modulaton	Unfaded BER	Faded BER
Coh BPSK	$\frac{1}{2}\text{erfc}\left(\sqrt{E_b/N_0}\right)$	$\frac{1}{4\left(\bar{E}_b/N_0\right)}$
Coh FSK	$\frac{1}{2}\text{erfc}\left(\sqrt{\frac{1}{2}E_b/N_0}\right)$	$\frac{1}{2\left(\bar{E}_b/N_0\right)}$

Observe that error rates decrease only inversely with SNR as against exponential decreases for the unfaded channel. Also note that for fading channels, coherent binary PSK is 3 dB better than coherent binary FSK, exactly the same advantage as in unfaded case. Even for modest target BER of 10^{-2} that is usually needed in mobile communications, the loss due to fading can be very high—17.2 dB.

To obtain the BER with maximal ratio diversity combining we have to average the BER expression for the unfaded BER with the distribution obtained for the maximal ratio combiner given in (12.5). Analytical expressions have been derived for these in [7]. For a branch SNR greater than 10 dB, the BER after maximal ratio diversity combining is given in Table 12.2.

We observe that the probability of error varies as $1/\bar{E}_b/N_0$ raised to the Lth power. Thus, diversity reduces the error rate exponentially as the number of independent branches increases.

Ford Motor Co.
Exhibit 1030
Page 294

TABLE 12.2 BER Performance for Coherent BPSK and FSK with Diversity

Modulaton	Post Diversity BER
Coherent BPSK	$\left(\frac{1}{4\ \bar{E}_b/N_0}\right)^L \binom{2L-1}{L}$
Coherent FSK	$\left(\frac{1}{2\bar{E}_b/N_0}\right)^L \binom{2L-1}{L}$

12.5 Concluding Remarks

Diversity provides a powerful technique for combating fading in mobile communication systems. Diversity techniques seek to generate and exploit multiple branches over which the signal shows low fade correlation. To obtain the best diversity performance, the multiple access, modulation, coding and antenna design of the wireless link must all be carefully chosen so as to provide a rich and reliable level of well-balanced, low-correlation diversity branches in the target propagation environment. Successful diversity exploitation can impact a mobile network in several ways. Reduced power requirements can result in increased coverage or improved battery life. Low signal outage improves voice quality and handoff performance. Finally, reduced fade margins directly translate to better reuse factors and, hence, increased system capacity.

Defining Terms

Automatic request for repeat: An error control mechanism in which received packets that cannot be corrected are retransmitted.

Channel coding/Forward error correction: A technique that inserts redundant bits during transmission to help detect and correct bit errors during reception.

Fading: Fluctuation in the signal level due to shadowing and multipath effects.

Frequency hopping: A technique where the signal bursts are transmitted at different frequencies separated by random spacing that are multiples of signal bandwidth.

Interleaving: A form of data scrambling that spreads burst of bit errors evenly over the received data allowing efficient forward error correction.

Outage probability: The probability that the signal level falls below a specified minimum level.

PCS: Personal Communications Services.

RAKE receiver: A receiver used in direct sequence spread spectrum signals. The receiver extracts energy in each path and then adds them together with appropriate weighting and delay.

References

[1] Adachi, F., Feeney, M.T., Williason, A.G., and Parsons, J.D., Crosscorrelation between the envelopes of 900 MHz signals received at a mobile radio base station site. *Proc. IEE,* 133(6), 506–512, 1986.

[2] Freeburg, T.A., Enabling technologies for in-building network communications—four technical challenges and four solutions. *IEEE Trans. Veh. Tech.,* 29(4), 58–64, 1991.

[3] Jakes, W.C., *Microwave Mobile Communications,* John Wiley & Sons, New York, 1974.

[4] Jefford, P.A., Turkmani, A.M.D., Arowojulu, A.A., and Kellet, C.J., An experimental evaluation of the performance of the two branch space and polarization schemes at 1800 MHz. *IEEE Trans. Veh. Tech.*, VT-44(2), 318–326, 1995.

[5] Lee, W.C.Y., *Mobile Communications Engineering*, McGraw-Hill, New York, 1982.

[6] Pahlavan, K. and Levesque, A.H., *Wireless Information Networks*, John Wiley & Sons, New York, 1995.

[7] Proakis, J.G., *Digital Communications*, McGraw-Hill, New York, 1989.

[8] Vaughan, R.G., Polarization diversity system in mobile communications. *IEEE Trans. Veh. Tech.*, VT-39(3), 177–186, 1990.

[9] Viterbi, A.J., *CDMA: Principle of Spread Spectrum Communications*, Addison-Wesley, Reading, MA, 1995.

Ford Motor Co.
Exhibit 1030
Page 296

13

Digital Communication System Performance[1]

13.1 Introduction 13-1
The Channel • The Link
13.2 Bandwidth and Power Considerations 13-3
The Bandwidth Efficiency Plane • M-ary Signalling • Bandwidth-Limited Systems • Power-Limited Systems • Minimum Bandwidth Requirements for MPSK and MFSK Signalling
13.3 Example 1: Bandwidth-Limited Uncoded System 13-7
Solution to Example 1
13.4 Example 2: Power-Limited Uncoded System 13-9
Solution to Example 2
13.5 Example 3: Bandwidth-Limited Power-Limited Coded System 13-10
Solution to Example 3 • Calculating Coding Gain
13.6 Example 4: Direct-Sequence Spread-Spectrum Coded System 13-15
Processing Gain • Channel Parameters for Example 13.4 • Solution to Example 13.4
13.7 Conclusion 13-19
Appendix: Received E_b/N_0 Is Independent of the Code Parameters 13-19
References 13-20
Further Information 13-20

Bernard Sklar
Communications Engineering Services

13.1 Introduction

In this section we examine some fundamental tradeoffs among bandwidth, power, and error performance of digital communication systems. The criteria for choosing modulation and coding schemes, based on whether a system is bandwidth limited or power limited, are reviewed for several system examples. Emphasis is placed on the subtle but straightforward relationships we encounter when transforming from data-bits to channel-bits to symbols to chips.

[1] A version of this chapter has appeared as a paper in the *IEEE Communications Magazine,* November 1993, under the title "Defining, Designing, and Evaluating Digital Communication Systems."

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 297

The design or definition of any digital communication system begins with a description of the communication link. The *link* is the name given to the communication transmission path from the modulator and transmitter, through the channel, and up to and including the receiver and demodulator. The *channel* is the name given to the propagating medium between the transmitter and receiver. A link description quantifies the average signal power that is received, the available bandwidth, the noise statistics, and other impairments, such as fading. Also needed to define the system are basic requirements, such as the data rate to be supported and the error performance.

13.1.1 The Channel

For radio communications, the concept of *free space* assumes a channel region free of all objects that might affect radio frequency (RF) propagation by absorption, reflection, or refraction. It further assumes that the atmosphere in the channel is perfectly uniform and nonabsorbing, and that the earth is infinitely far away or its reflection coefficient is negligible. The RF energy arriving at the receiver is assumed to be a function of distance from the transmitter (simply following the inverse-square law of optics). In practice, of course, propagation in the atmosphere and near the ground results in refraction, reflection, and absorption, which modify the free space transmission.

13.1.2 The Link

A radio transmitter is characterized by its average output signal power P_t and the gain of its transmitting antenna G_t. The name given to the product $P_t G_t$, with reference to an isotropic antenna is *effective radiated power* (*EIRP*) in watts (or dBW). The predetection average signal power S arriving at the output of the receiver antenna can be described as a function of the *EIRP*, the gain of the receiving antenna G_r, the path loss (or space loss) L_s, and other losses, L_o, as follows [14, 15]:

$$S = \frac{EIRP\, G_r}{L_s L_o} \tag{13.1}$$

The path loss L_s can be written as follows [15]:

$$L_s = \left(\frac{4\pi d}{\lambda}\right)^2 \tag{13.2}$$

where d is the distance between the transmitter and receiver and λ is the wavelength.

We restrict our discussion to those links distorted by the mechanism of additive white Gaussian noise (AWGN) only. Such a noise assumption is a very useful model for a large class of communication systems. A valid approximation for average received noise power N that this model introduces is written as follows [5, 9]:

$$N \cong kT^{\circ}W \tag{13.3}$$

where k is Boltzmann's constant (1.38×10^{-23} joule/K), T° is effective temperature in kelvin, and W is bandwidth in hertz. Dividing Eq. (13.3) by bandwidth, enables us to write the received noise-power spectral density N_0 as follows:

$$N_0 = \frac{N}{W} = kT^{\circ} \tag{13.4}$$

Ford Motor Co.
Exhibit 1030
Page 298

Dividing Eq. (13.1) by N_0 yields the received average signal-power to noise-power spectral density S/N_0 as

$$\frac{S}{N_0} = \frac{EIRP\, G_r/T^\circ}{kL_sL_o} \tag{13.5}$$

where G_r/T° is often referred to as the receiver figure of merit. A link budget analysis is a compilation of the power gains and losses throughout the link; it is generally computed in decibels, and thus takes on the bookkeeping appearance of a business enterprise, highlighting the assets and liabilities of the link. Once the value of S/N_0 is specified or calculated from the link parameters, we then shift our attention to optimizing the choice of signalling types for meeting system bandwidth and error performance requirements.

Given the received S/N_0, we can write the received bit-energy to noise-power spectral density E_b/N_0, for any desired data rate R, as follows:

$$\frac{E_b}{N_0} = \frac{ST_b}{N_0} = \frac{S}{N_0}\left(\frac{1}{R}\right) \tag{13.6}$$

Equation (13.6) follows from the basic definitions that received bit energy is equal to received average signal power times the bit duration and that bit rate is the reciprocal of bit duration. Received E_b/N_0 is a key parameter in defining a digital communication system. Its value indicates the apportionment of the received waveform energy among the bits that the waveform represents. At first glance, one might think that a system specification should entail the symbol-energy to noise-power spectral density E_s/N_0 associated with the arriving waveforms. We will show, however, that for a given S/N_0 the value of E_s/N_0 is a function of the modulation and coding. The reason for defining systems in terms of E_b/N_0 stems from the fact that E_b/N_0 depends only on S/N_0 and R and is unaffected by any system design choices, such as modulation and coding.

13.2 Bandwidth and Power Considerations

Two primary communications resources are the received power and the available transmission bandwidth. In many communication systems, one of these resources may be more precious than the other and, hence, most systems can be classified as either bandwidth limited or power limited. In bandwidth-limited systems, spectrally efficient modulation techniques can be used to save bandwidth at the expense of power; in power-limited systems, power efficient modulation techniques can be used to save power at the expense of bandwidth. In both bandwidth- and power-limited systems, error-correction coding (often called channel coding) can be used to save power or to improve error performance at the expense of bandwidth. Recently, trellis-coded modulation (TCM) schemes have been used to improve the error performance of bandwidth-limited channels without any increase in bandwidth [17], but these methods are beyond the scope of this chapter.

13.2.1 The Bandwidth Efficiency Plane

Figure 13.1 shows the abscissa as the ratio of bit-energy to noise-power spectral density E_b/N_0 (in decibels) and the ordinate as the ratio of throughput, R (in bits per second), that can be transmitted per hertz in a given bandwidth W. The ratio R/W is called bandwidth efficiency, since it reflects how efficiently the bandwidth resource is utilized. The plot stems from the Shannon–Hartley capacity theorem [12, 13, 15], which can be stated as

$$C = W \log_2\left(1 + \frac{S}{N}\right) \tag{13.7}$$

Ford Motor Co.
Exhibit 1030
Page 299

where S/N is the ratio of received average signal power to noise power. When the logarithm is taken to the base 2, the capacity C, is given in bits per second. The capacity of a channel defines the maximum number of bits that can be reliably sent per second over the channel. For the case where the data (information) rate R is equal to C, the curve separates a region of practical communication systems from a region where such communication systems cannot operate reliably [12, 15].

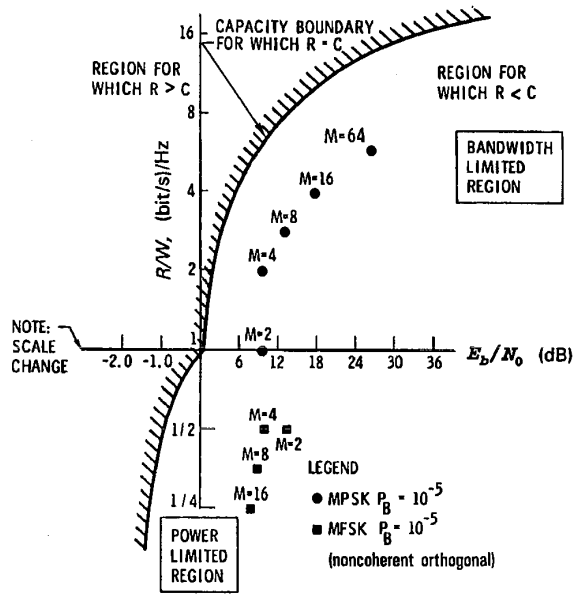


FIGURE 13.1 Bandwidth-efficiency plane.

13.2.2 *M*-ary Signalling

Each symbol in an M-ary alphabet can be related to a unique sequence of m bits, expressed as

$$M = 2^m \qquad \text{or} \qquad m = \log_2 M \tag{13.8}$$

where M is the size of the alphabet. In the case of digital transmission, the term symbol refers to the member of the M-ary alphabet that is transmitted during each symbol duration T_s. To transmit the symbol, it must be mapped onto an electrical voltage or current waveform. Because the waveform represents the symbol, the terms symbol and waveform are sometimes used interchangeably. Since one of M symbols or waveforms is transmitted during each symbol duration T_s, the data rate R in bits per second can be expressed as

$$R = \frac{m}{T_s} = \frac{\log_2 M}{T_s} \tag{13.9}$$

Data-bit-time duration is the reciprocal of data rate. Similarly, symbol-time duration is the reciprocal of symbol rate. Therefore, from Eq. (13.9), we write that the effective time

Ford Motor Co.
Exhibit 1030
Page 300

duration T_b of each bit in terms of the symbol duration T_s or the symbol rate R_s is

$$T_b = \frac{1}{R} = \frac{T_s}{m} = \frac{1}{mR_s} \tag{13.10}$$

Then, using Eqs. (13.8) and (13.10) we can express the symbol rate R_s in terms of the bit rate R as follows:

$$R_s = \frac{R}{\log_2 M} \tag{13.11}$$

From Eqs. (13.9) and (13.10), any digital scheme that transmits $m = \log_2 M$ bits in T_s seconds, using a bandwidth of W hertz, operates at a bandwidth efficiency of

$$\frac{R}{W} = \frac{\log_2 M}{WT_s} = \frac{1}{WT_b} \qquad \text{(b/s)/Hz} \tag{13.12}$$

where T_b is the effective time duration of each data bit.

13.2.3 Bandwidth-Limited Systems

From Eq. (13.12), the smaller the WT_b product, the more bandwidth efficient will be any digital communication system. Thus, signals with small WT_b products are often used with bandwidth-limited systems. For example, the European digital mobile telephone system known as Global System for Mobile Communications (GSM) uses Gaussian minimum shift keying (GMSK) modulation having a WT_b product equal to 0.3 Hz/(b/s), where W is the 3-dB bandwidth of a Gaussian filter [4].

For uncoded bandwidth-limited systems, the objective is to maximize the transmitted information rate within the allowable bandwidth, at the expense of E_b/N_0 (while maintaining a specified value of bit-error probability P_B). The operating points for coherent M-ary phase-shift keying (MPSK) at $P_B = 10^{-5}$ are plotted on the bandwidth-efficiency plane of Fig. 13.1. We assume Nyquist (ideal rectangular) filtering at baseband [10]. Thus, for MPSK, the required double-sideband (DSB) bandwidth at an intermediate frequency (IF) is related to the symbol rate as follows:

$$W = \frac{1}{T_s} = R_s \tag{13.13}$$

where T_s is the symbol duration and R_s is the symbol rate. The use of Nyquist filtering results in the minimum required transmission bandwidth that yields zero intersymbol interference; such ideal filtering gives rise to the name Nyquist minimum bandwidth.

From Eqs. (13.12) and (13.13), the bandwidth efficiency of MPSK modulated signals using Nyquist filtering can be expressed as

$$R/W = \log_2 M \qquad \text{(b/s)/Hz} \tag{13.14}$$

The MPSK points in Fig. 13.1 confirm the relationship shown in Eq. (13.14). Note that MPSK modulation is a bandwidth-efficient scheme. As M increases in value, R/W also increases. MPSK modulation can be used for realizing an improvement in bandwidth efficiency at the cost of increased E_b/N_0. Although beyond the scope of this chapter, many highly bandwidth-efficient modulation schemes are under investigation [1].

Ford Motor Co.
Exhibit 1030
Page 301

13.2.4 Power-Limited Systems

Operating points for noncoherent orthogonal M-ary FSK (MFSK) modulation at $P_B = 10^{-5}$ are also plotted on Fig. 13.1. For MFSK, the IF minimum bandwidth is as follows [15]

$$W = \frac{M}{T_s} = MR_s \tag{13.15}$$

where T_s is the symbol duration and R_s is the symbol rate. With MFSK, the required transmission bandwidth is expanded M-fold over binary FSK since there are M different orthogonal waveforms, each requiring a bandwidth of $1/T_s$. Thus, from Eqs. (13.12) and (13.15), the bandwidth efficiency of noncoherent orthogonal MFSK signals can be expressed as

$$\frac{R}{W} = \frac{\log_2 M}{M} \qquad \text{(b/s)/Hz} \tag{13.16}$$

The MFSK points plotted in Fig. 13.1 confirm the relationship shown in Eq. (13.16). Note that MFSK modulation is a bandwidth-expansive scheme. As M increases, R/W decreases. MFSK modulation can be used for realizing a reduction in required E_b/N_0 at the cost of increased bandwidth.

In Eqs. (13.13) and (13.14) for MPSK, and Eqs. (13.15) and (13.16) for MFSK, and for all the points plotted in Fig. 13.1, ideal filtering has been assumed. Such filters are not realizable! For realistic channels and waveforms, the required transmission bandwidth must be increased in order to account for realizable filters.

In the examples that follow, we will consider radio channels that are disturbed only by additive white Gaussian noise (AWGN) and have no other impairments, and for simplicity, we will limit the modulation choice to constant-envelope types, i.e., either MPSK or noncoherent orthogonal MFSK. For an uncoded system, MPSK is selected if the channel is bandwidth limited, and MFSK is selected if the channel is power limited. When error-correction coding is considered, modulation selection is not as simple, because coding techniques can provide power-bandwidth tradeoffs more effectively than would be possible through the use of any M-ary modulation scheme considered in this chapter [3].

In the most general sense, M-ary signalling can be regarded as a waveform-coding procedure, i.e., when we select an M-ary modulation technique instead of a binary one, we in effect have replaced the binary waveforms with better waveforms—either better for bandwidth performance (MPSK) or better for power performance (MFSK). Even though orthogonal MFSK signalling can be thought of as being a coded system, i.e., a first-order Reed-Muller code [8], we restrict our use of the term coded system to those traditional error-correction codes using redundancies, e.g., block codes or convolutional codes.

13.2.5 Minimum Bandwidth Requirements for MPSK and MFSK Signalling

The basic relationship between the symbol (or waveform) transmission rate R_s and the data rate R was shown in Eq. (13.11). Using this relationship together with Eqs. (13.13–13.16) and $R = 9600$ b/s, a summary of symbol rate, minimum bandwidth, and bandwidth efficiency for MPSK and noncoherent orthogonal MFSK was compiled for $M = 2, 4, 8, 16,$ and 32 (Table 13.1). Values of E_b/N_0 required to achieve a bit-error probability of 10^{-5} for MPSK and MFSK are also given for each value of M. These entries (which were computed using relationships that are presented later in this chapter) corroborate the tradeoffs shown in Fig. 13.1. As M increases, MPSK signalling provides more bandwidth efficiency at the

Ford Motor Co.
Exhibit 1030
Page 302

cost of increased E_b/N_0, whereas MFSK signalling allows for a reduction in E_b/N_0 at the cost of increased bandwidth.

TABLE 13.1 Symbol Rate, Minimum Bandwidth, Bandwidth Efficiency, and Required E_b/N_0 for MPSK and Noncoherent Orthogonal MFSK Signalling at 9600 bit/s

M	m	R (b/s)	R_s (symb/s)	MPSK Minimum Bandwidth (Hz)	MPSK R/W	MPSK E_b/N_0 (dB) $P_B = 10^{-5}$	Noncoherent Orthog MFSK Min Bandwidth (Hz)	MFSK R/W	MFSK E_b/N_0 (dB) $P_B = 10^{-5}$
2	1	9600	9600	9600	1	9.6	19,200	1/2	13.4
4	2	9600	4800	4800	2	9.6	19,200	1/2	10.6
8	3	9600	3200	3200	3	13.0	25,600	3/8	9.1
16	4	9600	2400	2400	4	17.5	38,400	1/4	8.1
32	5	9600	1920	1920	5	22.4	61,440	5/32	7.4

13.3 Example 1: Bandwidth-Limited Uncoded System

Suppose we are given a bandwidth-limited AWGN radio channel with an available bandwidth of $W = 4000$ Hz. Also, suppose that the link constraints (transmitter power, antenna gains, path loss, etc.) result in the ratio of received average signal-power to noise-power spectral density S/N_0 being equal to 53 dB-Hz. Let the required data rate R be equal to 9600 b/s, and let the required bit-error performance P_B be at most 10^{-5}. The goal is to choose a modulation scheme that meets the required performance. In general, an error-correction coding scheme may be needed if none of the allowable modulation schemes can meet the requirements. In this example, however, we shall find that the use of error-correction coding is not necessary.

13.3.1 Solution to Example 1

For any digital communication system, the relationship between received S/N_0 and received bit-energy to noise-power spectral density, E_b/N_0 was given in Eq. (13.6) and is briefly rewritten as

$$\frac{S}{N_0} = \frac{E_b}{N_0} R \tag{13.17}$$

Solving for E_b/N_0 in decibels, we obtain

$$\begin{aligned} \frac{E_b}{N_0}\ (\text{dB}) &= \frac{S}{N_0}\ (\text{dB-Hz}) - R\ (\text{dB-b/s}) \\ &= 53\ \text{dB-Hz} - (10 \times \log_{10} 9600)\ \text{dB-b/s} \\ &= 13.2\ \text{dB (or 20.89)} \end{aligned} \tag{13.18}$$

Since the required data rate of 9600 b/s is much larger than the available bandwidth of 4000 Hz, the channel is bandwidth limited. We therefore select MPSK as our modulation scheme. We have confined the possible modulation choices to be constant-envelope types; without such a restriction, we would be able to select a modulation type with greater bandwidth efficiency. To conserve power, we compute the *smallest possible* value of M such that the MPSK minimum bandwidth does not exceed the available bandwidth of 4000 Hz.

Table 13.1 shows that the smallest value of M meeting this requirement is $M = 8$. Next we determine whether the required bit-error performance of $P_B \leq 10^{-5}$ can be met by using 8-PSK modulation alone or whether it is necessary to use an error-correction coding scheme. Table 13.1 shows that 8-PSK alone will meet the requirements, since the required E_b/N_0 listed for 8-PSK is less than the received E_b/N_0 derived in Eq. (13.18). Let us imagine that we do not have Table 13.1, however, and evaluate whether or not error-correction coding is necessary.

Figure 13.2 shows the basic modulator/demodulator (MODEM) block diagram summarizing the functional details of this design. At the modulator, the transformation from data bits to symbols yields an output symbol rate R_s, that is, a factor $\log_2 M$ smaller than the input data-bit rate R, as is seen in Eq. (13.11). Similarly, at the input to the demodulator, the symbol-energy to noise-power spectral density E_S/N_0 is a factor $\log_2 M$ larger than E_b/N_0, since each symbol is made up of $\log_2 M$ bits. Because E_S/N_0 is larger than E_b/N_0 by the same factor that R_s is smaller than R, we can expand Eq. (13.17), as follows:

$$\frac{S}{N_0} = \frac{E_b}{N_0} R = \frac{E_s}{N_0} R_s \tag{13.19}$$

The demodulator receives a waveform (in this example, one of $M = 8$ possible phase shifts) during each time interval T_s. The probability that the demodulator makes a symbol error $P_E(M)$ is well approximated by the following equation for $M > 2$ [6]:

$$P_E(M) \cong 2Q\left[\sqrt{\frac{2E_s}{N_0}} \sin\left(\frac{\pi}{M}\right)\right] \tag{13.20}$$

where $Q(x)$, sometimes called the complementary error function, represents the probability under the tail of a zero-mean unit-variance Gaussian density function. It is defined as follows [18]:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{u^2}{2}\right) du \tag{13.21}$$

A good approximation for $Q(x)$, valid for $x > 3$, is given by the following equation [2]

$$Q(x) \cong \frac{1}{x\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{13.22}$$

In Fig. 13.2 and all of the figures that follow, rather than show explicit probability relationships, the generalized notation $f(x)$ has been used to indicate some functional dependence on x.

A traditional way of characterizing communication efficiency in digital systems is in terms of the received E_b/N_0 in decibels. This E_b/N_0 description has become standard practice, but recall that there are no bits at the input to the demodulator; there are only waveforms that have been assigned bit meanings. The received E_b/N_0 represents a bit-apportionment of the arriving waveform energy.

To solve for $P_E(M)$ in Eq. (13.20), we first need to compute the ratio of received symbol-energy to noise-power spectral density E_s/N_0. Since from Eq. (13.18)

$$\frac{E_b}{N_0} = 13.2 \text{ dB (or } 20.89)$$

and because each symbol is made up of $\log_2 M$ bits, we compute the following using $M = 8$.

$$\frac{E_s}{N_0} = (\log_2 M) \frac{E_b}{N_0} = 3 \times 20.89 = 62.67 \tag{13.23}$$

Ford Motor Co.
Exhibit 1030
Page 304

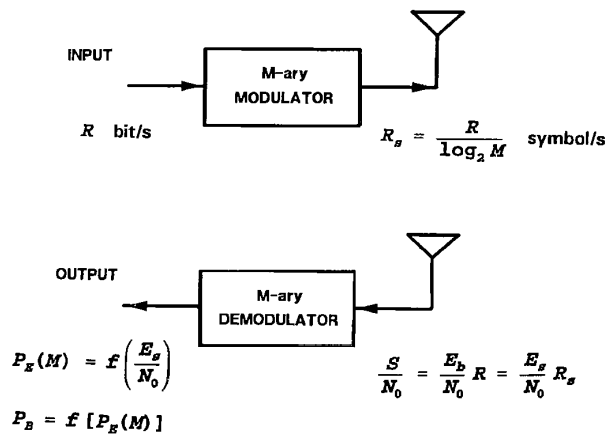


FIGURE 13.2 Basic modulator/demodulator (MODEM) without channel coding.

Using the results of Eq. (13.23) in Eq. (13.20), yields the symbol-error probability $P_E = 2.2 \times 10^{-5}$. To transform this to bit-error probability, we use the relationship between bit-error probability P_B and symbol-error probability P_E, for multiple-phase signalling [8] for $P_E \ll 1$ as follows:

$$P_B \cong \frac{P_E}{\log_2 M} = \frac{P_E}{m} \tag{13.24}$$

which is a good approximation when Gray coding is used for the bit-to-symbol assignment [6]. This last computation yields $P_B = 7.3 \times 10^{-6}$, which meets the required bit-error performance. No error-correction coding is necessary, and 8-PSK modulation represents the design choice to meet the requirements of the bandwidth-limited channel, which we had predicted by examining the required E_b/N_0 values in Table 13.1.

13.4 Example 2: Power-Limited Uncoded System

Now, suppose that we have exactly the same data rate and bit-error probability requirements as in Example 1, but let the available bandwidth W be equal to 45 kHz, and the available S/N_0 be equal to 48 dB-Hz. The goal is to choose a modulation or modulation/coding scheme that yields the required performance. We shall again find that error-correction coding is not required.

13.4.1 Solution to Example 2

The channel is clearly not bandwidth limited since the available bandwidth of 45 kHz is more than adequate for supporting the required data rate of 9600 bit/s. We find the received E_b/N_0 from Eq. (13.18), as follows:

$$\frac{E_b}{N_0}\ (\text{dB}) = 48\ \text{dB-Hz} - (10 \times \log_{10} 9600)\ \text{dB-b/s} = 8.2\ \text{dB (or 6.61)} \tag{13.25}$$

Ford Motor Co.
Exhibit 1030
Page 305

Since there is abundant bandwidth but a relatively small E_b/N_0 for the required bit-error probability, we consider that this channel is power limited and choose MFSK as the modulation scheme. To conserve power, we search for the *largest possible* M such that the MFSK minimum bandwidth is not expanded beyond our available bandwidth of 45 kHz. A search results in the choice of $M = 16$ (Table 13.1). Next, we determine whether the required error performance of $P_B \leq 10^{-5}$ can be met by using 16-FSK alone, i.e., without error-correction coding. Table 13.1 shows that 16-FSK alone meets the requirements, since the required E_b/N_0 listed for 16-FSK is less than the received E_b/N_0 derived in Eq. (13.25). Let us imagine again that we do not have Table 13.1, and evaluate whether or not error-correction coding is necessary.

The block diagram in Fig. 13.2 summarizes the relationships between symbol rate R_s, and bit rate R, and between E_s/N_0 and E_b/N_0, which is identical to each of the respective relationships in Example 1. The 16-FSK demodulator receives a waveform (one of 16 possible frequencies) during each symbol time interval T_s. For noncoherent orthogonal MFSK, the probability that the demodulator makes a symbol error $P_E(M)$ is approximated by the following upper bound [20]:

$$P_E(M) \leq \frac{M-1}{2} \exp\left(-\frac{E_s}{2N_0}\right) \tag{13.26}$$

To solve for $P_E(M)$ in Eq. (13.26), we compute E_S/N_0, as in Example 1. Using the results of Eq. (13.25) in Eq. (13.23), with $M = 16$, we get

$$\frac{E_s}{N_0} = (\log_2 M)\frac{E_b}{N_0} = 4 \times 6.61 = 26.44 \tag{13.27}$$

Next, using the results of Eq. (13.27) in Eq. (13.26), yields the symbol-error probability $P_E = 1.4 \times 10^{-5}$. To transform this to bit-error probability, P_B, we use the relationship between P_B and P_E for orthogonal signalling [20], given by

$$P_B = \frac{2^{m-1}}{(2^m - 1)} P_E \tag{13.28}$$

This last computation yields $P_B = 7.3 \times 10^{-6}$, which meets the required bit-error performance. Thus, we can meet the given specifications for this power-limited channel by using 16-FSK modulation, without any need for error-correction coding, as we had predicted by examining the required E_b/N_0 values in Table 13.1.

13.5 Example 3: Bandwidth-Limited and Power-Limited Coded System

We start with the same channel parameters as in Example 1 ($W = 4000$ Hz, $S/N_0 = 53$ dB-Hz, and $R = 9600$ b/s), with one exception.

Ford Motor Co.
Exhibit 1030
Page 306

In this example, we specify that P_B must be at most 10^{-9}. Table 13.1 shows that the system is both bandwidth limited and power limited, based on the available bandwidth of 4000 Hz and the available E_b/N_0 of 20.2 dB, from Eq. (13.18); 8-PSK is the only possible choice to meet the bandwidth constraint; however, the available E_b/N_0 of 20.2 dB is certainly insufficient to meet the required P_B of 10^{-9}. For this small value of P_B, we need to consider the performance improvement that error-correction coding can provide within the available bandwidth. In general, one can use convolutional codes or block codes.

The Bose–Chaudhuri–Hocquenghem (BCH) codes form a large class of powerful error-correcting cyclic (block) codes [7]. To simplify the explanation, we shall choose a block code from the BCH family. Table 13.2 presents a partial catalog of the available BCH codes in terms of n, k, and t, where k represents the number of information (or data) bits that the code transforms into a longer block of n coded bits (or channel bits), and t represents the largest number of incorrect channel bits that the code can correct within each n-sized block. The rate of a code is defined as the ratio k/n; its inverse represents a measure of the code's redundancy [7].

TABLE 13.2
BCH Codes (Partial Catalog)

n	k	t
7	4	1
15	11	1
	7	2
	5	3
31	26	1
	21	2
	16	3
	11	5
63	57	1
	51	2
	45	3
	39	4
	36	5
	30	6
127	120	1
	113	2
	106	3
	99	4
	92	5
	85	6
	78	7
	71	9
	64	10

13.5.1 Solution to Example 3

Since this example has the same bandwidth-limited parameters given in Example 1, we start with the same 8-PSK modulation used to meet the stated bandwidth constraint. We now employ error-correction coding, however, so that the bit-error probability can be lowered to $P_B \leq 10^{-9}$.

To make the optimum code selection from Table 13.2, we are guided by the following goals.

1. The output bit-error probability of the combined modulation/coding system must meet the system error requirement.
2. The rate of the code must not expand the required transmission bandwidth beyond the available channel bandwidth.
3. The code should be as simple as possible. Generally, the shorter the code, the simpler will be its implementation.

The uncoded 8-PSK minimum bandwidth requirement is 3200 Hz (Table 13.1) and the allowable channel bandwidth is 4000 Hz, and so the uncoded signal bandwidth can be increased by no more than a factor of 1.25 (i.e., an expansion of 25%). The very first step in this (simplified) code selection example is to eliminate the candidates in Table 13.2 that would expand the bandwidth by more than 25%. The remaining entries form a much reduced set of bandwidth-compatible codes (Table 13.3).

In Table 13.3, a column designated Coding Gain G (for MPSK at $P_B = 10^{-9}$) has been added. Coding gain in decibels is defined as follows:

$$G = \left(\frac{E_b}{N_0}\right)_{\text{uncoded}} - \left(\frac{E_b}{N_0}s\right)_{\text{coded}} \tag{13.29}$$

G can be described as the reduction in the required E_b/N_0 (in decibels) that is needed due to the error-performance properties of the channel coding. G is a function of the

Ford Motor Co.
Exhibit 1030
Page 307

TABLE 13.3 Bandwidth-Compatible BCH Codes

n	k	t	Coding Gain, G (dB) MPSK, $P_B = 10^{-9}$
31	26	1	2.0
63	57	1	2.2
	51	2	3.1
127	120	1	2.2
	113	2	3.3
	106	3	3.9

modulation type and bit-error probability, and it has been computed for MPSK at $P_B = 10^{-9}$ (Table 13.3). For MPSK modulation, G is relatively independent of the value of M. Thus, for a particular bit-error probability, a given code will provide about the same coding gain when used with any of the MPSK modulation schemes. Coding gains were calculated using a procedure outlined in the subsequent Calculating Coding Gain section.

A block diagram summarizes this system, which contains both modulation and coding (Fig. 13.3). The introduction of encoder/decoder blocks brings about additional transformations. The relationships that exist when transforming from R b/s to R_c channel-b/s to R_s symbol/s are shown at the encoder/modulator. Regarding the channel-bit rate R_c, some authors prefer to use the units of channel-symbol/s (or code-symbol/s). The benefit is that error-correction coding is often described more efficiently with nonbinary digits. We reserve the term symbol for that group of bits mapped onto an electrical waveform for transmission, and we designate the units of R_c to be channel-b/s (or coded-b/s).

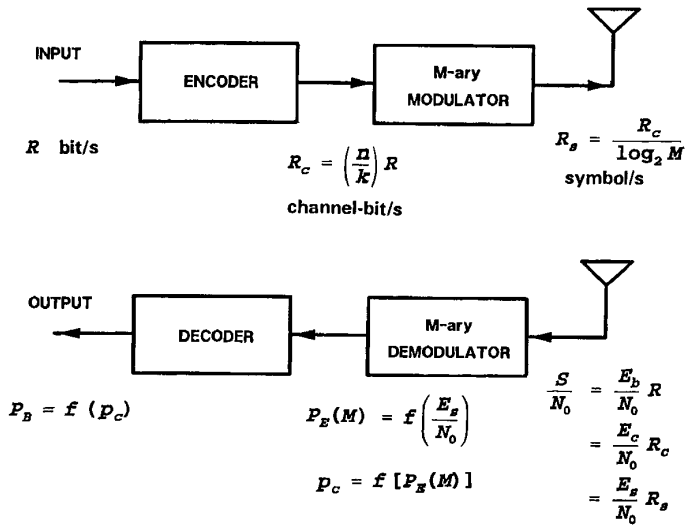


FIGURE 13.3 MODEM with channel coding.

We assume that our communication system cannot tolerate any message delay, so that the channel-bit rate R_c must exceed the data-bit rate R by the factor n/k. Further, each symbol is made up of $\log_2 M$ channel bits, and so the symbol rate R_s is less than R_c by the factor $\log_2 M$. For a system containing both modulation and coding, we summarize the

Ford Motor Co.
Exhibit 1030
Page 308

rate transformations as follows:

$$R_c = \left(\frac{n}{k}\right) R \tag{13.30}$$

$$R_s = \frac{R_c}{\log_2 M} \tag{13.31}$$

At the demodulator/decoder in Fig.13.3, the transformations among data-bit energy, channel- bit energy, and symbol energy are related (in a reciprocal fashion) by the same factors as shown among the rate transformations in Eqs. (13.30) and (13.31). Since the encoding transformation has replaced k data bits with n channel bits, then the ratio of channel-bit energy to noise-power spectral density E_c/N_0 is computed by decrementing the value of E_b/N_0 by the factor k/n. Also, since each transmission symbol is made up of $\log_2 M$ channel bits, then E_S/N_0, which is needed in Eq. (13.20) to solve for P_E, is computed by incrementing E_c/N_0 by the factor $\log_2 M$. For a system containing both modulation and coding, we summarize the energy to noise-power spectral density transformations as follows:

$$\frac{E_c}{N_0} = \left(\frac{k}{n}\right) \frac{E_b}{N_0} \tag{13.32}$$

$$\frac{E_s}{N_0} = (\log_2 M) \frac{E_c}{N_0} \tag{13.33}$$

Using Eqs. (13.30) and (13.31), we can now expand the expression for S/N_0 in Eq. (13.19), as follows (Appendix).

$$\frac{S}{N_0} = \frac{E_b}{N_0} R = \frac{E_c}{N_0} R_c = \frac{E_s}{N_0} R_s \tag{13.34}$$

As before, a standard way of describing the link is in terms of the received E_b/N_0 in decibels. However, there are no data bits at the input to the demodulator, and there are no channel bits; there are only waveforms that have bit meanings and, thus, the waveforms can be described in terms of bit-energy apportionments.

Since S/N_0 and R were given as 53 dB-Hz and 9600 b/s, respectively, we find as before, from Eq. (13.18), that the received $E_b/N_0 = 13.2$ dB. The received E_b/N_0 is fixed and independent of n, k, and t (Appendix). As we search, in Table 13.3 for the ideal code to meet the specifications, we can iteratively repeat the computations suggested in Fig. 13.3. It might be useful to program on a personal computer (or calculator) the following four steps as a function of n, k, and t. Step 1 starts by combining Eqs. (13.32) and (13.33), as follows.

Step 1:

$$\frac{E_s}{N_0} = (\log_2 M) \frac{E_c}{N_0} = (\log_2 M) \left(\frac{k}{n}\right) \frac{E_b}{N_0} \tag{13.35}$$

Step 2:

$$P_E(M) \cong 2Q \left[\sqrt{\frac{2E_s}{N_0}} \sin\left(\frac{\pi}{M}\right) \right] \tag{13.36}$$

which is the approximation for symbol-error probability P_E rewritten from Eq. (13.20). At each symbol-time interval, the demodulator makes a symbol decision, but it delivers a channel-bit sequence representing that symbol to the decoder. When the channel-bit output of the demodulator is quantized to two levels, 1 and 0, the demodulator is said to make hard decisions. When the output is quantized to more than two levels, the demodulator

Ford Motor Co.
Exhibit 1030
Page 309

is said to make soft decisions [15]. Throughout this paper, we shall assume hard-decision demodulation.

Now that we have a decoder block in the system, we designate the channel-bit-error probability out of the demodulator and into the decoder as p_c, and we reserve the notation P_B for the bit-error probability out of the decoder. We rewrite Eq. (13.24) in terms of p_c for $P_E \ll 1$ as follows.

Step 3:

$$p_c \cong \frac{P_E}{\log_2 M} = \frac{P_E}{m} \tag{13.37}$$

relating the channel-bit-error probability to the symbol-error probability out of the demodulator, assuming Gray coding, as referenced in Eq. (13.24).

For traditional channel-coding schemes and a given value of received S/N_0, the value of E_s/N_0 with coding will always be less than the value of E_s/N_0 without coding. Since the demodulator with coding receives less E_s/N_0, it makes more errors! When coding is used, however, the system error-performance does not only depend on the performance of the demodulator, it also depends on the performance of the decoder. For error-performance improvement due to coding, the decoder must provide enough error correction to more than compensate for the poor performance of the demodulator.

The final output decoded bit-error probability P_B depends on the particular code, the decoder, and the channel-bit-error probability p_c. It can be expressed by the following approximation [11].

Step 4:

$$P_B \cong \frac{1}{n} \sum_{j=t+1}^{n} j \binom{n}{j} p_c^j (1-p_c)^{n-j} \tag{13.38}$$

where t is the largest number of channel bits that the code can correct within each block of n bits. Using Eqs. (13.35–13.38) in the four steps, we can compute the decoded bit-error probability P_B as a function of n, k, and t for each of the codes listed in Table 13.3. The entry that meets the stated error requirement with the largest possible code rate and the smallest value of n is the double-error correcting (63, 51) code. The computations are as follows.

Step 1:

$$\frac{E_s}{N_0} = 3\left(\frac{51}{63}\right) 20.89 = 50.73$$

where $M = 8$, and the received $E_b/N_0 = 13.2$ dB (or 20.89).

Step 2:

$$P_E \cong 2Q\left[\sqrt{101.5} \times \sin\left(\frac{\pi}{8}\right)\right] = 2Q(3.86) = 1.2 \times 10^{-4}$$

Step 3:

$$p_c \cong \frac{1.2 \times 10^{-4}}{3} = 4 \times 10^{-5}$$

Step 4:

$$\begin{aligned} P_B &\cong \frac{3}{63}\binom{63}{3}\left(4 \times 10^{-5}\right)^3 \left(1 - 4 \times 10^{-5}\right)^{60} \\ &\quad + \frac{4}{63}\binom{63}{4}\left(4 \times 10^{-5}\right)^4 \left(1 - 4 \times 10^{-5}\right)^{59} + \cdots \\ &= 1.2 \times 10^{-10} \end{aligned}$$

Ford Motor Co.
Exhibit 1030
Page 310

where the bit-error-correcting capability of the code is $t = 2$. For the computation of P_B in step 4, we need only consider the first two terms in the summation of Eq. (13.38) since the other terms have a vanishingly small effect on the result. Now that we have selected the (63, 51) code, we can compute the values of channel-bit rate R_c and symbol rate R_s using Eqs. (13.30) and (13.31), with $M = 8$,

$$R_c = \left(\frac{n}{k}\right) R = \left(\frac{63}{51}\right) 9600 \approx 11{,}859 \text{ channel-b/s}$$

$$R_s = \frac{R_c}{\log_2 M} = \frac{11859}{3} = 3953 \text{ symbol/s}$$

13.5.2 Calculating Coding Gain

Perhaps a more direct way of finding the simplest code that meets the specified error performance is to first compute how much coding gain G is required in order to yield $P_B = 10^{-9}$ when using 8-PSK modulation alone; then, from Table 13.3, we can simply choose the code that provides this performance improvement. First, we find the uncoded E_s/N_0 that yields an error probability of $P_B = 10^{-9}$, by writing from Eqs. (13.24) and (13.36), the following:

$$P_B \cong \frac{P_E}{\log_2 M} \cong \frac{2Q\left[\sqrt{\frac{2E_s}{N_0}}\sin\left(\frac{\pi}{M}\right)\right]}{\log_2 M} = 10^{-9} \tag{13.39}$$

At this low value of bit-error probability, it is valid to use Eq. (13.22) to approximate $Q(x)$ in Eq. (13.39) By trial and error (on a programmable calculator), we find that the uncoded $E_s/N_0 = 120.67 = 20.8$ dB, and since each symbol is made up of $\log_2 8 = 3$ bits, the required $(E_b/N_0)_{\text{uncoded}} = 120.67/3 = 40.22 = 16$ dB. From the given parameters and Eq. (13.18), we know that the received $(E_b/N_0)_{\text{coded}} = 13.2$ dB. Using Eq. (13.29), the required coding gain to meet the bit-error performance of $P_B = 10^{-9}$ in decibels is

$$G = \left(\frac{E_b}{N_0}\right)_{\text{uncoded}} - \left(\frac{E_b}{N_0}\right)_{\text{coded}} = 16 - 13.2 = 2.8$$

To be precise, each of the E_b/N_0 values in the preceding computation must correspond to exactly the same value of bit-error probability (which they do not). They correspond to $P_B = 10^{-9}$ and $P_B = 1.2 \times 10^{-10}$, respectively. At these low probability values, however, even with such a discrepancy, this computation still provides a good approximation of the required coding gain. In searching Table 13.3 for the simplest code that will yield a coding gain of at least 2.8 dB, we see that the choice is the (63, 51) code, which corresponds to the same code choice that we made earlier.

13.6 Example 4: Direct-Sequence (DS) Spread-Spectrum Coded System

Spread-spectrum systems are not usually classified as being bandwidth- or power-limited. They are generally perceived to be power-limited systems, however, because the bandwidth occupancy of the information is much larger than the bandwidth that is intrinsically needed for the information transmission. In a direct-sequence spread-spectrum (DS/SS) system, spreading the signal bandwidth by some factor permits lowering the signal-power spectral density by the same factor (the total average signal power is the same as before spreading).

The bandwidth spreading is typically accomplished by multiplying a relatively narrowband data signal by a wideband spreading signal. The spreading signal or spreading code is often referred to as a pseudorandom code or PN code.

13.6.1 Processing Gain

A typical DS/SS radio system is often described as a two-step BPSK modulation process. In the first step, the carrier wave is modulated by a bipolar data waveform having a value +1 or −1 during each data-bit duration; in the second step, the output of the first step is multiplied (modulated) by a bipolar PN-code waveform having a value +1 or −1 during each PN-code-bit duration. In reality, DS/SS systems are usually implemented by first multiplying the data waveform by the PN-code waveform and then making a single pass through a BPSK modulator. For this example, however, it is useful to characterize the modulation process in two separate steps—the outer modulator/demodulator for the data, and the inner modulator/demodulator for the PN code (Fig. 13.4).

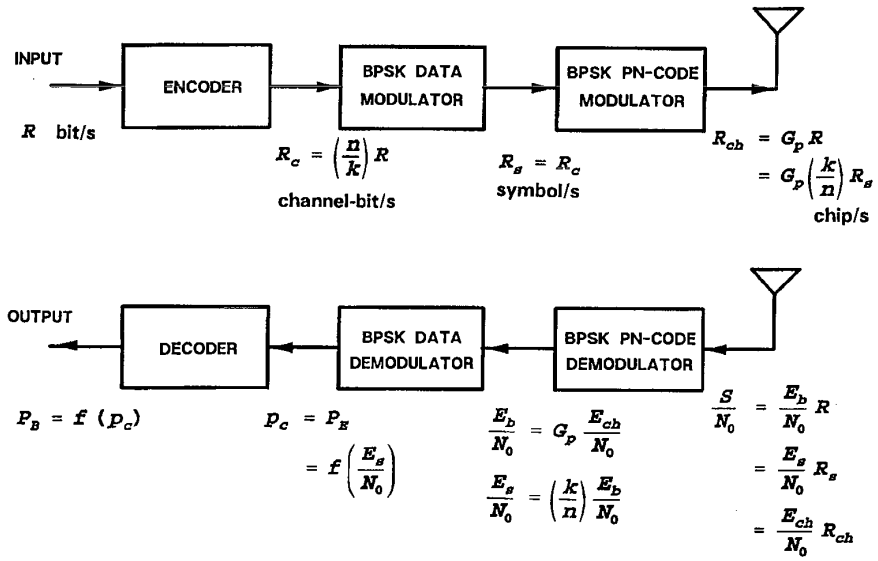


FIGURE 13.4 Direct-sequence spread-spectrum MODEM with channel coding.

A spread-spectrum system is characterized by a processing gain G_p, that is defined in terms of the spread-spectrum bandwidth W_{ss} and the data rate R as follows [20]:

$$G_p = \frac{W_{\text{ss}}}{R} \tag{13.40}$$

For a DS/SS system, the PN-code bit has been given the name chip, and the spread-spectrum signal bandwidth can be shown to be about equal to the chip rate R_{ch} as follows:

$$G_p = \frac{R_{\text{ch}}}{R} \tag{13.41}$$

Some authors define processing gain to be the ratio of the spread-spectrum bandwidth to

Ford Motor Co.
Exhibit 1030
Page 312

the symbol rate. This definition separates the system performance that is due to bandwidth spreading from the performance that is due to error-correction coding. Since we ultimately want to relate all of the coding mechanisms relative to the information source, we shall conform to the most usually accepted definition for processing gain, as expressed in Eqs. (13.40) and (13.41).

A spread-spectrum system can be used for interference rejection and for multiple access (allowing multiple users to access a communications resource simultaneously). The benefits of DS/SS signals are best achieved when the processing gain is very large; in other words, the chip rate of the spreading (or PN) code is much larger than the data rate. In such systems, the large value of G_p allows the signalling chips to be transmitted at a power level well below that of the thermal noise. We will use a value of $G_p = 1000$. At the receiver, the despreading operation correlates the incoming signal with a synchronized copy of the PN code and, thus, accumulates the energy from multiple (G_p) chips to yield the energy per data bit. The value of G_p has a major influence on the performance of the spread-spectrum system application. We shall see, however, that the value of G_p has no effect on the received E_b/N_0. In other words, spread spectrum techniques offer no error-performance advantage over thermal noise. For DS/SS systems, there is no disadvantage either! Sometimes such spread-spectrum radio systems are employed only to enable the transmission of very small power-spectral densities and thus avoid the need for FCC licensing [16].

13.6.2 Channel Parameters for Example 13.4

Consider a DS/SS radio system that uses the same (63, 51) code as in the previous example. Instead of using MPSK for the data modulation, we shall use BPSK. Also, we shall use BPSK for modulating the PN-code chips. Let the received $S/N_0 = 48$ dB-Hz, the data rate $R = 9600$ b/s, and the required $P_B \leq 10^{-6}$. For simplicity, assume that there are no bandwidth constraints. Our task is simply to determine whether or not the required error performance can be achieved using the given system architecture and design parameters. In evaluating the system, we will use the same type of transformations used in the previous examples.

13.6.3 Solution to Example 13.4

A typical DS/SS system can be implemented more simply than the one shown in Fig. 13.4. The data and the PN code would be combined at baseband, followed by a single pass through a BPSK modulator. We will, however, assume the existence of the individual blocks in Fig. 13.4 because they enhance our understanding of the transformation process. The relationships in transforming from data bits, to channel bits, to symbols, and to chips Fig. 13.4 have the same pattern of subtle but straightforward transformations in rates and energies as previous relationships (Figs. 13.2 and 13.3). The values of R_c, R_s, and R_{ch} can now be calculated immediately since the (63, 51) BCH code has already been selected. From Eq. (13.30) we write

$$R_c = \left(\frac{n}{k}\right) R = \left(\frac{63}{51}\right) 9600 \approx 11{,}859 \text{ channel-b/s}$$

Since the data modulation considered here is BPSK, then from Eq. (13.31) we write

$$R_s = R_c \approx 11{,}859 \text{ symbol/s}$$

Ford Motor Co.
Exhibit 1030
Page 313

and from Eq. (13.41), with an assumed value of $G_p = 1000$

$$R_{\text{ch}} = G_p R = 1000 \times 9600 = 9.6 \times 10^6 \text{ chip/s}$$

Since we have been given the same S/N_0 and the same data rate as in Example 2, we find the value of received E_b/N_0 from Eq. (13.25) to be 8.2 dB (or 6.61). At the demodulator, we can now expand the expression for S/N_0 in Eq. (13.34) and the Appendix as follows:

$$\frac{S}{N_0} = \frac{E_b}{N_0}R = \frac{E_c}{N_0}R_c = \frac{E_s}{N_0}R_s = \frac{E_{\text{ch}}}{N_0}R_{\text{ch}} \tag{13.42}$$

Corresponding to each transformed entity (data bit, channel bit, symbol, or chip) there is a change in rate and, similarly, a reciprocal change in energy-to-noise spectral density for that received entity. Equation (13.42) is valid for any such transformation when the rate and energy are modified in a reciprocal way. There is a kind of *conservation of power* (or energy) phenomenon that exists in the transformations. The total received average power (or total received energy per symbol duration) is fixed regardless of how it is computed, on the basis of data bits, channel bits, symbols, or chips.

The ratio E_{ch}/N_0 is much lower in value than E_b/N_0. This can be seen from Eqs. (13.42) and (13.41), as follows:

$$\frac{E_{\text{ch}}}{N_0} = \frac{S}{N_0}\left(\frac{1}{R_{\text{ch}}}\right) = \frac{S}{N_0}\left(\frac{1}{G_p R}\right) = \left(\frac{1}{G_p}\right)\frac{E_b}{N_0} \tag{13.43}$$

But, even so, the despreading function (when properly synchronized) accumulates the energy contained in a quantity G_p of the chips, yielding the same value $E_b/N_0 = 8.2$ dB, as was computed earlier from Eq. (13.25). Thus, the DS spreading transformation has no effect on the error performance of an AWGN channel [15], and the value of G_p has no bearing on the value of P_B in this example.

From Eq. (13.43), we can compute, in decibels,

$$\begin{aligned} \frac{E_{\text{ch}}}{N_0} &= E_b/N_0 - G_p \\ &= 8.2 - (10 \times \log_{10} 1000) \\ &= -21.8 \end{aligned} \tag{13.44}$$

The chosen value of processing gain ($G_p = 1000$) enables the DS/SS system to operate at a value of chip energy well below the thermal noise, with the same error performance as without spreading.

Since BPSK is the data modulation selected in this example, each message symbol therefore corresponds to a single channel bit, and we can write

$$\frac{E_s}{N_0} = \frac{E_c}{N_0} = \left(\frac{k}{n}\right)\frac{E_b}{N_0} = \left(\frac{51}{63}\right) \times 6.61 = 5.35 \tag{13.45}$$

where the received $E_b/N_0 = 8.2$ dB (or 6.61). Out of the BPSK data demodulator, the symbol-error probability P_E (and the channel-bit error probability p_c) is computed as follows [15]:

$$p_c = P_E = Q\left(\sqrt{\frac{2E_c}{N_0}}\right) \tag{13.46}$$

Using the results of Eq. (13.45) in Eq. (13.46) yields

$$p_c = Q(3.27) = 5.8 \times 10^{-4}$$

Ford Motor Co.
Exhibit 1030
Page 314

Finally, using this value of p_c in Eq. (13.38) for the (63,51) double-error correcting code yields the output bit-error probability of $P_B = 3.6 \times 10^{-7}$. We can, therefore, verify that for the given architecture and design parameters of this example the system does, in fact, achieve the required error performance.

13.7 Conclusion

The goal of this section has been to review fundamental relationships used in evaluating the performance of digital communication systems. First, we described the concept of a link and a channel and examined a radio system from its transmitting segment up through the output of the receiving antenna. We then examined the concept of bandwidth-limited and power-limited systems and how such conditions influence the system design when the choices are confined to MPSK and MFSK modulation. Most important, we focused on the definitions and computations involved in transforming from data bits to channel bits to symbols to chips. In general, most digital communication systems share these concepts; thus, understanding them should enable one to evaluate other such systems in a similar way.

Appendix: Received E_b/N_0 Is Independent of the Code Parameters

Starting with the basic concept that the received average signal power S is equal to the received symbol or waveform energy, E_s, divided by the symbol-time duration, T_s (or multiplied by the symbol rate, R_s), we write

$$\frac{S}{N_0} = \frac{E_s/T_s}{N_0} = \frac{E_s}{N_0} R_s \tag{A13.1}$$

where N_0 is noise-power spectral density.

Using Eqs. (13.27) and (13.25), rewritten as

$$\frac{E_s}{N_0} = (\log_2 M)\frac{E_c}{N_0} \qquad \text{and} \qquad R_s = \frac{R_c}{\log_2 M}$$

let us make substitutions into Eq. (A13.1), which yields

$$\frac{S}{N_0} = \frac{E_c}{N_0} R_c \tag{A13.2}$$

Next, using Eqs. (13.26) and (13.24), rewritten as

$$\frac{E_c}{N_0} = \left(\frac{k}{n}\right)\frac{E_b}{N_0} \qquad \text{and} \qquad R_c = \left(\frac{n}{k}\right) R$$

let us now make substitutions into Eq. (A13.2), which yields the relationship expressed in Eq. (13.11)

$$\frac{S}{N_0} = \frac{E_b}{N_0} R \tag{A13.3}$$

Hence, the received E_b/N_0 is only a function of the received S/N_0 and the data rate R. It is independent of the code parameters, n, k, and t. These results are summarized in Fig. 13.3.

Ford Motor Co.
Exhibit 1030
Page 315

References

[1] Anderson, J.B. and Sundberg, C.-E.W., Advances in constant envelope coded modulation, *IEEE Commun., Mag.*, 29(12), 36–45, 1991.

[2] Borjesson, P.O. and Sundberg, C.E., Simple approximations of the error function $Q(x)$ for communications applications, *IEEE Trans. Comm.*, COM-27, 639–642, Mar. 1979.

[3] Clark Jr., G.C. and Cain, J.B., *Error-Correction Coding for Digital Communications,* Plenum Press, New York, 1981.

[4] Hodges, M.R.L., The GSM radio interface, *British Telecom Technol. J.*, 8(1), 31–43, 1990.

[5] Johnson, J.B., Thermal agitation of electricity in conductors, *Phys. Rev.*, 32, 97–109, Jul. 1928.

[6] Korn, I., *Digital Communications,* Van Nostrand Reinhold Co., New York, 1985.

[7] Lin, S. and Costello Jr., D.J., *Error Control Coding: Fundamentals and Applications,* Prentice-Hall, Englewood Cliffs, NJ, 1983.

[8] Lindsey, W.C. and Simon, M.K., *Telecommunication Systems Engineering,* Prentice-Hall, Englewood Cliffs, NJ, 1973.

[9] Nyquist, H., Thermal agitation of electric charge in conductors, *Phys. Rev.*, 32, 110–113, Jul. 1928.

[10] Nyquist, H., Certain topics on telegraph transmission theory, *Trans. AIEE,* 47, 617–644, Apr. 1928.

[11] Odenwalder, J.P., *Error Control Coding Handbook.* Linkabit Corp., San Diego, CA, Jul. 15, 1976.

[12] Shannon, C.E., A mathematical theory of communication, *BSTJ.* 27, 379–423, 623–657, 1948.

[13] Shannon, C.E., Communication in the presence of noise, *Proc. IRE.* 37(1), 10–21, 1949.

[14] Sklar, B., What the system link budget tells the system engineer or how I learned to count in decibels, *Proc. of the Intl. Telemetering Conf.*, San Diego, CA, Nov. 1979.

[15] Sklar, B., *Digital Communications: Fundamentals and Applications,* Prentice-Hall, Englewood Cliffs, NJ, 1988.

[16] Title 47, *Code of Federal Regulations,* Part 15 Radio Frequency Devices.

[17] Ungerboeck, G., Trellis-coded modulation with redundant signal sets, Pt. I and II, *IEEE Comm. Mag.*, 25, 5–21. Feb. 1987.

[18] Van Trees, H.L., *Detection, Estimation, and Modulation Theory,* Pt. I, John Wiley & Sons, New York, 1968.

[19] Viterbi, A.J., *Principles of Coherent Communication,* McGraw-Hill, New York, 1966.

[20] Viterbi, A.J., Spread spectrum communications—myths and realities, *IEEE Comm. Mag.*, 11–18, May, 1979.

Further Information

A useful compilation of selected papers can be found in: *Cellular Radio & Personal Communications–A Book of Selected Readings,* edited by Theodore S. Rappaport, Institute of Electrical and Electronics Engineers, Inc., Piscataway, New Jersey, 1995. Fundamental design issues, such as propagation, modulation, channel coding, speech coding, multiple-accessing and networking, are well represented in this volume.

Another useful sourcebook that covers the fundamentals of mobile communications in great detail is: *Mobile Radio Communications,* edited by Raymond Steele, Pentech Press, London 1992. This volume is also available through the Institute of Electrical and Electronics Engineers, Inc., Piscataway, New Jersey.

Ford Motor Co.
Exhibit 1030
Page 316

For spread spectrum systems, an excellent reference is: *Spread Spectrum Communications Handbook,* by Marvin K. Simon, Jim K. Omura, Robert A. Scholtz, and Barry K. Levitt, McGraw-Hill Inc., New York, 1994.

14

Telecommunications Standardization

14.1 Introduction 14-1
14.2 Global Standardization 14-2
ITU-T • ITU-R • BDT • ISO/IEC JTC 1
14.3 Regional Standardization 14-6
CITEL
14.4 National Standardization 14-9
ANSI T1 • TIA • TTC
14.5 Intellectual Property 14-12
14.6 Standards Coordination 14-13
14.7 Scientific 14-14
14.8 Standards Development Cycle 14-14
Defining Terms 14-15
Further Information 14-16

Spiros Dimolitsas
Lawrence Livermore National Laboratory

Michael Onufry
COMSAT Laboratories

14.1 Introduction

National economies are increasingly becoming information based, where networking and information transport provide a foundation for productivity and economic growth. Concurrently, many countries are rapidly adopting deregulation policies that are resulting in a telecommunications industry that is increasingly multicarrier and multivendor based, and where interconnectivity and compatibility between different networks is emerging as key to the success of this technological and regulatory transition. The communications industry has, consequently, become more interested in standardization; **standards** give manufacturers, service providers, and users freedom of choice at reasonable cost.

In this chapter, a review is provided of the primary telecommunications standards setting bodies. As will be seen, these bodies are often driven by slightly different underlying philosophies, but the output of their activities, i.e., the standards, possess essentially the same characteristics. An all-encompassing review of standardization bodies is not attempted here; this would clearly take many volumes to describe. Furthermore, as country after country increasingly deregulates its telecommunication industry, new standards setting bodies emerge to fill in the void of the de-facto (but no longer existing) standards setting bodies: the national telecommunications administration.

The principal communications standards bodies that will be covered are the following: the International Telecommunications Union (**ITU**); the United States **ANSI** Committee

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 318

T1 on Telecommunications; the Telecommunications Industry Association (**TIA**); the European Telecommunications Standards Institute (**ETSI**); the Inter-American Telecommunications Commission (**CITEL**); the Japanese Telecommunications Technology Committee (**TTC**); and the Institute of Electrical and Electronics Engineers (**IEEE**). Not addressed explicitly are other standards setting bodies that are either national or regional in character, even though it is recognized that sometimes there is overlap in scope with the bodies explicitly covered here.

Most notably, standards setting bodies that are not covered, but that are worth noting, include: the United States ANSI Committee X3; the International Standards Organization (**ISO**), the International Electrotechnical Commission (**IEC**) [except ISO/IEC joint technical committee (JTC) 1], the Telecommunications Standards Advisory Council of Canada (TSACC), the Australian Telecommunications Standardization Committee (ATSC), the Telecommunication Technology Association (TTA) in Korea, and several forums (whose scope is, in principle, somewhat different) such as the asynchronous transfer mode (ATM) forum, the frame relay forum, the integrated digital services network (ISDN) users' forum, and telocator. As will be described later, many of these bodies operate in a coherent fashion through a mechanism developed by the Interregional Telecommunications Standards Conference (ITSC) and its successor, the Global Standards Collaboration (**GSC**).

14.2 Global Standardization

When it comes to setting global communications standards, the ITU comes to the forefront. The ITU is an intergovernmental organization, whereby each sovereign state that is a member of the United Nations may become a member of the ITU. Member governments (in most cases represented by their telecommunications administrations) are constitutional members with a right to vote. Other organizations, such as network and service providers, manufacturers, and scientific and industrial organizations also participate in ITU activities but with a lower legal status.

ITU traces its history back to 1865 in the era of telegraphy. The supreme organ of the ITU is the plenipotentiary conference, which is held not less than every five years and plays a major role in the management of ITU. In 1993 the ITU as a U.N.-specialized agency was reorganized into three sectors (see Fig. 14.1): The *telecommunications standardization* sector (**ITU-T**), the *radiocommunications* sector (**ITU-R**), and the *development* sector (BDT). These sectors' activities are, respectively, standardization of telecommunications, including radio communications; regulation of telecommunications (mainly for radio communications); and development of telecommunications.

It should be noted that, in general, the ITU-T is the successor of the international telephone and telegraph consultative committee (**CCITT**) of the ITU with additional responsibilities for standardization of network-related radio communications. Similarly, the ITU-R is the successor of the international radio consultative committee (**CCIR**) and the international frequency registration bureau (IFRB) of the ITU (after transferring some of its standardization activities to the ITU-T). The BDT is a new sector, which became operational in 1989.

14.2.1 ITU-T

Within the ITU structure, standardization work is undertaken by a number of study groups (SG) dealing with specific areas of communications. There are currently 14 study groups, as shown in Table 14.1.

Ford Motor Co.
Exhibit 1030
Page 319

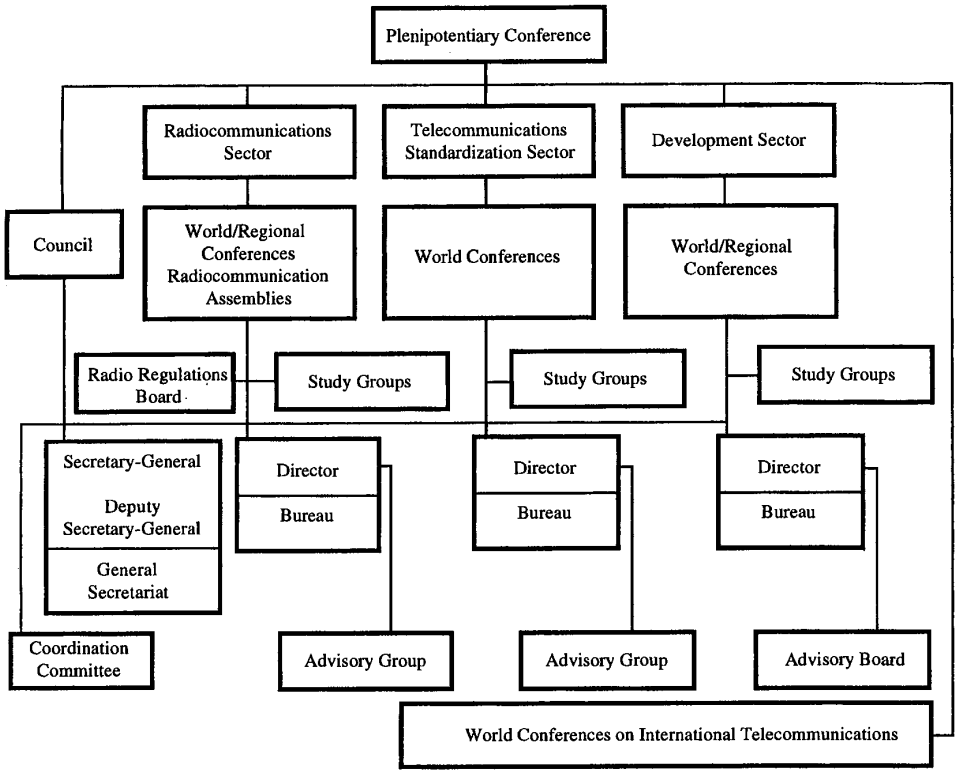


FIGURE 14.1 The ITU structure.

Study groups develop standards for their respective work areas, which then have to be agreed upon by consensus—a process that for the time being is reserved to administrations only. The standards so developed are called **recommendations** to indicate their legal nonbinding nature. Technically, however, there is no distinction between recommendations developed by the ITU and standards developed by other standards setting bodies.

The study groups' work is undertaken by delegation members, sent or sponsored by their national administrations, and delegates from recognized private operating organizations (RPOA), scientific and industrial organizations (SIO) or international organizations. Because an ITU-T study group can typically have from 100 to more than 500 participating members and deal with 20–50 project standards, the work of each study group is often divided among working parties (WP). Such working parties are usually split further into experts' groups led by a chair or "rapporteur" with responsibility for leading the work defined in an approved active question or subelement of a question.

To coordinate standardization work that spans several study groups, two joint coordination groups (**JCG**) have also been established (not shown in Fig. 14.1): International Mobile Communications (IMT-2000) and Satellite Matters.

Such groups do not have executive powers but are merely there to coordinate work of pervasive interest within the ITU-T sector.

Also part of the ITU-T structure is the telecommunications standardization bureau (TSB) or, as it was formerly called, the CCITT secretariat. The TSB is responsible for the organization of numerous meetings held by the sector each year as well as all other support

Ford Motor Co.
Exhibit 1030
Page 320

TABLE 14.1 ITU-T Study Group Structure

SG 2	Network and service operation Lead SG on Service definition, Numbering, Routing and Global Mobility
SG 3	Tariff and accounting principles
SG 4	TMN and network maintenance Lead SG on Telecommunication management network (TMN) studies
SG 5	Protection against electromagnetic environmental effects
SG 6	Outside plant
SG 7	Data networks and open systems communications Lead SG on Open Distributed Processing (ODP), Frame Relay and for Communications System Security
SG 8	Characteristics of telematic services Lead SG on Facsimile
SG 9	Television and sound transmission
SG 10	Languages and general software aspects for telecommunications systems
SG 11	Signalling requirements and protocols Lead SG on Intelligent Network and IMT-2000
SG 12	End-to-end transmission performance of networks and terminals
SG 13	General network aspects Lead SG on General network aspects Global Information Infrastructures and Broadband ISDN
SG 15	Transport networks, systems and equipment Lead SG on Access Network Transport
SG 16	Transmission systems and equipment Lead SG on Multimedia services and systems

services required to ensure the smooth and efficient operation of the sector (including, but not limited to, document production and distribution). The TSB is headed by a director, who holds the executive power and, in collaboration with the study groups, bears full responsibility for the ITU-T activities. In this structure, unlike other U.N. organizations, the secretary general is the legal representative of the ITU, with the executive powers being vested in the director.

Finally, the ITU-T is supported by an advisory group, i.e., the telecommunications standardization advisory group (TSAG), which together with interested ITU members, the ITU-T Director, and ITU-T SG chairman, guides standardization activities.

14.2.2 ITU-R

The radiocommunications sector emphasizes the regulatory and pure radio-interface aspects. The functional structure of the ITU-R currently includes eight study groups, (shown in Table 14.2) a radiocommunications bureau, and an advisory board. The role of the latter two elements is very similar to the ITU-T and, thus, need not be repeated here.

As within the ITU-T, there are areas of pervasive interest, and so areas of common interest can be found between the ITU-T and ITU-R where activities need to be coordinated. To achieve this objective, two intersector coordination groups (**ICG**) have been established (not shown in Fig. 14.1) dealing with international mobile telecommunications (IMT-2000), and satellite matters.

Three major special activities have been organized within ITU-R:

- IMT-2000 (formerly known as Future Public Land Mobile Telecommunications Systems FPLMTS). The objective of the International Mobile Telecommunications (IMT)-2000 activity is to provide seamless satellite and terrestrial operation

Ford Motor Co.
Exhibit 1030
Page 321

TABLE 14.2 ITU-R Study Group Structure

SG 1	Spectrum management
SG 3	Radio wave propagation
SG 4	Fixed satellite service
SG 7	Science services
SG 8	Mobile, radio determination, amateur and related satellite services
SG 9	Fixed service
SG 10	Broadcasting services: sound
SG 11	Broadcasting services: television

of mobile terminals throughout the world—anywhere, anytime—where communication coverage requires interoperation of satellite and terrestrial networks. This is to be accomplished using technology available around the year 2000.

- Mobile-satellite and radionavigation-satellite service (MSS-RNSS). The rapid growth of service in these areas has created a need to focus attention on interference and spectrum allocation.
- Wireless Access Systems (WAS). This is an application of radio technology and personal communications systems directed toward lowering the installation and maintenance cost of the local access network. The traditional high cost has prevented penetration of basic telephone service in evolving and developing countries of the world. Overcoming this barrier is an objective of the BDT, described next.

14.2.3 BDT

Unlike the ITU-T (and to some extent ITU-R), which deals with standardization, the BDT deals with aspects that promote the integration and deployment of communications in developing countries. Typical outputs from this sector include implementation guides that expand the utility of ITU recommendations and ensure their expeditious implementation. Communications has been recognized as a necessary element for economic growth. The BDT also seeks to arrange special financing involving communication suppliers and governments or authorized carriers within developing countries to enable provision of basic communications service where otherwise it would not be possible.

14.2.4 ISO/IEC JTC 1

Two global organizations are active in the information processing systems area, the International Standards Organization (ISO) and the International Electrotechnical Commission (IEC), particularly through the Joint Technical Committee 1 (JTC 1).

The ISO comprises national standards bodies, which have the responsibility for promoting and distributing ISO standards within their own countries. ISO technical work is carried out by some 200 technical committees (TC). Technical committees are established by the ISO council and their work program is approved by the technical board on behalf of the council.

The IEC comprises national committees (one from each country) and deals with almost all spheres of electrotechnology, including power, electronics, telecommunications, and nuclear energy. IEC technical work is performed by some 200 TCs set up by its council and some 700 working groups. Part of this organization, a President's Advisory Committee on future technology (PACT) advises the IEC president on new technologies which require preliminary

Ford Motor Co.
Exhibit 1030
Page 322

or immediate standardization work. PACT is designed to form a direct link with private and public research and development activities, keeping the IEC abreast of accelerating technological changes and the accompanying demand for new standards. Small industrial project teams examine new work initiatives which can be introduced into the regular IEC working structure.

In 1987 a joint technical committee was established incorporating ISO TC97, IEC TC83, and subcommittee 47B to deal with generic information technology. The international standards developed by JTC1 are published under the ISO and IEC logos. The activities of ISO/IEC/JTC 1 are listed in Table 14.3 expressed in terms of its subcommittees (SC).

TABLE 14.3 ISO/IEC/JTC1 Subcommittees

SC 1	Vocabulary
SC 2	Coded character sets
SC 6	Telecommunications information exchange between systems
SC 7	Software engineering
SC 11	Flexible magnetic media for digital data interchange
SC 17	Identification cards and related devices
SC 22	Programming languages, their environments and systems software interfaces
SC 23	Optical disk cartridges for information interchange
SC 24	Computer graphics and image processing
SC 25	Interconnection information technology management
SC 26	Microprocessor systems
SC 27	IT security techniques
SC 28	Office equipment
SC 29	Coding of audio, picture, multimedia and hypermedia information
SC 31	Automatic data capture
SC 32	Data management services
SC 33	Distributed application services

The ISO and IEC jointly issue directions for the work of the technical committees. The scope (or area of activity) of each technical committee (TC)/subcommittee (SC) is defined by the TC/SC itself, and then submitted to the Committee of Action (CA)/parent TC for approval. The TCs/SCs prepare technical documents on specific subjects within their respective scopes, which area then submitted to the National Committees for voting with a view to their approval as international standards.

14.3 Regional Standardization

Today the ETSI comes closest to being a true regional standards setting body, together with CITEL, the regional (Latin-American) standardization body.

ETSI is the result of the Single Act of the European community and the EC commission green paper in 1987 that analyzed the consequences of the Single Act and recommended that a European telecommunications standards body be created to develop common standards for telecommunications equipment and networks. Out of this recommendation, the Committee for Harmonization (CCH) and the European Conference for Post and Telecommunications (**CEPT**) evolved into ETSI, which formally came into being in March 1988. It should be noted, however, that even though ETSI attributes at least part of its existence to the European Community, its membership is wider than just the European Union Nations.

Ford Motor Co.
Exhibit 1030
Page 323

Because of the way ETSI came into being, ETSI is characterized by a unique aspect, namely, it is often called upon by the European Commission to develop standards that are necessary to implement legislation. Such standards, which are referred to as technical basis reports (TBR) and whose application is usually mandatory, are often needed in public procurements, as well as in provisioning for open network interconnection as national telecommunications administrations are being deregulated. Like ITU, however, ETSI also develops voluntary standards in accordance with common international understanding against which industry is not obliged to produce conforming products. These standards fall into either the European technical standard (**ETS**) class when fully approved, or into the interim-ETS class, when not fully stable or proven.

ETSI standards are typically sought when either the subject matter is not studied at the global level (such as when it may be required to support some piece of legislation), or the development of the standard is justified by market needs that exist in Europe and not in other parts of the world. In some cases, it may be necessary to adapt ITU standards for the European continent, although a simple endorsement of an ITU standard as a European standard is also possible. A more delicate case arises when both the ITU and ETSI are pursuing parallel standards activities, in which case close coordination with the ITU is sought either through member countries that may input ETSI standards to the ITU for consideration or through the global standards collaboration process.

The highest authority of ETSI is the general assembly, which determines ETSI's policy, appoints its director and deputy, adopts the budget, and approves the audited accounts. The more technical issues are addressed by the technical assembly, which approves technical standards, advises on the work to be undertaken, and sets priorities. The ETSI technical committees are listed in Table 14.4.

TABLE 14.4 ETSI Technical Committees

TCEE	Environmental engineering
TCHF	Human factors
TCMTS	Methods for testing and specification
TCSEC	Security
TCSPS	Signalling protocols and switching
TCTM	Transmission and multiplexing
TCERM	EMC and radio spectrum matters
TCICC	Integrated circuit cards
TCNA	Network aspects
TCSES	Satellite earth stations and systems
TCSTQ	Speech processing, transmission and quality
TCTMN	Telecommunications management networks
ECMA TC32	Communication, networks and systems interconnection
EBU/CENELEC/ETSI JTC	Joint technical committee

It can be seen that ETSI currently comprises 14 technical committees reporting to the technical assembly. These committees are responsible for the development of technical standards. In addition, these committees are responsible for prestandardization activities, that is, activities lead to ETSI technical reports (ETR) that eventually become the basis for future standards.

Ford Motor Co.
Exhibit 1030
Page 324

In addition to the technical assembly, a strategic review committee (SRC) is responsible for prospective examination of a single technical domain, whereas an intellectual property rights committee defines ETSI's policy in the area of intellectual property. Although by no means unique to ETSI, the rapid pace of technological progress has resulted in more standards being adopted that embrace technologies that are still under patent protection. This creates a fundamental conflict between the private, exclusive nature of industrial property rights, and the open, public nature of standards. Harmonizing those conflicting claims has emerged as a thorny issue in all standards organizations; ETSI has established a formal function for this purpose. Finally, the ETS/EBU technical committee coordinates activities with the European broadcasting union (EBU), whereas the ISDN committee is in charge of managing and coordinating the standardization process for narrowband ISDN.

14.3.1 CITEL

On June 11, 1993, the Organization of American States (OAS) General Assembly revised the existing Inter-American Telecommunication Commission (CITEL) strengthening and reorganizing the activities of CITEL, creating a position for the executive secretariat of CITEL and opening the doors, as associate members, to enterprises, organisms, and private telecommunication organizations, to act as observers of the permanent consultative committees of CITEL and its working groups.

CITEL's objectives include facilitating and promoting the continuous development of telecommunications in the hemisphere. It serves as the organization's principal advisory body on matters related to telecommunications. The commission represents all the members states. It has a permanent executive committee consisting of 11 members, and three permanent consultative committees. The permanent consultative committees, whose members are all member states of the organization, also have associate members that represent various private telecommunications agencies or companies.

The general assembly of CITEL, through resolution CITEL Res.8(I-94) established the following specific mandates for the three permanent consultative committees and the steering committee.

Permanent Consultative Committee I: Public Telecommunication Services. To promote and watch over the integration and strengthening of networks and public telecommunication services operating in the countries of the Americas, taking into account the need for modernization of networks and promotion of universal telephone basic services, as well as for increasing the public availability of specialized services and the promotion of the use of international ITU standards and radio regulations.

Permanent Consultative Committee II: Broadcasting. To stimulate and encourage the regional presence of broadcasting services, promoting the use of modern technologies and improving the public availability of such communication media, including audio and video systems, and the promotion of the use of international ITU standards and radio regulations.

Permanent Consultative Committee III: Radiocommunications. To promote the harmonization of radiocommunication services bearing especially in mind the need for a reduction to the minimum of those factors that may cause harmful interferences in the performance and operation of networks and services. To promote the use of modern technologies and the application of the ITU radio regulations and standards.

Steering Committee. The Steering Committee shall be formed by the chairman and vice-chairman of COM/CITEL and the chairman of the PCCs. The committee will be responsible for the revision and proposal to COM/CITEL of the continuous updating of

Ford Motor Co.
Exhibit 1030
Page 325

the regulations, mandates and work programs of CITEL bodies; the executive secretary of CITEL will act as the secretary of said committee.

14.4 National Standardization

As standardization moves from global to regional and then to national levels, the number of actual participating entities rapidly grows. Here, the function of two national standards bodies are reviewed, primarily because these have been in existence the longest and secondarily because they also represent major markets for commercial communications.

14.4.1 ANSI T1

Unlike the ETSI, which came into being partly as a consequence of legislative recommendations, the ANSI Committee T1 on telecommunications came into being as a result of the realization that with the breakup of the Bell System, de-facto standards could no longer be expected. In fact, T1 came into being the very same year (1984) that the breakup of the Bell System came into effect.

The T1 membership comprises four types of interest groups: users and general interest groups, manufacturers, interexchange carriers, and exchange carriers. This rather broad membership is reflected, to some extent, by the scope to which T1 standards are being applied; this means that nontraditional telecommunications service providers are utilizing the technologies standardized by committee T1. This situation is the result of the rapid evolution and convergence of the telecommunications, computer, and cable television industries in the United States, and advances in wireless technology.

Committee T1 currently addresses approximately 150 approved projects, which led to the establishment of six, primarily functionally oriented, technical subcommittees (TSC), as shown in Table 14.5 and Fig. 14.2 [although not evident from Table 14.3, subcommittee T1P1 has primary responsibility for management of activities on personal communications systems (PCS)]. In-turn, each of these six subcommittees is divided into a number of subtending working groups, and subworking groups.

TABLE 14.5 T1 Subcommittee Structure

TSC: T1A1	Performance and signal processing
TSC: T1E1	Network interfaces and environmental considerations
TSC: T1M1	Interwork operations, administration, maintenance, and provisioning
TSC: T1P1	Systems engineering, standards planning, and program management
TSC: T1S1	Services, architecture, and signalling
TSC: T1X1	Digital hierarchy and synchronization

Committee T1 also has an advisory group (**T1AG**) made up of elected representatives from each of the four interest groups to carry out committee T1 directives and to develop proposals for consideration by the T1 membership.

In parallel to serving as the forum that establishes ANSI telecommunications network standards, committee T1 technical subcommittees draft candidate U.S. technical contributions to the ITU. These contributions are submitted to the U.S. Department of State National Committee for the ITU, which administers U.S. participation and contributions to the ITU (see Fig. 14.3). In this manner, activities within T1 are coordinated with

Ford Motor Co.
Exhibit 1030
Page 326

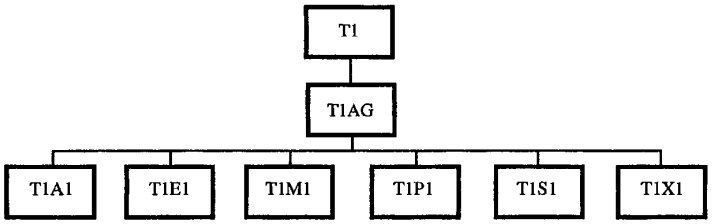


FIGURE 14.2 T1 committee structure.

those of the ITU. This coordination with other standards setting bodies is also reflected in T1's involvement with Latin-American standards, through the formation of an ad hoc group with CITEL's permanent technical committee 1 (PTC 1/T1). Further coordination with ETSI and other standards setting bodies is accomplished through the global standards collaboration process.

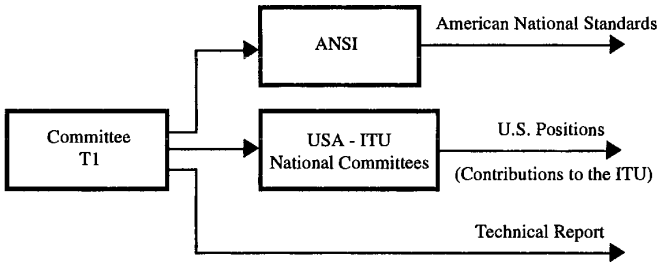


FIGURE 14.3 Committee T1 output.

14.4.2 TIA

The TIA is a full-service trade organization that provides its members with numerous services including government relations, market support activities, educational programs, and standards setting activities.

TIA is a member-driven organization. Policy is formulated by 25 board members selected from member companies, and is carried out by a permanent professional staff located in Washington D.C. TIA comprises six issue-oriented standing committees, each of which is chaired by a board member. The six committees are membership scope and development, international, marketing and trade shows, public policy and government relations, and technical. It is this last committee that in 1992 was accredited by ANSI in the United States to standardize telecommunications products. Technology standardization activities are reflected by TIA's four product-oriented divisions, namely, user premises equipment, network equipment, mobile and personal communications equipment, and fiber optics.

In these divisions the legislative and regulatory concerns of product manufacturers and the preparation of standards dealing with performance testing and compatibility are addressed. For example, modem and telematic standards, as well as much of the cellular standards technology, has been standardized in the United States under the mandate of TIA.

Ford Motor Co.
Exhibit 1030
Page 327

14.4.3 TTC

The third national committee to be addressed is the TTC in Japan. TTC was established in October 1985 to develop and disseminate Japanese domestic standards for deregulated technical items and protocols. It is a nongovernmental, nonprofit standards setting organization established to ensure fair and transparent standardization procedures.

TTC's primary emphasis is to develop, conduct studies and research, and disseminate protocols and standards for the connection of telecommunications networks. TTC is organized along six technical subcommittees that report to a board of directors through a technical assembly (see Fig. 14.4).

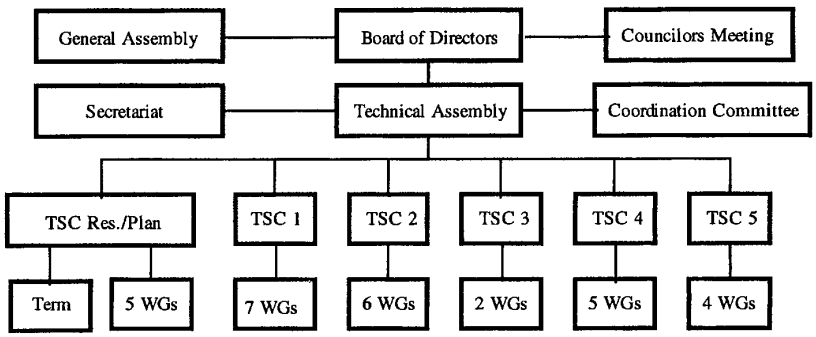


FIGURE 14.4 Organization of TTC.

The TTC organization comprises a general assembly, which is in charge of matters such as business plans and budgets. The councilors meeting examines standards development procedures in order to assure impartiality and clarity. The secretariat provides overall support to the organization; the technical assembly develops standards and handles technical matters including surveys and research. Each technical subcommittee is partitioned into two or more working groups (WG). The coordination committee handles all issues in or between the TSCs and WGs, and it assures the smooth running of all technical committee meetings.

Under the coordination committee, a subcommittee examines users' requests and studies their applicability to the five-year standardization-project plan. This subcommittee also conducts user-request surveys. The areas of involvement of each of the five subcommittees are shown in Table 14.6.

TTC membership is divided into four categories. Type I telecommunications carriers, that is, those carriers that own telecommunications circuits and facilities; type II telecom-

TABLE 14.6 TTC SubCommittees

Strategic Research and Planning Committee: Technical Survey and International Collaboration	
TSC 1	Network-to-network interfaces, mobile communications
TSC 2	User-network interfaces
TSC 3	PBX, LAN
TSC 4	Higher level protocols
TSC 5	Voice and video signal coding scheme and systems

Ford Motor Co.
Exhibit 1030
Page 328

munications carriers, that is, those with telecommunications circuits leased from type I carriers; related equipment manufacturers; and others, including users.

Underlying objectives that guide TTC's approach to standards development are 1) to conform to international recommendations or standards; 2) standardize items, where either international recommendations or standards are not clear, or where national standards need to be set, and where a consensus is achieved; and 3) to conduct further studies into any of the items just mentioned whenever the technical assembly is unable to arrive at a consensus.

These objectives, which give highest priority in developing standards that are compatible with international recommendations or standards, have often driven TTC to adapt international standards for national use through the use of supplements that:

- Give guidelines for users of TTC standards on how to apply them
- Help clarify the contents of standards
- Help with the implementation of standards in terminal equipment and adaptors
- Assure interconnection between terminal equipment and adaptors
- Provide background information regarding the content of standards
- Assure interconnection.

These supplements also include questions and answers that help in implementing the standards, including encoding examples of various parameters and explanation of the practical meaning of a standard.

14.5 Intellectual Property

In the deregulating telecommunication arena patents have become increasingly more important. New ideas that are incorporated in standards often have global market potential and patent holders are seeking to obtain an income from their intellectual property as well as from products. In addition, the general effort to develop standards quickly places them closer to the leading edge of technology. There are some cases, for example speech encoding algorithms, where terms of reference for performance are typically set as objectives that no one can meet when the objectives are defined. The state of the art is being pushed by goals of the standards development organization. In this environment, incorporation of some intellectual property in standards is practically unavoidable.

With regard to intellectual property rights in the ITU, the TSB has developed a "code of practice" which may be summarized as follows.

The TSB requests members putting forth standards to draw the attention of the TSB to any known patent or patent pending application relevant to the developing standard. Where such information has been declared to the TSB, a log of registered patent holders for each affected recommendation is maintained for the convenience of users of ITU standards. If a recommendation, which is a nonbinding international standard, is developed and contains patented intellectual property there are three situations that may arise.

- The patent holder waives the rights and the recommendation is freely accessible to everybody.
- The patent holder will not waive the rights but is willing to negotiate licenses with other parties on a nondiscriminatory basis and on reasonable terms and conditions. What is reasonable is not defined, and the ITU-T will not participate in such negotiations.

Ford Motor Co.
Exhibit 1030
Page 329

- The patent holder is not willing to comply with either of the above two situations, in which case the ITU-T will not approve a recommendation containing such intellectual property.

The patent policy of the American National Standards Institute (ANSI), which governs all standards development organizations accredited by ANSI, is defined in ANSI procedures 1.2.11. It is similar to that of the ITU in that it requires a statement from patent holders or identified parties to indicate granting of a royalty-free license, willingness to license on reasonable and nondiscriminatory terms and conditions, or a disclaimer of no patent. Unlike the ITU, ANSI advises that is prepared to get involved in resolving disputes of what is considered "nondiscriminatory" and "reasonable." Additional information on ANSI patent guidelines can be found at `http://web.ansi.org/public/library/guides/ppguide.html`.

As mentioned earlier ETSI produces a combination of mandatory and voluntary standards. This can create additional complications when intellectual property issues are encapsulated within the standards. To formally address these issues an intellectual property rights committee defines ETSI's policy in the area of intellectual property.

Given the different patent policies adopted by various standards organizations, it is recommended that companies developing products based on standards investigate and understand the patent policy of the associated standards body and the patent statements filed regarding the standard being implemented.

14.6 Standards Coordination

The pace of technological advancements coupled with deregulation has given rise to increased global telecommunications standards activities. At the same time a growth of regional standards bodies has occurred which has increased the potential for duplication of work, wasting resources, and creating conflicting standards. This potentially adverse situation was addressed by a number of interregional telecommunications standardization conferences (ITSCs) that were held in the early 1990s. A global standards collaboration (GSC) group was established to oversee collaborative activities including electronic document handling (EDH) and five high-interest standards subjects:

- Broadband integrated services digital network (B-ISDN)
- Intelligent Networks (IN)
- Transmission management network (TMN)
- Universal personal telecommunications (UPT)
- Synchronous digital hierarchy/synchronous optical network (SDH/SONET)

This early activity was successful in avoiding duplication of effort and coordinating activities on these major standardization efforts. Today the level of cooperative activities, again driven by the pressure to avoid wasting valuable resources and reaching agreed standards more rapidly, are being driven to lower levels through the use of liaison statements between regional standards groups and permitting "documents of information" to flow between standards development organizations. The processes for this information flow are evolving and the electronic addresses provided at the end of this chapter should be consulted for the current interstandards organization communication mechanisms.

Ford Motor Co.
Exhibit 1030
Page 330

14.7 Scientific

Another global, scientifically based organization that has been particularly active in standards development (more recently emphasizing information processing) is the IEEE. Responsibility for standards adoption within the IEEE lies with the IEEE standards board. The board is supported by nine standing committees (see Fig. 14.5).

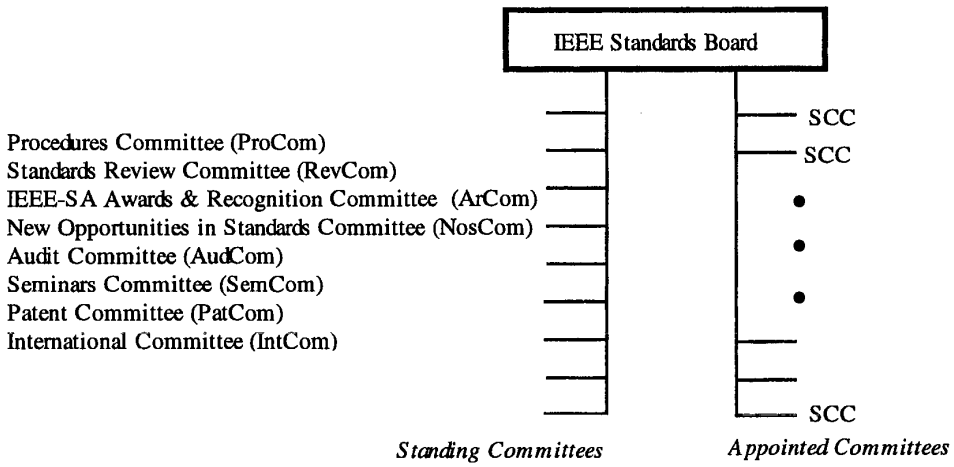


FIGURE 14.5 IEEE standards board organization.

Proposed standards are normally developed in the technical committees of the IEEE societies. There are occasions, however, when the scope of activity is too broad to be encompassed by a single society or where the societies are not able to do so for other reasons. In this case the standards board establishes its own standards developing committees, namely, the standards coordinating committees (**SCC**), to perform this function.

The adoption of IEEE standards is based on projects that have been approved by the IEEE standards board, while each project is the responsibility of a sponsor. Sponsors need not be an SCC, but can also include technical committees of IEEE societies; a standards, or standards coordinating committee of an IEEE Society; an accredited standards committee; or another organization approved by the IEEE standards board.

14.8 Standards Development Cycle

Although the manner in which standards are developed and approved somewhat varies between standards organizations, there are common characteristics to be found.

For most standards, first a set of requirements is defined. This may be done either by the standards committee actually developing the standard or by another entity in collaboration with such a committee. Subsequently, the technical details of a standard are developed. The actual entity developing a standard may be a member of the standards committee, or the actual standards committee itself. Outsiders may also contribute to standards development but, typically, only if sponsored by a committee member. Membership in the standards committee and the right to contribute technical information towards the development of

Ford Motor Co.
Exhibit 1030
Page 331

the standard differs among the various standards' organizations, as indicated. This process is illustrated in Fig. 14.6.

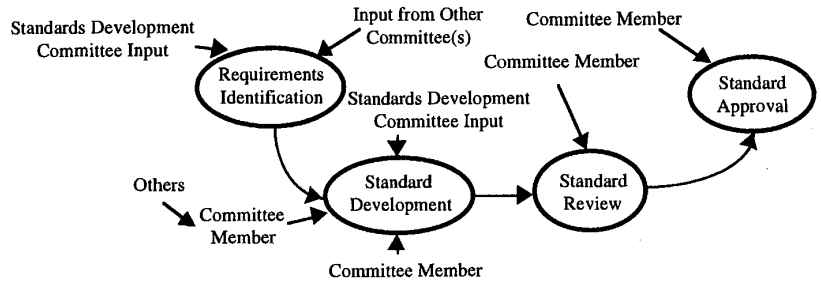


FIGURE 14.6 Typical standards development and approval process.

Finally, once the standard has been fully developed, it is placed under an approval cycle. Each standards setting body typically has precisely defined and often complex procedures for reviewing and then approving proposed standards, which although different in detail, are typically consensus driven.

Defining Terms

ANSI: The American National Standards Institute.

CCIR: The International Radio Consultative Committee, the predecessor of the ITU-R.

CCITT: The International Telephone and Telegraph Consultative Committee, the predecessor of the ITU-T.

CEPT: The European Conference for Post and Telecommunications, a predecessor of ETSI.

CITEL: Inter-American Telecommunications Commission, a standards setting body for the Americas.

ETS: A European (ETSI) technical standard.

ETSI: The European Telecommunications Standards Institute.

GSC: The Global Standards Collaboration group.

ICG: Intersector Coordination Group, a group which coordinates activities between the ITU-T and ITU-R.

IEC: The International Electrotechnical Commission.

IEEE: The Institute of Electrical and Electronics Engineers.

ISO: The International Standards Organization.

ITU: The International Telecommunications Union, an international treaty organization, which is part of the United Nations.

ITU-R: The radio communications sector of the ITU, the successor of the CCIR.

ITU-T: The standardization sector of the ITU, the successor of the CCITT.

JCG: The Joint Coordination Group, which oversees the coordination of common work between ITU-T study groups.

Ford Motor Co.
Exhibit 1030
Page 332

Recommendation: An ITU technical standard.

SCC: A standard's coordinating committee within the IEEE organization.

Standard: A publicly approved technical specification.

T1: An ANSI-approved standards body, which develops telecommunications standards in the United States.

T1AG: The primary advisory group within ANSI Committee T1 on Telecommunications.

TIA: The Telecommunications Industry Association, which is an ANSI-approved standards body that develops terminal equipment standards.

TTC: The Telecommunications Technology Committee, a Japanese standards setting body.

Further Information

[1] Irmer, T., Shaping future telecommunications: the challenge of global standardization, *IEEE Comm. Mag.*, 32(1), 20–28, 1994.

[2] Matute, M.A., CITEL: formulating telecommunications in the Americas. *IEEE Comm. Mag.*, 32(1), 38–39, 1994.

[3] Robin, G., The European perspective for telecommunications standards. *IEEE Comm. Mag.*, 32(1), 40–50, 1994.

[4] Reilly, A.K., A U.S. perspective on standards development. *IEEE Comm. Mag.*, 32(1), 30–36, 1994.

[5] Iida, T., Domestic standards in a changing world. *IEEE Comm. Mag.*, 32(1), 46–50, 1994.

[6] Habara, K., Cooperation in standardization. *IEEE Comm. Mag.*, 32(1), 78–84, 1994.

[7] *IEEE Standards Board Bylaws.* Institute of Electrical and Electronics Engineers. Dec. 1993.

[8] Chiarottino, W. and Pirani, G., International telecommunications standards organizations, *CSELT Tech. Repts.*, XXI(2), 207–236, 1993.

[9] ITU, Book No. 1. Resolutions; Recommendations on the organization of the work of ITU-T (series A); study groups and other groups; list of study questions (1993-1996). World Standardization Conf. Helsinki, 1–12, Mar. 1993.

[10] Standards Committee T1., *Telecommunications.* Procedures Manual. 7th Iss. Jun. 1992.

The standards' organizations often undergo structural and substantive changes. It is recommended that the following web sites be visited for the most updated information.

ANSI	`http://www.ansi.org/`
CITEL	`http://www.oas.org`
ETSI	`http://www.etsi.org`
IEC	`http://www.iec.ch`
IEEE	`http://www.ieee.org`
ISO	`http://www.iso.ch`
ITU	`http://www.itu.ch`
T1	`http://www.t1.org`
TIA	`http://www.tia.org`
TTC	`http://www.ttc.or.jp`

Ford Motor Co.
Exhibit 1030
Page 333

II

Wireless

15 Wireless Personal Communications: A Perspective *Donald C. Cox* 15-1
Introduction • Background and Issues • Evolution of Technologies, Systems, and Services • Evolution Toward the Future and to Low-Tier Personal Communications Services • Comparisons with Other Technologies • Quality, Capacity, and Economic Issues • Other Issues • Infrastructure Networks • Conclusion

16 Modulation Methods *Gordon L. Stüber* 16-1
Introduction • Basic Description of Modulated Signals • Analog Frequency Modulation • Phase Shift Keying (PSK) and $\pi/4$-QPSK • Continuous Phase Modulation (CPM) and MSK • Gaussian Minimum Shift Keying • Orthogonal Frequency Division Multiplexing (OFDM) • Conclusions

17 Access Methods *Bernd-Peter Paris* 17-1
Introduction • Relevant Wireless Communication System Characteristics • Frequency Division Multiple Access • Time Division Multiple Access • Code Division Multiple Access • Comparison and Outlook

18 Rayleigh Fading Channels *Bernard Sklar* 18-1
Introduction • The Challenge of a Fading Channel • Mobile-Radio Propagation: Large-Scale Fading and Small-Scale Fading • Signal Time-Spreading Viewed in the Time-Delay Domain: Figure 18.1, Block 7—The Multipath Intensity Profile • Signal Time-Spreading Viewed in the Frequency Domain: Figure 18.1, Block 10—The Spaced-Frequency Correlation Function • Typical Examples of Flat Fading and Frequency-Selective Fading Manifestations • Time Variance Viewed in the Time Domain: Figure 18.1, Block 13—The Spaced-Time Correlation Function • Time Variance Viewed in the Doppler-Shift Domain: Figure 18.1, Block 16—The Doppler Power Spectrum • Analogy Between Spectral Broadening in Fading Channels and Spectral Broadening in Digital Signal Keying • Degradation Categories due to Time Variance, Viewed in the Doppler-Shift Domain • Mitigation Methods • Summary of the Key Parameters Characterizing Fading Channels • The Viterbi Equalizer as Applied to GSM • The Rake Receiver Applied to Direct-Sequence Spread-Spectrum (DS/SS) Systems • Conclusion

19 Space-Time Processing *Arogyaswami J. Paulraj* 19-1
Introduction • The Space-Time Wireless Channel • Signal Models • ST Receive Processing (Base) • ST Transmit Processing (Base) • Summary

Ford Motor Co.
Exhibit 1030
Page 334

20 Location Strategies for Personal Communications Services *Ravi Jain, Yi-Bing Lin, and Seshadri Mohan* 20-1
Introduction • An Overview of PCS • IS-41 Preliminaries • Global System for Mobile Communications • Analysis of Database Traffic Rate for IS-41 and GSM • Reducing Signalling During Call Delivery • Per-User Location Caching • Caching Threshold Analysis • Techniques for Estimating Users' LCMR • Discussion • Conclusions

21 Cell Design Principles *Michel Daoud Yacoub* 21-1
Introduction • Cellular Principles • Performance Measures and System Requirements • System Expansion Techniques • Basic Design Steps • Traffic Engineering • Cell Coverage • Interference • Conclusions

22 Microcellular Radio Communications *Raymond Steele* 22-1
Introducing Microcells • Highway Microcells • City Street Microcells • Indoor Microcells • Microcellular Infrastructure • Multiple Access Issues • Discussion

23 Fixed and Dynamic Channel Assignment *Bijan Jabbari* 23-1
Introduction • The Resource Assignment Problem • Fixed Channel Assignment • Enhanced Fixed Channel Assignment • Dynamic Channel Assignment • CDMA Systems • Conclusion

24 Radiolocation Techniques *Gordon L. Stüber and James J. Caffery, Jr.* 24-1
Introduction • Description of Radiolocation Methods • Location Algorithms • Measures of Location Accuracy • Location in Cellular Systems • Sources of Location Error • Summary and Conclusions

25 Power Control *Roman Pichna and Qiang Wang* 25-1
Introduction • Cellular Systems and Power Control • Power Control Examples • Summary

26 Enhancements in Second Generation Systems *Marc Delprat and Vinod Kumar* 26-1
Introduction • Overview of Second Generation Systems • Capacity Enhancement • Quality Enhancement • High Bit Rate Data Transmission • Conclusion

27 The Pan-European Cellular System *Lajos Hanzo* 27-1
Introduction • Overview • Logical and Physical Channels • Speech and Data Transmission • Transmission of Control Signals • Synchronization Issues • Gaussian Minimum Shift Keying Modulation • Wideband Channel Models • Adaptive Link Control • Discontinuous Transmission • Summary

28 Speech and Channel Coding for North American TDMA Cellular Systems *Paul Mermelstein* 28-1
Introduction • Modulation of Digital Voice and Data Signals • Speech Coding Fundamentals • Channel Coding Considerations • VSELP Encoder • Linear Prediction Analysis and Quantization • Bandwidth Expansion • Quantizing and Encoding the Reflection Coefficients • VSELP Codebook Search • Long-Term Filter Search • Orthogonalization of the Codebooks • Quantizing the Excitation and Signal Gains • Channel Coding and Interleaving • Bad Frame Masking • ACELP Encoder • Algebraic Codebook Structure and Search • Quantization of the Gains for ACELP Encoding • Channel Coding for ACELP Encoding • Conclusions

29 The British Cordless Telephone Standard: CT-2 *Lajos Hanzo* 29-1
History and Background • The CT-2 Standard • The Radio Interface • Burst Formats • Signalling Layer Two (L2) • CPP-Initiated Link Setup Procedures • CFP-Initiated Link Setup Procedures • Handshaking • Main Features of the CT-2 System

30 Half-Rate Standards *Wai-Yip Chan, Ira Gerson, and Toshio Miki* 30-1

Ford Motor Co.
Exhibit 1030
Page 335

Introduction • Speech Coding for Cellular Mobile Radio Communications • Codec Selection and Performance Requirements • Speech Coding Techniques in the Half-Rate Standards • Channel Coding Techniques in the Half-Rate Standards • The Japanese Half-Rate Standard • The European GSM Half-Rate Standard • Conclusions

31 Wireless Video Communications *Madhukar Budagavi and Raj Talluri* 31-1
Introduction • Wireless Video Communications • Error Resilient Video Coding • MPEG-4 Error Resilience Tools • H.263 Error Resilience Tools • Discussion

32 Wireless LANs *Suresh Singh* 32-1
Introduction • Physical Layer Design • MAC Layer Protocols • Network Layer Issues • Transport Layer Design • Conclusions

33 Wireless Data *Allen H. Levesque and Kaveh Pahlavan* 33-1
Introduction • Characteristics of Wireless Data Networks • Market Issues • Modem Services Over Cellular Networks • Private Data Networks • Cellular Data Networks and Services • Other Developing Standards • Conclusions

34 Wireless ATM: Interworking Aspects *Melbourne Barton, Matthew Cheng, and Li Fung Chang* 34-1
Introduction • Background and Issues • Wireless Interworking With Transit ATM Networks • The PCS-to-ATM Interworking Scenario • QoS Support • Conclusions

35 Wireless ATM: QoS and Mobility Management *Bala Rajagopalan and Daniel Reininger* 35-1
Introduction • QoS in Wireless ATM • Mobility Management in Wireless ATM • Summary and Conclusions

36 An Overview of cdma2000, WCDMA, and EDGE *Tero Ojanperä and Steven D. Gray* 36-1
Introduction • CDMA-Based Schemes • CDMA System Design Issues • WCDMA • cdma2000 • TDMA-Based Schemes • Time Division Duplex (TDD) • Conclusions

Ford Motor Co.
Exhibit 1030
Page 336

15

Wireless Personal Communications: A Perspective

15.1 Introduction 15-1
15.2 Background and Issues 15-3
Mobility and Freedom from Tethers
15.3 Evolution of Technologies, Systems, and Services 15-5
Cordless Telephones • Cellular Mobile Radio Systems • Wide-Area Wireless Data Systems • High-Speed Wireless Local-Area Networks (WLANs) • Paging/Messaging Systems • Satellite-Based Mobile Systems • {Fixed Point-to-Multipoint Wireless Loops • Reality Check
15.4 Evolution Toward the Future and to Low-Tier Personal Communications Services 15-21
15.5 Comparisons with Other Technologies 15-25
Complexity/Coverage Area Comparisons • {Coverage, Range, Speed, and Environments
15.6 Quality, Capacity, and Economic Issues 15-29
Capacity, Quality, and Complexity • Economics, System Capacity, and Coverage Area Size • {Loop Evolution and Economics
15.7 Other Issues 15-34
Improvement of Batteries • People Only Want One Handset • Other Environments • Speech Quality Issues • New Technology • High-Tier to Low-Tier or Low-Tier to High-Tier Dual Mode
15.8 Infrastructure Networks 15-44
15.9 Conclusion 15-45
References 15-46

Donald C. Cox
Stanford University

{ This chapter has been updated using { } as indicators of inserts into the text of the original chapter of the same title that appeared in the first edition of this Handbook in 1996. }

15.1 Introduction

Wireless personal communications has captured the attention of the media and with it, the imagination of the public. Hardly a week goes by without one seeing an article on the subject appearing in a popular U.S. newspaper or magazine. Articles ranging from a short

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 337

paragraph to many pages regularly appear in local newspapers, as well as in nationwide print media, e.g., *The Wall Street Journal, The New York Times, Business Week,* and *U.S. News and World Report.* Countless marketing surveys continue to project enormous demand, often projecting that at least half of the households, or half of the people, want wireless personal communications. Trade magazines, newsletters, conferences, and seminars on the subject by many different names have become too numerous to keep track of, and technical journals, magazines, conferences, and symposia continue to proliferate and to have ever increasing attendance and numbers of papers presented. It is clear that wireless personal communications is, by any measure, the fastest growing segment of telecommunications. { The explosive growth of wireless personal communications has continued unabated worldwide. Cellular and high-tier PCS pocketphones, pagers, and cordless telephones have become so common in many countries that few people even notice them anymore. These items have become an expected part of everyday life in most developed countries and in many developing countries around the world. }

If you look carefully at the seemingly endless discussions of the topic, however, you cannot help but note that they are often describing different things, i.e., different versions of wireless personal communications [29, 50]. Some discuss pagers, or messaging, or data systems, or access to the national information infrastructure, whereas others emphasize cellular radio, or cordless telephones, or dense systems of satellites. Many make reference to popular fiction entities such as Dick Tracy, Maxwell Smart, or *Star Trek.* { In addition to the things noted above, the topic of wireless loops [24], [30], [32] has also become popular in the widespread discussions of wireless communications. As discussed in [30], this topic includes several fixed wireless applications as well as the low-tier PCS application that was discussed originally under the wireless loop designation [24, 32]. The fixed wireless applications are aimed at reducing the cost of wireline loop-ends, i.e., the so-called "last mile" or "last km" of wireline telecommunications. }

Thus, it appears that almost everyone wants wireless personal communications, but *What is it?* There are many different ways to segment the complex topic into different communications applications, modes, functions, extent of coverage, or mobility [29, 30, 50]. The complexity of the issues has resulted in considerable confusion in the industry, as evidenced by the many different wireless systems, technologies, and services being offered, planned, or proposed. Many different industry groups and regulatory entities are becoming involved. The confusion is a natural consequence of the massive dislocations that are occurring, and will continue to occur, as we progress along this large change in the paradigm of the way we communicate. Among the different changes that are occurring in our communications paradigm, perhaps the major constituent is the change from wired fixed place-to-place communications to wireless mobile person-to-person communications. Within this major change are also many other changes, e.g., an increase in the significance of data and message communications, a perception of possible changes in video applications, and changes in the regulatory and political climates. { The fixed wireless loop applications noted earlier do not fit the new mobile communications paradigm. After many years of decline of fixed wireless communications applications, e.g., intercontinental HF radio and later satellites, point-to-point terrestrial microwave radio, and tropospheric scatter, it is interesting to see this rebirth of interest in fixed wireless applications. This rebirth is riding on the gigantic "wireless wave" resulting from the rapid public acceptance of mobile wireless communications. It will be interesting to observe this rebirth to see if communications history repeats; certainly mobility is wireless, but there is also considerable historical evidence that wireless is also mobility. }

This chapter attempts to identify different issues and to put many of the activities in wireless into a framework that can provide perspective on what is driving them, and perhaps

Ford Motor Co.
Exhibit 1030
Page 338

even to yield some indication of where they appear to be going in the future. Like any attempt to categorize many complex interrelated issues, however, there are some that do not quite fit into neat categories, and so there will remain some dangling loose ends. Like any major paradigm shift, there will continue to be considerable confusion as many entities attempt to interpret the different needs and expectations associated with the new paradigm.

15.2 Background and Issues

15.2.1 Mobility and Freedom from Tethers

Perhaps the clearest constituents in all of the wireless personal communications activity are the desire for mobility in communications and the companion desire to be free from tethers, i.e., from physical connections to communications networks. These desires are clear from the very rapid growth of mobile technologies that provide primarily two-way voice services, even though economical wireline voice services are readily available. For example, cellular mobile radio has experienced rapid growth. Growth rates have been between 35 and 60% per year in the United States for a decade, with the total number of subscribers reaching 20 million by year-end 1994. The often neglected wireless companions to cellular radio, i.e., cordless telephones, have experienced even more rapid, but harder to quantify, growth with sales rates often exceeding 10 million sets a year in the United States, and with an estimated usage significantly exceeding 50 million in 1994. Telephones in airlines have also become commonplace. Similar or even greater growth in these wireless technologies has been experienced throughout the world. { The explosive growth in cellular and its identical companion, high-tier PCS, has continued to about 55 million subscribers in the U.S. at year-end 1997 and a similar number worldwide. In Sweden the penetration of cellular subscribers by 1997 was over one-third of the total population, i.e., the total including every man, woman, and child! And the growth has continued since. Similar penetrations of mobile wireless services are seen in some other developed nations, e.g., Japan. The growth in users of cordless telephones also has continued to the point that they have become the dominant subscriber terminal on wireline telephone loops in the U.S. It would appear that, taking into account cordless telephones and cellular and high-tier PCS phones, half of all telephone calls in the U.S. terminate with at least one end on a wireless device. }

{ Perhaps the most significant event in wireless personal communications since the writing of this original chapter was the widespread deployment and start of commercial service of personal handphone (PHS) in Japan in July of 1995 and its very rapid early acceptance by the consumer market [53]. By year-end 1996 there were 5 million PHS subscribers in Japan with the growth rate exceeding one-half million/month for some months. The PHS "phenomena" was one of the fastest adoptions of a new technology ever experienced. However, the PHS success story [41] peaked at a little over 7 million subscribers in 1997 and has declined slightly to a little under 7 million in mid-1998. This was the first mass deployment of a low-tier-like PCS technology (see later sections of this chapter), but PHS has some significant limitations. Perhaps the most significant limitation is the inability to successfully handoff at vehicular speeds. This handoff limitation is a result of the cumbersome radio link structure and control algorithms used to implement dynamic channel allocation (DCA) in PHS. DCA significantly increases channel occupancy (base station capacity) but incurs considerable complexity in implementing handoff. Another significant limitation of the PHS standard has been insufficient receiver sensitivity to permit "adequate" coverage from a "reasonably" dense deployment of base stations. These technology deficiencies coupled with heavy price cutting by the cellular service providers to compete with the rapid advancing of the PHS market were significant contributors to the leveling out of PHS growth. It is

Ford Motor Co.
Exhibit 1030
Page 339

again evident, as with CT-2 phone point discussed in a later section, that low-tier PCS has very attractive features that can attract many subscribers, but it must also provide vehicle speed handoff and widespread coverage of highways as well as populated regions.

Others might point out the deployment and start of service of CDMA systems as a significant event since the first edition. However, the major significance of this CDMA activity is that it confirmed that CDMA performance was no better than other less-complex technologies and that those, including this author, who had been branded as "unbelieving skeptics" were correct in their assessments of the shortcomings of this technology. The overwhelming failure of CDMA technology to live up to the early claims for it can hardly be seen as a significant positive event in the evolution of wireless communication. It was, of course, a significant negative event. After years of struggling with the problems of this technology, service providers still have significantly fewer subscribers on CDMA worldwide than there are PHS subscribers in Japan alone! CDMA issues are discussed more in later sections dealing with technology issues. }

Paging and associated messaging, although not providing two-way voice, do provide a form of tetherless mobile communications to many subscribers worldwide. These services have also experienced significant growth { and have continued to grow since 1996. } There is even a glimmer of a market in the many different specialized wireless data applications evident in the many wireless local area network (WLAN) products on the market, the several wide area data services being offered, and the specialized satellite-based message services being provided to trucks on highways. { Wireless data technologies still have many supporters, but they still have fallen far short of the rapid deployment and growth of the more voice oriented wireless technologies. However, hope appears to be eternal in the wireless data arena. }

The topics discussed in the preceding two paragraphs indicate a dominant issue separating the different evolutions of wireless personal communications. That issue is the voice versus data communications issue that permeates all of communications today; this division also is very evident in fixed networks. The packet-oriented computer communications community and the circuit-oriented voice telecommunications (telephone) community hardly talk to each other and often speak different languages in addressing similar issues. Although they often converge to similar overall solutions at large scales (e.g., hierarchical routing with exceptions for embedded high-usage routes), the small-scale initial solutions are frequently quite different. Asynchronous transfer mode (ATM-) based networks are an attempt to integrate, at least partially, the needs of both the packet-data and circuit-oriented communities.

Superimposed on the voice-data issue is an issue of competing modes of communications that exist in both fixed and mobile forms. These different modes include the following.

Messaging is where the communication is not real time but is by way of message transmission, storage, and retrieval. This mode is represented by voice mail, electronic facsimile (fax), and electronic mail (e-mail), the latter of which appears to be a modern automated version of an evolution that includes telegraph and telex. Radio paging systems often provide limited one-way messaging, ranging from transmitting only the number of a calling party to longer alpha-numeric text messages.

Real-time two-way communications are represented by the telephone, cellular mobile radio telephone, and interactive text (and graphics) exchange over data networks. Two-way video phone always captures significant attention and fits into this mode; however, its benefit/cost ratio has yet to exceed a value that customers are willing to pay.

Paging, i.e., broadcast with no return channel, alerts a paged party that someone wants to communicate with him/her. Paging is like the ringer on a telephone without having the capability for completing the communications.

Ford Motor Co.
Exhibit 1030
Page 340

Agents are new high-level software applications or entities being incorporated into some computer networks. When launched into a data network, an agent is aimed at finding information by some title or characteristic and returning the information to the point from which the agent was launched. { The rapid growth of the worldwide web is based on this mode of communications. }

There are still other ways in which wireless communications have been segmented in attempts to optimize a technology to satisfy the needs of some particular group. Examples include 1) user location, which can be differentiated by indoors or outdoors, or on an airplane or a train and 2) degree of mobility, which can be differentiated either by speed, e.g., vehicular, pedestrian, or stationary, or by size of area throughout which communications are provided. { As noted earlier, wireless local loop with stationary terminals has become a major segment in the pursuit of wireless technology. }

At this point one should again ask; wireless personal communications—*What is it?* The evidence suggests that what is being sought by users, and produced by providers, can be categorized according to the following two main characteristics.

Communications portability and mobility on many different scales:

- Within a house or building [cordless telephone, (WLANs)]
- Within a campus, a town, or a city (cellular radio, WLANs, wide area wireless data, radio paging, extended cordless telephone)
- Throughout a state or region (cellular radio, wide area wireless data, radio paging, satellite-based wireless)
- Throughout a large country or continent (cellular radio, paging, satellite-based wireless)
- Throughout the world?

Communications by many different modes for many different applications:

- Two-way voice
- Data
- Messaging
- Video?

Thus, it is clear why wireless personal communications today is not one technology, not one system, and not one service but encompasses many technologies, systems, and services optimized for different applications.

15.3 Evolution of Technologies, Systems, and Services

Technologies and systems [27, 29, 30, 39, 50, 59, 67, 87], that are currently providing, or are proposed to provide, wireless communications services can be grouped into about seven relatively distinct groups, { the seven previous groups are still evident in the technology but with the addition of the fixed point-to-multipoint wireless loops there are now eight, } although there may be some disagreement on the group definitions, and in what group some particular technology or system belongs. All of the technologies and systems are evolving as technology advances and perceived needs change. Some trends are becoming evident in the evolutions. In this section, different groups and evolutionary trends are explored along with factors that influence the characteristics of members of the groups. The grouping is generally with respect to scale of mobility and communications applications or modes.

Ford Motor Co.
Exhibit 1030
Page 341

15.3.1 Cordless Telephones

Cordless telephones [29, 39, 50] generally can be categorized as providing low-mobility, low-power, two-way tetherless voice communications, with low mobility applying both to the range and the user's speed. Cordless telephones using analog radio technologies appeared in the late 1970s, and have experienced spectacular growth. They have evolved to digital radio technologies in the forms of second-generation cordless telephone (CT-2), and digital European cordless telephone (DECT) standards in Europe, and several different industrial scientific medical (ISM) band technologies in the United States.[1]

{ Personal handyphone (PHS) noted earlier and discussed in later sections and inserts can be considered either as a quite advanced digital cordless telephone similar to DECT or as a somewhat limited low-tier PCS technology. It has most of the attributes of similarity of the digital cordless telephones listed later in this section except that PHS uses $\pi/4$ QPSK modulation. }

Cordless telephones were originally aimed at providing economical, tetherless voice communications inside residences, i.e., at using a short wireless link to replace the cord between a telephone base unit and its handset. The most significant considerations in design compromises made for these technologies are to minimize total cost, while maximizing the talk time away from the battery charger. For digital cordless phones intended to be carried away from home in a pocket, e.g., CT-2 or DECT, handset weight and size are also major factors. These considerations drive designs toward minimizing complexity and minimizing the power used for signal processing and for transmitting.

Cordless telephones compete with wireline telephones. Therefore, high circuit quality has become a requirement. Early cordless sets had marginal quality. They were purchased by the millions, and discarded by the millions, until manufacturers produced higher-quality sets. Cordless telephones sales then exploded. Their usage has become commonplace, approaching, and perhaps exceeding, usage of corded telephones.

The compromises accepted in cordless telephone design in order to meet the cost, weight, and talk-time objectives are the following.

- Few users per megahertz
- Few users per base unit (many link together a particular handset and base unit)
- Large number of base units per unit area; one or more base units per wireline access line (in high-rise apartment buildings the density of base units is very large)
- Short transmission range

There is no added network complexity since a base unit looks to a telephone network like a wireline telephone. These issues are also discussed in [29, 50].

Digital cordless telephones in Europe have been evolving for a few years to extend their domain of use beyond the limits of inside residences. Cordless telephone, second generation, (CT-2) has evolved to provide telepoint or phone-point services. Base units are located in places where people congregate. e.g., along city streets and in shopping malls, train stations, etc. Handsets registered with the phone-point provider can place calls when within range of a telepoint. CT-2 does not provide capability for transferring (handing off) active wireless calls from one phone point to another if a user moves out of range of the one to which the

[1] These ISM technologies either use spread spectrum techniques (direct sequence or frequency hopping) or very low-transmitter power ($<\sim$ 1 mW) as required by the ISM band regulations.

Ford Motor Co.
Exhibit 1030
Page 342

call was initiated. A CT-2+ technology, evolved from CT-2 and providing limited handoff capability, is being deployed in Canada. { CT-2+ deployment was never completed. } Phone-point service was introduced in the United Kingdom twice, but failed to attract enough customers to become a viable service. In Singapore and Hong Kong, however, CT-2 phone point has grown rapidly, reaching over 150,000 subscribers in Hong Kong [75] in mid-1994. The reasons for success in some places and failure in others are still being debated, but it is clear that the compactness of the Hong Kong and Singapore populations make the service more widely available, using fewer base stations than in more spreadout cities. Complaints of CT-2 phone-point users in trials have been that the radio coverage was not complete enough, and/or they could not tell whether there was coverage at a particular place, and the lack of handoff was inconvenient. In order to provide the alerting or ringing function for phone-point service, conventional radio pagers have been built into some CT-2 handsets. (The telephone network to which a CT-2 phone point is attached has no way of knowing from which base units to send a ringing message, even though the CT-2 handsets can be rung from a home base unit). { CT-2 phone points in Hong Kong peaked at about 170,000 subscribers. There was then a precipitous decline as these subscribers abandoned the limited service CT-2 in favor of high-tier PCS and cellular services which had reduced their prices to compete with phone points. Phone points have now been removed from Hong Kong. CT-2 phone points were also deployed in Korea, again with initial success followed by decline. The message is clear from the CT-2 and PHS experiences. The attributes of cordless phone-like low-tier PCS are very attractive, but need widespread coverage and vehicle speed handoff in order to be long term attractive and viable in the market. }

Another European evolution of cordless telephones is DECT, which was optimized for use inside buildings. Base units are attached through a controller to private branch exchanges (PBXs), key telephone systems, or phone company CENTREX telephone lines. DECT controllers can hand off active calls from one base unit to another as users move, and can page or ring handsets as a user walks through areas covered by different base units. { DECT has increased in deployment in some countries, notably Italy, for several applications including cordless telephone and wireless loop. There are reported to be perhaps 7 million or more DECT users in Europe. }

These cordless telephone evolutions to more widespread usage outside and inside with telepoints and to usage inside large buildings are illustrated in Fig. 15.1, along with the integration of paging into handsets to provide alerting for phone-point services. They represent the first attempts to increase the service area of mobility for low-power cordless telephones.

Some of the characteristics of the digital cordless telephone technologies, CT-2 and DECT, are listed in Table 15.1. Additional information can be found in [32, 50]. Even though there are significant differences between these technologies, e.g., multiple access technology [frequency division multiple access (FDMA) or time division multiple access (TDMA)/FDMA], and channel bit rate, there are many similarities that are fundamental to the design objectives discussed earlier and to a user's perception of them. These similarities and their implications are as follows.

32-kb/s Adaptive Differential Pulse Code Modulation (ADPCM) Digital Speech Encoding: This is a low-complexity (low-signal processing power) speech encoding process that provides wireline speech quality and is an international standard.

Average Transmitter Power $\leq$ 10 mW: This permits many hours of talk time with small, low-cost, lightweight batteries, but provides limited radio range.

Low-Complexity Radio Signal Processing: There is no forward error correction and no complex multipath mitigation (i.e., no equalization or spread spectrum).

Ford Motor Co.
Exhibit 1030
Page 343

TABLE 15.1 Wireless PCS Technologies

	High-Power Systems				Low-Power Systems			
	Digital Cellular (High-Tier PCS)				Low-Tier PCS		Digital Cordless	
System	IS-54	IS-95 (DS)	GSM	DCS-1800	WACS/PACS	Handi-Phone	DECT	CT-2
Multiple Access	TDMA/ FDMA	CDMA/ FDMA	TDMA/ FDMA	TDMA/ FDMA	TDMA/ FDMA	TDMA/ FDMA	TDMA/ FDMA	FDMA
Freq. band, MHz						1895–1907	1880–1990	864–868
Uplink, MHz	869–894	869–894	935–960	1710–1785	Emerg.			
Downlink, MHz	824–849 (USA)	824–849 (USA)	890–915 (Eur.)	1805–1880 (UK)	Tech.* (USA)	(Japan)	(Eur.)	(Eur. and Asia)
RF ch. spacing						300	1728	100
Downlink, KHz	30	1250	200	200	300			
Uplink, KHz	30	1250	200	200	300			
Modulation	$\pi/4$ DQPSK	BPSK/QPSK	GMSK	GMSK	$\pi/4$ QPSK	$\pi/4$ DQPSK	GFSK	GFSK
Portable txmit Power, max./avg.	600 mW/ 200 mW	600 mW	1 W/ 125 mW	1 W/ 125 mW	200 mW/ 25 mW	80 mW/ 10 mW	250 mW/ 10 mW	10 mW/ 5 mW
Speech coding	VSELP	QCELP	RPE-LTP	RPE-LTP	ADPCM	ADPCM	ADPCM	ADPCM
Speech rate, kb/s	7.95	8 (var.)	13	13	32/16/8	32	32	32
Speech ch./RF ch.	3	—	8	8	8/16/32	4	12	1
Ch. Bit rate, kb/s						384	1152	72
Uplink, kb/s	48.6		270.833	270.833	384			
Downlink, kb/s	48.6		270.833	270.833	384			
Ch. coding	1/2 rate conv.	1/2 rate fwd. 1/3 rate rev.	1/2 rate conv.	1/2 rate conv.	CRC	CRC	CRC (control)	None
Frame, ms	40	20	4.615	4.615	2.5	5	10	2

* Spectrum is 1.85–2.2 GHz allocated by the FCC for emerging technologies; DS is direct sequence.

Ford Motor Co.
Exhibit 1030
Page 344

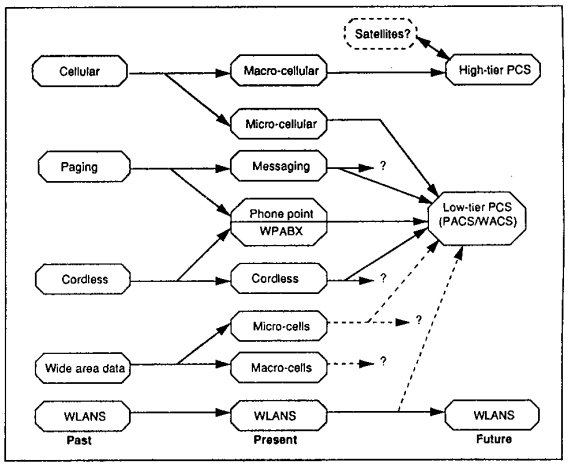


FIGURE 15.1

Low Transmission Delay, e.g., < 50 ms, and for CT-2 < 10-ms Round Trip: This is a speech-quality and network-complexity issue. A maximum of 10 ms should be allowed, taking into account additional inevitable delay in long-distance networks. Echo cancellation is generally required for delays > 10 ms.

Simple Frequency-Shift Modulation and Noncoherent Detection: Although still being low in complexity, the slightly more complex 4QPSK modulation with coherent detection provides significantly more spectrum efficiency, range, and interference immunity.

Dynamic Channel Allocation: Although this technique has potential for improved system capacity, the cordless-telephone implementations do not take full advantage of this feature for handoff and, thus, cannot reap the full benefit for moving users [15, 19].

Time Division Duplex (TDD): This technique permits the use of a single contiguous frequency band and implementation of diversity from one end of a radio link. Unless all base station transmissions are synchronized in time, however, it can incur severe cochannel interference penalties in outside environments [15, 16]. Of course, for cordless telephones used inside with base stations not having a propagation advantage, this is not a problem. Also, for small indoor PBX networks, synchronization of base station transmission is easier than is synchronization throughout a widespread outdoor network, which can have many adjacent base stations connected to different geographic locations for central control and switching.

15.3.2 Cellular Mobile Radio Systems

Cellular mobile radio systems are becoming known in the United States as high-tier personal communications service (PCS), particularly when implemented in the new 1.9-GHz PCS bands [20]. These systems generally can be categorized as providing high-mobility, wide-ranging, two-way tetherless voice communications. In these systems, high mobility refers to vehicular speeds, and also to widespread regional to nationwide coverage [27, 29, 50]. Mobile radio has been evolving for over 50 years. Cellular radio integrates wireless access with large-scale networks having sophisticated intelligence to manage mobility of users.

Cellular radio was designed to provide voice service to wide-ranging vehicles on streets and highways [29, 39, 50, 82], and generally uses transmitter power on the order of 100

times that of cordless telephones ($\approx$ 2 W for cellular). Thus, cellular systems can only provide reduced service to handheld sets that are disadvantaged by using somewhat lower transmitter power (< 0.5 W) and less efficient antennas than vehicular sets. Handheld sets used inside buildings have the further disadvantage of attenuation through walls that is not taken into account in system design.

Cellular radio or high-tier PCS has experienced large growth as noted earlier. In spite of the limitations on usage of handheld sets already noted, handheld cellular sets have become very popular, with their sales becoming comparable to the sales of vehicular sets. Frequent complaints from handheld cellular users are that batteries are too large and heavy, and both talk time and standby time are inadequate. { Cellular and high-tier PCS pocket handsets have continued to decrease in size and weight and more efficient lithium batteries have been incorporated. This has increased their attractiveness (more on this in the later section "Reality Check"). For several years there have been many more pocket handsets sold than vehicular mounted sets every year. However, despite the improvements in these handsets and batteries, the complaints of weight and limited talk time still persist. The electronics have become essentially weightless compared to the batteries required for these high-tier PCS and cellular handsets. }

Cellular radio at 800 MHz has evolved to digital radio technologies [29, 39, 50] in the forms of the deployed systems standards

- Global Standard for Mobile (GSM) in Europe
- Japanese or personal digital cellular (JDC or PDC) in Japan
- U.S. TDMA digital cellular known as USDC or IS-54.

and in the form of the code division multiple access (CDMA) standard, IS-95, which is under development but not yet deployed. { Since the first edition was published, CDMA systems have been deployed in the U.S., Korea, Hong Kong, and other countries after many months (years) of redesign, reprogramming, and adjustment. These CDMA issues are discussed later in the section "New Technology." }

The most significant consideration in the design compromises made for the U.S. digital cellular or high-tier PCS systems was the high cost of cell sites (base stations). A figure often quoted is U.S. $1 million for a cell site. This consideration drove digital system designs to maximize users per megahertz and to maximize the users per cell site.

Because of the need to cover highways running through low-population-density regions between cities, the relatively high transmitter power requirement was retained to provide maximum range from high antenna locations.

Compromises that were accepted while maximizing the two just cited parameters are as follows.

- High transmitter power consumption.
- High user-set complexity, and thus high signal-processing power consumption.
- Low circuit quality.
- High network complexity, e.g., the new IS-95 technology will require complex new switching and control equipment in the network, as well as high-complexity wireless-access technology.

Cellular radio or high-tier PCS has also been evolving for a few years in a different direction, toward very small coverage areas or microcells. This evolution provides increased capacity in areas having high user density, as well as improved coverage of shadowed areas. Some microcell base stations are being installed inside, in conference center lobbies and

Ford Motor Co.
Exhibit 1030
Page 346

similar places of high user concentrations. Of course, microcells, also permit lower transmitter power that conserves battery power when power control is implemented, and base stations inside buildings circumvent the outside wall attenuation. Low-complexity microcell base stations also are considerably less expensive than conventional cell sites, perhaps two orders of magnitude less expensive. Thus, the use of microcell base stations provides large increases in overall system capacity, while also reducing the cost per available radio channel and the battery drain on portable subscriber equipment. This microcell evolution, illustrated in Fig. 15.1, moves handheld cellular sets in a direction similar to that of the expanded-coverage evolution of cordless telephones to phone points and wireless PBX.

Some of the characteristics of digital-cellular or high-tier PCS technologies are listed in Table 15.1 for IS-54, IS-95, and GSM at 900 MHz, and DCS-1800, which is GSM at 1800 MHz. { The technology listed here as IS-54 has also become known as IS-136 having more sophisticated digital control channels. These technologies, IS-54/IS-136 are also sometimes known as DAMPS (i.e., Digital AMPS), as U.S. TDMA or North American TDMA, or sometimes just as "TDMA." } Additional information can be found in [29, 39, 50]. The JDC or PDC technology, not listed, is similar to IS-54. As with the digital cordless technologies, there are significant differences among these cellular technologies, e.g., modulation type, multiple access technology, and channel bit rate. There are also many similarities, however, that are fundamental to the design objectives discussed earlier. These similarities and their implications are as follows.

Low Bit-Rate Speech Coding $\leq$13 kb/s with Some $\leq$8 kb/s: Low bit-rate speech coding obviously increases the number of users per megahertz and per cell site. However, it also significantly reduces speech quality [29], and does not permit speech encodings in tandem while traversing a network; see also the section on Other Issues later in this chapter.

Some Implementations Make Use of Speech Inactivity: This further increases the number of users per cell site, i.e., the cell-site capacity. It also further reduces speech quality [29], however, because of the difficulty of detecting the onset of speech. This problem is even worse in an acoustically noisy environment like an automobile.

High Transmission Delay; $\approx$200-ms Round Trip: This is another important circuit-quality issue. Such large delay is about the same as one-way transmission through a synchronous-orbit communications satellite. A voice circuit with digital cellular technology on both ends will experience the delay of a full satellite circuit. It should be recalled that one reason long-distance circuits have been removed from satellites and put onto fiber-optic cable is because customers find the delay to be objectionable. This delay in digital cellular technology results from both computation for speech bit-rate reduction and from complex signal processing, e.g., bit interleaving, error correction decoding, and multipath mitigation [equalization or spread spectrum code division multiple access (CDMA)].

High-Complexity Signal Processing, Both for Speech Encoding and for Demodulation: Signal processing has been allowed to grow without bound and is about a factor of 10 greater than that used in the low-complexity digital cordless telephones [29]. Since several watts are required from a battery to produce the high transmitter power in a cellular or high-tier PCS set, signal-processing power is not as significant as it is in the low-power cordless telephones; see also the section on Complexity/Coverage Area Comparisons later in this chapter.

Fixed Channel Allocation: The difficulties associated with implementing capacity-increasing dynamic channel allocation to work with handoff [15, 19] have impeded its adoption in systems requiring reliable and frequent handoff.

Frequency Division Duplex (FDD): Cellular systems have already been allocated paired-frequency bands suitable for FDD. Thus, the network or system complexity required

Ford Motor Co.
Exhibit 1030
Page 347

for providing synchronized transmissions [15, 16] from all cell sites for TDD has not been embraced in these digital cellular systems. Note that TDD has not been employed in IS-95 even though such synchronization is required for other reasons.

Mobile/Portable Set Power Control: The benefits of increased capacity from lower overall cochannel interference and reduced battery drain have been sought by incorporating power control in the digital cellular technologies.

15.3.3 Wide-Area Wireless Data Systems

Existing wide area data systems generally can be categorized as providing high mobility, wide-ranging, low-data-rate digital data communications to both vehicles and pedestrians [29, 50]. These systems have not experienced the rapid growth that the two-way voice technologies have, even though they have been deployed in many cities for a few years and have established a base of customers in several countries. Examples of these packet data systems are shown in Table 15.2.

TABLE 15.2 Wide-Area Wireless Packet Data Systems

	CDPD[1]	RAM Mobile (Mobitex)	ARDIS[2] (KDT)	Metricom (MDN)[3]
Data rate, kb/s	19.2	8 (19.2)	4.8 (19.2)	76
Modulation	GMSK BT = 0.5	GMSK	GMSK	GMSK
Frequency, MHz	800	900	800	915
Chan. spacing, kHz	30	12.5	25	160
Status	1994 service	Full service	Full service	In service
Access means	Unused AMPS channels	Slotted Aloha CSMA		FH SS (ISM)
Transmit power, W			40	1

Note: Data in parentheses ( ) indicates proposed.
[1]Cellular Digital Packet Data
[2]Advanced Radio Data Information Service
[3]Microcellular Data Network

The earliest and best known of these systems in the United States are the ARDIS network developed and run by Motorola, and the RAM mobile data network based on Ericsson Mobitex Technology. These technologies were designed to make use of standard, two-way voice, land mobile-radio channels, with 12.5- or 25-kHz channel spacing. In the United States these are specialized mobile radio services (SMRS) allocations around 450 MHz and 900 MHz. Initially, the data rates were low: 4.8 kb/s for ARDIS and 8 kb/s for RAM. The systems use high transmitter power (several tens of watts) to cover large regions from a few base stations having high antennas. The relatively low data capacity of a relatively expensive base station has resulted in economics that have not favored rapid growth.

The wide-area mobile data systems also are evolving in several different directions in an attempt to improve base station capacity, economics, and the attractiveness of the service. The technologies used in both the ARDIS and RAM networks are evolving to higher channel bit rates of 19.2 kb/s.

The cellular carriers and several manufacturers in the United States are developing and deploying a new wide area packet data network as an overlay to the cellular radio networks. This cellular digital packet data (CDPD) technology shares the 30-kHz spaced 800-MHz

Ford Motor Co.
Exhibit 1030
Page 348

voice channels used by the analog FM advanced mobile phone service (AMPS) systems. Data rate is 19.2 kb/s. The CDPD base station equipment also shares cell sites with the voice cellular radio system. The aim is to reduce the cost of providing packet data service by sharing the costs of base stations with the better established and higher cell-site capacity cellular systems. This is a strategy similar to that used by nationwide fixed wireline packet data networks that could not provide an economically viable data service if they did not share costs by leasing a small amount of the capacity of the interexchange networks that are paid for largely by voice traffic. { CDPD has been deployed in many U.S. cities for several years. However, it has not lived up to early expectations and has become "just another" wireless data service with some subscribers, but not with the large growth envisioned earlier. }

Another evolutionary path in wide-area wireless packet data networks is toward smaller coverage areas or microcells. This evolutionary path also is indicated on Fig. 15.1. The microcell data networks are aimed at stationary or low-speed users. The design compromises are aimed at reducing service costs by making very small and inexpensive base stations that can be attached to utility poles, the sides of buildings and inside buildings and can be widely distributed throughout a region. Base-station-to-base-station wireless links are used to reduce the cost of the interconnecting data network. In one network this decreases the overall capacity to serve users, since it uses the same radio channels that are used to provide service. Capacity is expected to be made up by increasing the number of base stations that have connections to a fixed-distribution network as service demand increases. Another such network uses other dedicated radio channels to interconnect base stations. In the high-capacity limit, these networks will look more like a conventional cellular network architecture, with closely spaced, small, inexpensive base stations, i.e., microcells, connected to a fixed infrastructure. Specialized wireless data networks have been built to provide metering and control of electric power distributions, e.g., Celldata and Metricom in California.

A large microcell network of small inexpensive base stations has been installed in the lower San Francisco Bay Area by Metricom, and public packet-data service was offered during early 1994. Most of the small (shoe-box size) base stations are mounted on street light poles. Reliable data rates are about 75 kb/s. The technology is based on slow frequency-hopped spread spectrum in the 902–928 MHz U.S. ISM band. Transmitter power is 1 W maximum, and power control is used to minimize interference and maximize battery life time. { The metricom network has been improved and significantly expanded in the San Francisco Bay Area and has been deployed in Washington, D.C. and a few other places in the U.S. However, like all wireless data services so far, it has failed to grow as rapidly or to attract as many subscribers as was originally expected. Wireless data overall has had only very limited success compared to that of the more voice-oriented technologies, systems, and services. }

15.3.4 High-Speed Wireless Local-Area Networks (WLANs)

Wireless local-area data networks can be categorized as providing low-mobility high-data-rate data communications within a confined region, e.g., a campus or a large building. Coverage range from a wireless data terminal is short, tens to hundreds of feet, like cordless telephones. Coverage is limited to within a room or to several rooms in a building. WLANs have been evolving for a few years, but overall the situation is chaotic, with many different products being offered by many different vendors [29, 59]. There is no stable definition of the needs or design objectives for WLANs, with data rates ranging from hundreds of kb/s to more than 10 Mb/s, and with several products providing one or two Mb/s wireless

Ford Motor Co.
Exhibit 1030
Page 349

link rates. The best description of the WLAN evolutionary process is: having severe birth pains. An IEEE standards committee, 802.11, has been attempting to put some order into this topic, but their success has been somewhat limited. A partial list of some advertised products is given in Table 15.3. Users of WLANs are not nearly as numerous as the users of more voice-oriented wireless systems. Part of the difficulty stems from these systems being driven by the computer industry that views the wireless system as just another plug-in interface card, without giving sufficient consideration to the vagaries and needs of a reliable radio system. { This section still describes the WLAN situation in spite of some attempts at standards in the U.S. and Europe, and continuing industry efforts. Some of the products in Table 15.3 have been discontinued because of lack of market and some new products have been offered, but the manufacturers still continue to struggle to find enough customers to support their efforts. Optimism remains high in the WLAN community that "eventually" they will find the "right" technology, service, or application to make WLANs "take off" — but the world still waits. Success is still quite limited. }

There are two overall network architectures pursued by WLAN designers. One is a centrally coordinated and controlled network that resembles other wireless systems. There are base stations in these networks that exercise overall control over channel access [44].

The other type of network architecture is the self-organizing and distributed controlled network where every terminal has the same function as every other terminal, and networks are formed ad hoc by communications exchanges among terminals. Such ad hoc networks are more like citizen band (CB) radio networks, with similar expected limitations if they were ever to become very widespread.

Nearly all WLANs in the United States have attempted to use one of the ISM frequency bands for unlicensed operation under part 15 of the FCC rules. These bands are 902–928 MHz, 2400–2483.5 MHz, and 5725–5850 MHz, and they require users to accept interference from any interfering source that may also be using the frequency. The use of ISM bands has further handicapped WLAN development because of the requirement for use of either frequency hopping or direct sequence spread spectrum as an access technology, if transmitter power is to be adequate to cover more than a few feet. One exception to the ISM band implementations is the Motorola ALTAIR, which operates in a licensed band at 18 GHz. { It appears that ALTAIR has been discontinued because of the limited market. } The technical and economic challenges of operation at 18 GHz have hampered the adoption of this 10–15 Mb/s technology. The frequency-spectrum constraints have been improved in the United States with the recent FCC allocation of spectrum from 1910–1930 MHz for unlicensed data PCS applications. Use of this new spectrum requires implementation of an access etiquette incorporating listen before transmit in an attempt to provide some coordination of an otherwise potentially chaotic, uncontrolled environment [68]. Also, since spread spectrum is not a requirement, access technologies and multipath mitigation techniques more compatible with the needs of packet-data transmission [59], e.g., multipath equalization or multicarrier transmission can be incorporated into new WLAN designs. { The FCC is allocating spectrum at 5 GHz for wideband wireless data for internet and next generation data network access, BUT it remains to be seen whether this initiative is any more successful than past wireless data attempts. Optimism is again high, BUT... }

Three other widely different WLAN activities also need mentioning. One is a large European Telecommunications Standards Institute (ETSI) activity to produce a standard for high performance radio local area network (HIPERLAN), a 20-Mb/s WLAN technology to operate near 5 GHz. Other activities are large U.S. Advance Research Projects Agency-(ARPA-) sponsored, WLAN research projects at the Universities of California at Berkeley (UCB), and at Los Angeles (UCLA). The UCB Infopad project is based on a coordinated network architecture with fixed coordinating nodes and direct-sequence spread spectrum

Ford Motor Co.
Exhibit 1030
Page 350

TABLE 15.3 Partial List of WLAN Products

Product Company Location	Freq., MHz	Link Rate, Mb/s	User Rate	Protocol(s)	Access	No. of chan. or Spread Factor	Mod./ Coding	Power, mW	Network Topology
Altair Plus Motorola Arlington Hts, IL	18–19 GHz	15	5.7 Mb/s	Ethernet			4-level FSK	25 peak	Eight devices/ radio; radio to base to ethernet
WaveLAN NCR/AT&T Dayton, OH	902–928	2	1.6 Mb/s	Ethernet-like	DS SS		DQPSK	250	Peer-to-peer
AirLan Solectek San Diego, CA	902–928		2 Mb/s	Ethernet	DS SS		DQPSK	250	PCMCIA w/ant.; radio to hub
Freeport Windata Inc. Northboro, MA	902–928	16	5.7 Mb/s	Ethernet	DS SS	32 chips/bit	16 PSK trellis coding	650	Hub
Intersect Persoft Inc. Madison, WI	902–928		2 Mb/s	Ethernet token ring	DS SS		DQPSK	250	Hub
LAWN O'Neill Comm. Horsham, PA	902–928		38.4 kb/s	AX.25	SS	20 users/chan.; max. 4 chan.		20	Peer-to-peer
WILAN Wi-LAN Inc. Calgary, Alberta	902–928	20	1.5 Mb/s/ chan.	Ethernet, token ring	CDMA/ TDMA	3 chan. 10–15 links each	unconven-tional	30	Peer-to-peer
RadioPort ALPS Electric USA	902–928		242 kb/s	Ethernet	SS	7/3 channels		100	Peer-to-peer
ArLAN 600 Telesys. SLW Don Mills, Ont.	902–928; 2.4 GHz		1.35 Mb/s	Ethernet	SS			1 W max	PCs with ant.; radio to hub
Radio Link Cal. Microwave Sunnyvale, CA	902–928; 2.4 GHz	250 kb/s	64 kb/s		FH SS	250 ms/hop 500 kHz space			Hub
Range LAN Proxim, Inc. Mountain View, CA	902–928		242 kb/s	Ethernet, token ring	DS SS	3 chan.		100	
RangeLAN 2 Proxim, Inc. Mountain View, CA	2.4 GHz	1.6	50 kb/s max.	Ethernet, token ring	FH SS	10 chan. at 5 kb/s; 15 sub-ch. each		100	Peer-to-peer bridge
Netwave Xircom Calabasas, CA	2.4 GHz	1/adaptor		Ethernet, token ring	FH SS	82 1-MHz chn. or "hops"			Hub
Freelink Cabletron Sys. Rochester, NH	2.4 and 5.8 GHz		5.7 Mb/s	Ethernet	DS SS	32 chips/bit	16 PSK trellis coding	100	Hub

Ford Motor Co.
Exhibit 1030
Page 351

(CDMA), whereas, the UCLA project is aimed at peer-to-peer networks and uses frequency hopping. Both ARPA sponsored projects are concentrated on the 900-MHz ISM band.

As computers shrink in size from desktop to laptop to palmtop, mobility in data network access is becoming more important to the user. This fact, coupled with the availability of more usable frequency spectrum, and perhaps some progress on standards, may speed the evolution and adoption of wireless mobile access to WLANs. From the large number of companies making products, it is obvious that many believe in the future of this market. { It should be noted that the objective for 10 MB/s data service with widespread coverage from a sparse distribution of widely separated base stations equivalent to cellular is unrealistic and unrealizable. This can be readily seen by considering a simple example. Consider a cellular coverage area that requires full cellular power of 0.5 watt to cover from a handset. Consider the handset to use a typical digital cellular bit rate of about 10 kb/s (perhaps 8 kb/s speech coding + overhead). With all else in the system the same, e.g., antennas, antenna height, receiver noise figure, detection sensitivity, etc., the 10 MB/s data would require 10 MB/s $\div$ 10 kb/s = 1000 times as much power as the 10 kb/s cellular. Thus, it would require 0.5 $\times$ 1000 = 500 watts for the wireless data transmitter. This is a totally unrealistic situation. If the data system operates at a higher frequency (e.g., 5 GHz) than the cellular system (e.g., 1 or 2 GHz) then there will be even more power required to overcome the additional loss at a higher frequency. The sometimes expressed desire by the wireless data community for a system to provide network access to users in and around buildings and to provide 10 MB/s over 10 miles with 10 milliwatts of transmitter power and costing $10.00 is totally impossible. It requires violation of the "laws of physics." }

15.3.5 Paging/Messaging Systems

Radio paging began many years ago as a one-bit messaging system. The one bit was: some one wants to communicate with you. More generally, paging can be categorized as one-way messaging over wide areas. The one-way radio link is optimized to take advantage of the asymmetry. High transmitter power (hundreds of watts to kilowatts), and high antennas at the fixed base stations permit low-complexity, very low-power-consumption, pocket paging receivers that provide long usage time from small batteries. This combination provides the large radio-link margins needed to penetrate walls of buildings without burdening the user set battery. Paging has experienced steady rapid growth for many years and serves about 15 million subscribers in the United States.

Paging also has evolved in several different directions. It has changed from analog tone coding for user identification to digitally encoded messages. It has evolved from the 1-b message, someone wants you, to multibit messages from, first, the calling party's telephone number to, now, short e-mail text messages. This evolution is noted in Fig. 15.1.

The region over which a page is transmitted has also increased from 1) local, around one transmitting antenna; to 2) regional, from multiple widely-dispersed antennas; to 3) nationwide, from large networks of interconnected paging transmitters. The integration of paging with CT-2 user sets for phone-point call alerting was noted previously.

Another evolutionary paging route sometimes proposed is two-way paging. This is an ambiguous and unrealizable concept, however, since the requirement for two-way communications destroys the asymmetrical link advantage so well exploited by paging. Two-way paging puts a transmitter in the user's set and brings along with it all of the design compromises that must be faced in such a two-way radio system. Thus, the word paging is not appropriate to describe a system that provides two-way communications. { The two-way paging situation is as unrealistic as that noted earlier for wide-area, high-speed, low-power wireless data. This can be seen by looking at the asymmetry situation in paging. In or-

Ford Motor Co.
Exhibit 1030
Page 352

der to achieve comparable coverage uplink and downlink, a 500-watt paging transmitter downlink advantage must be overcome in the uplink. Even considering the relatively high cellular handset transmit power levels on the order of 0.5 watt results in a factor of 1000 disadvantage, and 0.5 watt is completely incompatible with the low power paging receiver power requirements. If the same uplink and downlink coverage is required for an equivalent set of system parameters, then the only variable left to work with is bandwidth. If the paging link bit rate is taken to be 10 kb/sec (much higher than many paging systems), then the usable uplink rate is 10 kb/s/1000 = 10 B/s, an unusably low rate. Of course, some uplink benefit can be gained because of better base station receiver noise figure and by using forward error correction and perhaps ARQ. However, this is unlikely to raise the allowable rate to greater than 100 B/s which even though likely overoptimistic is still unrealistically low and we have assumed an unrealistically high transmit power in the two-way "pager!" }

15.3.6 Satellite-Based Mobile Systems

Satellite-based mobile systems are the epitome of wide-area coverage, expensive base station systems. They generally can be categorized as providing two-way (or one-way) limited quality voice and/or very limited data or messaging to very wide-ranging vehicles (or fixed locations). These systems can provide very widespread, often global, coverage, e.g., to ships at sea by INMARSAT. There are a few messaging systems in operation, e.g., to trucks on highways in the United States by Qualcomm's Omnitracs system.

A few large-scale mobile satellite systems have been proposed and are being pursued: perhaps the best known is Motorola's Iridium; others include Odyssey, Globalstar, and Teledesic. The strength of satellite systems is their ability to provide large regional or global coverage to users outside buildings. However, it is very difficult to provide adequate link margin to cover inside buildings, or even to cover locations shadowed by buildings, trees, or mountains. A satellite system's weakness is also its large coverage area. It is very difficult to provide from Earth orbit the small coverage cells that are necessary for providing high overall systems capacity from frequency reuse. This fact, coupled with the high cost of the orbital base stations, results in low capacity along with the wide overall coverage but also in expensive service. Thus, satellite systems are not likely to compete favorably with terrestrial systems in populated areas or even along well-traveled highways. They can complement terrestrial cellular or PCS systems in low-population-density areas. It remains to be seen whether there will be enough users with enough money in low-population-density regions of the world to make satellite mobile systems economically viable. { Some of the mobile satellite systems have been withdrawn, e.g., Odyssey. Some satellites in the Iridium and Globalstar systems have been launched. The industry will soon find out whether these systems are economically viable. }

Proposed satellite systems range from 1) low-Earth-orbit systems (LEOS) having tens to hundreds of satellites through 2) intermediate- or medium-height systems (MEOS) to 3) geostationary or geosynchronous orbit systems (GEOS) having fewer than ten satellites. LEOS require more, but less expensive, satellites to cover the Earth, but they can more easily produce smaller coverage areas and, thus, provide higher capacity within a given spectrum allocation. Also, their transmission delay is significantly less (perhaps two orders of magnitude!), providing higher quality voice links, as discussed previously. On the other hand, GEOS require only a few, somewhat more expensive, satellites (perhaps only three) and are likely to provide lower capacity within a given spectrum allocation and suffer severe transmission-delay impairment on the order of 0.5 s. Of course, MEOS fall in between these extremes. The possible evolution of satellite systems to complement high-tier PCS is indicated in Fig. 15.1.

Ford Motor Co.
Exhibit 1030
Page 353

15.3.7 {Fixed Point-to-Multipoint Wireless Loops

Wideband point-to-multipoint wireless loop technologies sometimes have been referred to earlier as "wireless cable" when they were proposed as an approach for providing interactive video services to homes [30]. However, as the video application started to appear less attractive, the application emphasis shifted to providing wideband data access for the internet, the worldwide web, and future wideband data networks. Potentially lower costs are the motivation for this wireless application. As such, these technologies will have to compete with existing coaxial cable and fiber/coax distribution by CATV companies, with satellites, and with fiber and fiber/coax systems being installed or proposed by telephone companies and other entities [30]. Another competitor is asymmetric digital subscriber line technology, which uses advanced digital signal processing to provide high-bandwidth digital distribution over twisted copper wire pairs.

In the U.S. two widely different frequency bands are being pursued for fixed point-to-multipoint wireless loops. These bands are at 28 GHz for local multipoint distribution systems or services (LMDS) [52] and 2.5 to 2.7 GHz for microwave or metropolitan distribution systems (MMDS) [74]. The goal of low-cost fixed wireless loops is based on the low cost of point-to-multipoint line-of-sight wireless technology. However, significant challenges are presented by the inevitable blockage by trees, terrain, and houses, and by buildings in heavily built-up residential areas. Attenuation in rainstorms presents an additional problem at 28 GHz in some localities. Even at the 2.5-GHz MMDS frequencies, the large bandwidth required for distribution of many video channels presents a challenge to provide adequate radio-link margin over obstructed paths. From mobile satellite investigations it is known that trees can often produce over 15 dB additional path attenuation [38]. Studies of blockage by buildings in cities have shown that it is difficult to have line-of-sight access to more than 60% of the buildings from a single base station [55]. Measurements in a region in Brooklyn, NY [60], suggest that access from a single base station can range from 25% to 85% for subscriber antenna heights of 10 to 35 ft and a base station height of about 290 ft. While less blockage by houses could be expected in residential areas, such numbers would suggest that greater than 90% access to houses could be difficult, even from multiple elevated locations, when mixes of one- and two-story houses, trees, and hills are present. In regions where tree cover is heavy, e.g., the eastern half of the U.S., tree cover in many places will present a significant obstacle. Heavy rainfall is an additional problem at 28 GHz in some regions. In spite of these challenges, the lure of low-cost wireless loops is attracting many participants, both service providers and equipment manufacturers. }

15.3.8 Reality Check

Before we go on to consider other applications and compromises, perhaps it would be helpful to see if there is any indication that the previous discussion is valid. For this check, we could look at cordless telephones for telepoint use (i.e., pocketphones) and at pocket cellular telephones that existed in the 1993 time frame.

Two products from one United States manufacturer are good for this comparison. One is a third-generation hand-portable analog FM cellular phone from this manufacturer that represents their second generation of pocketphones. The other is a first-generation digital cordless phone built to the United Kingdom CT-2 common air interface (CAI) standard. Both units are of flip phone type with the earpiece on the main handset body and the mouthpiece formed by or on the flip-down part. Both operate near 900 MHz and have 1/4 wavelength pull-out antennas. Both are fully functional within their class of operation (i.e., full number of U.S. cellular channels, full number of CT-2 channels, automatic channel

Ford Motor Co.
Exhibit 1030
Page 354

setup, etc.) Table 15.4 compares characteristics of these two wireless access pocketphones from the same manufacturer.

TABLE 15.4 Comparison of CT-2 and Cellular Pocket Size Flip-Phones from the Same Manufacturer

Characteristics/Parameter	CT-2	Cellular
Weight, oz		
Flip phone only	5.2	4.2
Battery[1] only	1.9	3.6
Total unit	7.1	7.8
Size (max.dimensions), in		
Flip phone only	5.9 × 2.2 × 0.95 8.5 in^3	5.5 × 2.4 × 0.9 —
Battery[1] only	1.9 × 1.3 × 0.5 internal	4.7 × 2.3 × 0.4 external
Total unit	5.9 × 2.2 × 0.95 8.5 in^3	5.9 × 2.4 × 1.1 11.6 in^3
Talk-time, min (h)		
Rechargeable battery[2]	180 (3)	45
Nonrechargeable battery	600 (10)	N/A
Standby time, h		
Rechargeable battery	30	8
Nonrechargeable battery	100	N/A
Speech quality	32 kb/s telephone quality	30 kHz FM depends on channel quality
Transmit power avg., W	0.005	0.5

[1] Rechargeable battery.
[2] Ni-cad battery.

The following are the most important items to note in the Table 15.4 comparison.

1. The talk time of the low-power pocketphone is four times that of the high-power pocket-phone.
2. The battery inside the low-power pocketphone is about one-half the weight and size of the battery attached to the high-power pocketphone.
3. The battery-usage ratio, talk time/weight of battery, is eight times greater, almost an order of magnitude, for the low-power pocketphone compared to the high-power pocketphone!
4. Additionally, the lower power (5 mW) digital cordless pocketphone is slightly smaller and lighter than the high-power (500 mW) analog FM cellular mobile pocketphone.

{ Similar comparisons can be made between PHS advanced cordless/low-tier PCS phones and advanced cellular/high-tier PCS pocketphones. New lithium batteries have permitted increased talk time in pocketphones. Digital control/paging channels facilitate significantly extended standby time. Advances in solid-state circuits have reduced the size and weight of cellular pocketphone electronics so that they are almost insignificant compared to the battery required for the high power transmitter and complex digital signal processing. However, even with all these changes, there is still a very significant weight and talk time benefit in the low complexity PHS handsets compared to the most advanced cellular/high-tier PCS handsets. Picking typical minimum size and weight handsets for both technologies results in the following comparisons.

	PHS	Cellular
weight, oz		
total unit	3	4.5
size		
total unit	4.2 in^3	—
talk-time, h	8	3
standby time, h	600	48

From the table, the battery usage ratio has been reduced to a factor of about 4 from a factor of 8, but this is based on total weight, not battery weight alone as used for the earlier CT-2 and cellular comparison.

Thus, there is still a large talk time and weight benefit for low-power low-complexity low-tier PCS compared to higher power high-complexity, high tier PCS and cellular. }

The following should also be noted.

1. The room for technology improvement of the CT-2 cordless phone is greater since it is first generation and the cellular phone is second/third generation.
2. A digital cellular phone built to the IS-54, GSM, or JDC standard, or in the proposed United States CDMA technology, would either have less talk time or be heavier and larger than the analog FM phone, because: a) the low-bit-rate digital speech coder is more complex and will consume more power than the analog speech processing circuits; b) the digital units have complex digital signal-processing circuits for forward error correction—either for delay dispersion equalizing or for spread-spectrum processing—that will consume significant amounts of power and that have no equivalents in the analog FM unit; and c) power amplifiers for the shaped-pulse nonconstant-envelope digital signals will be less efficient than the amplifiers for constant-envelope analog FM. Although it may be suggested that transmitter power control will reduce the weight and size of a CDMA handset and battery, if that handset is to be capable of operating at full power in fringe areas, it will have to have capabilities similar to other cellular sets. Similar power control applied to a CT-2-like low-maximum-power set would also reduce its power consumption and thus also its weight and size.

The major difference in size, weight, and talk time between the two pocketphones is directly attributable to the two orders of magnitude difference in average transmitter power. The generation of transmitter power dominates power consumption in the analog cellular phone. Power consumption in the digital CT-2 phone is more evenly divided between transmitter-power generation and digital signal processing. Therefore, power consumption in complex digital signal processing would have more impact on talk time in small low-power personal communicators than in cellular handsets where the transmitter-power generation is so large. Other than reducing power consumption for both functions, the only alternative for increasing talk time and reducing battery weight is to invent new battery technology having greater density; see section on Other Issues later in this chapter.

In contrast, lowering the transmitter power requirement, modestly applying digital signal processing, and shifting some of the radio coverage burden to a higher density of small, low-power, low-complexity, low-cost fixed radio ports has the effect of shifting some of the talk time, weight, and cost constraints from battery technology to solid state electronics technology, which continues to experience orders-of-magnitude improvements in the span of several years. Digital signal-processing complexity, however, cannot be permitted to overwhelm power consumption in low-power handsets; whereas small differences in complexity

Ford Motor Co.
Exhibit 1030
Page 356

will not matter much, orders-of-magnitude differences in complexity will continue to be significant.

Thus, it can be seen from Table 15.4 that the size, weight, and quality arguments in the preceding sections generally hold for these examples. It also is evident from the preceding paragraphs that they will be even more notable when comparing digital cordless pocketphones with digital cellular pocketphones of the same development generations.

15.4 Evolution Toward the Future and to Low-Tier Personal Communications Services

After looking at the evolution of several wireless technologies and systems in the preceding sections it appears appropriate to ask again: wireless personal communications, What is it? All of the technologies in the preceding sections claim to provide wireless personal communications, and all do to some extent. All have significant limitations, however, and all are evolving in attempts to overcome the limitations. It seems appropriate to ask, what are the likely endpoints? Perhaps some hint of the endpoints can be found by exploring what users see as limitations of existing technologies and systems and by looking at the evolutionary trends.

In order to do so, we summarize some important clues from the preceding sections and project them, along with some U.S. standards activity, toward the future.

Digital Cordless Telephones

- Strengths: good circuit quality; long talk time; small lightweight battery; low-cost sets and service.
- Limitations: limited range; limited usage regions.
- Evolutionary trends: phone points in public places; wireless PBX in business.
- Remaining limitations and issues: limited usage regions and coverage holes; limited or no handoff; limited range.

{ Experience with PHS and CT-2 phone point have provided more emphasis on the need for vehicle speed handoff and continuous widespread coverage of populated areas and of highways in between. }

Digital Cellular Pocket Handsets

- Strength: widespread service availability.
- Limitations: limited talk time; large heavy batteries; high-cost sets and service; marginal circuit quality; holes in coverage and poor in-building coverage; limited data capabilities; complex technologies.
- Evolutionary trends: microcells to increase capacity and in-building coverage and to reduce battery drain; satellite systems to extend coverage.
- Remaining limitations and issues: limited talk time and large battery; marginal circuit quality; complex technologies.

Wide Area Data

- Strength: digital messages.
- Limitations: no voice, limited data rate; high cost.
- Evolutionary trends: microcells to increase capacity and reduce cost; share facilities with voice systems to reduce cost.

Ford Motor Co.
Exhibit 1030
Page 357

- Remaining limitations and issues: no voice; limited capacity.

Wireless Local Area Networks (WLANs)

- Strength: high data rate.
- Limitations: insufficient capacity for voice, limited coverage; no standards; chaos.
- Evolutionary trends: hard to discern from all of the churning.

Paging/messaging

- Strengths: widespread coverage; long battery life; small lightweight sets and batteries; economical.
- Limitations: one-way message only; limited capacity.
- Evolutionary desire: two-way messaging and/or voice; capacity.
- Limitations and issues: two-way link cannot exploit the advantages of one-way link asymmetry.

{**Fixed Wireless Loops**

- Strength: High data rates.
- Limitations: no mobility. }

There is a strong trajectory evident in these systems and technologies aimed at providing the following features.

High Quality Voice and Data

- To small, lightweight, pocket carried communicators.
- Having small lightweight batteries.
- Having long talk time and long standby battery life.
- Providing service over large coverage regions.
- For pedestrians in populated areas (but not requiring high population density).
- Including low to moderate speed mobility with handoff. { It has become evident from the experience with PHS and CT-2 phone point that vehicle speed handoff is essential so that handsets can be used in vehicles also. }

Economical Service

- Low subscriber-set cost.
- Low network-service cost.

Privacy and Security of Communications

- Encrypted radio links.

This trajectory is evident in all of the evolving technologies but can only be partially satisfied by any of the existing and evolving systems and technologies! Trajectories from all of the evolving technologies and systems are illustrated in Fig. 15.1 as being aimed at low-tier personal communications systems or services, i.e., low-tier PCS. Taking characteristics from cordless, cellular, wide-area data and, at least moderate-rate, WLANs, suggests the following attributes for this low-tier PCS.

Ford Motor Co.
Exhibit 1030
Page 358

1. 32 kb/s ADPCM speech encoding in the near future to take advantage of the low complexity and low power consumption, and to provide low-delay high-quality speech.
2. Flexible radio link architecture that will support multiple data rates from several kilobits per second. This is needed to permit evolution in the future to lower bit rate speech as technology improvements permit high quality without excessive power consumption or transmission delay and to provide multiple data rates for data transmission and messaging.
3. Low transmitter power ($\leq$ 25 mW average) with adaptive power control to maximize talk time and data transmission time. This incurs short radio range that requires many base stations to cover a large region. Thus, base stations must be small and inexpensive, like cordless telephone phone points or the Metricom wireless data base stations. { The lower power will require somewhat closer spacing of base stations in cluttered environments with many buildings, etc. This issue is dealt with in more detail in Section 15.5. The issues associated with greater base station spacing along highways are also considered in Section 15.5. }
4. Low-complexity signal processing to minimize power consumption. Complexity one-tenth that of digital cellular or high-tier PCS technologies is required [29]. With only several tens of milliwatts (or less under power control) required for transmitter power, signal processing power becomes significant.
5. Low cochannel interference and high coverage area design criteria. In order to provide high-quality service over a large region, at least 99% of any covered area must receive good or better coverage and be below acceptable cochannel interference limits. This implies less than 1% of a region will receive marginal service. This is an order-of-magnitude higher service requirement than the 10% of a region permitted to receive marginal service in vehicular cellular system (high-tier PCS) design criteria.
6. Four-level phase modulation with coherent detection to maximize radio link performance and capacity with low complexity.
7. Frequency division duplexing to relax the requirement for synchronizing base station transmissions over a large region. { PHS uses time division duplexing and requires base station synchronization. In first deployments, one provider did not implement this synchronization. The expected serious performance degradation prompted system upgrades to provide the needed synchronization. While this is not a big issue, it does add complexity to the system and decreases the overall robustness. }
8. { As noted previously, experience with PHS and CT-2 phone point have emphasized the need for vehicular speed handoff in these low-tier PCS systems. Such handoff is readily implemented in PACS and has been demonstrated in the field [51]. This issue is discussed in more detail later in this section. }

Such technologies and systems have been designed, prototyped, and laboratory and field tested and evaluated for several years [7, 23, 24, 25, 26, 27, 28, 29, 31, 32, 50]. The viewpoint expressed here is consistent with the progress in the Joint Technical Committee (JTC) of the U.S. standards bodies, Telecommunications Industry Association (TIA) and Committee T1 of the Alliance for Telecommunications Industry Solutions (ATIS). Many technologies and systems were submitted to the JTC for consideration for wireless PCS in the new 1.9-GHz frequency bands for use in the United States [20]. Essentially all of the technologies and systems listed in Table 15.1, and some others, were submitted in late 1993. It was evident

Ford Motor Co.
Exhibit 1030
Page 359

that there were at least two and perhaps three distinctly different classes of submissions. No systems optimized for packet data were submitted, but some of the technologies are optimized for voice.

One class of submissions was the group labeled high-power systems, digital cellular (high-tier PCS) in Table 15.1. These are the technologies discussed previously in this chapter. They are highly optimized for low-bit-rate voice and, therefore, have somewhat limited capability for serving packet-data applications. Since it is clear that wireless services to wide ranging high-speed mobiles will continue to be needed, and that the technology already described for low-tier PCS may not be optimum for such services, Fig. 15.1 shows a continuing evolution and need in the future for high-tier PCS systems that are the equivalent of today's cellular radio. There are more than 100 million vehicles in the United States alone. In the future, most, if not all, of these will be equipped with high-tier cellular mobile phones. Therefore, there will be a continuing and rapidly expanding market for high-tier systems.

Another class of submissions to the JTC [20] included the Japanese personal handyphone system (PHS) and a technology and system originally developed at Bellcore but carried forward to prototypes and submitted to the JTC by Motorola and Hughes Network Systems. This system was known as wireless access communications systems (WACS).[2] These two submissions were so similar in their design objectives and system characteristics that, with the agreement of the delegations from Japan and the United States, the PHS and WACS submissions were combined under a new name, personal access communication systems (PACS), that was to incorporate the best features of both. This advanced, low-power wireless access system, PACS, was to be known as low-tier PCS. Both WACS/PACS and Handyphone (PHS) are shown in Table 15.1 as low-tier PCS and represent the evolution to low-tier PCS in Fig. 15.1. The WACS/PACS/ UDPC system and technology are discussed in [7, 23, 24, 25, 26, 28, 29, 31, 32, 50].

In the JTC, submissions for PCS of DECT and CT-2 and their variations were also lumped under the class of low-tier PCS, even though these advanced digital cordless telephone technologies were somewhat more limited in their ability to serve all of the low-tier PCS needs. They are included under digital cordless technologies in Table 15.1. Other technologies and systems were also submitted to the JTC for high-tier and low-tier applications, but they have not received widespread industry support.

One wireless access application discussed earlier that is not addressed by either high-tier or low-tier PCS is the high-speed WLAN application. Specialized high-speed WLANs also are likely to find a place in the future. Therefore, their evolution is also continued in Fig. 15.1. The figure also recognizes that widespread low-tier PCS can support data at several hundred kilobits per second and, thus, can satisfy many of the needs of WLAN users.

It is not clear what the future roles are for paging/messaging, cordless telephone appliances, or wide-area packet-data networks in an environment with widespread contiguous coverage by low-tier and high-tier PCS. Thus, their extensions into the future are indicated with a question mark in Fig. 15.1.

Those who may object to the separation of wireless PCS into high-tier and low-tier should review this section again, and note that we have two tiers of PCS now. On the voice side there is cellular radio, i.e., high-tier PCS, and cordless telephone, i.e., an early form of low-tier PCS. On the data side there is wide-area data, i.e., high-tier data PCS, and WLANs,

[2]WACS was known previously as Universal Digital Portable Communications (UDPC).

Ford Motor Co.
Exhibit 1030
Page 360

i.e., perhaps a form of low-tier data PCS. In their evolutions, these all have the trajectories discussed and shown in Fig. 15.1 that point surely toward low-tier PCS. It is this low-tier PCS that marketing studies continue to project is wanted by more than half of the U.S. households or by half of the people, a potential market of over 100 million subscribers in the United States alone. Similar projections have been made worldwide.

{ PACS technology [6] has been prototyped by several manufacturers. In 1995 field demonstrations were run in Boulder, CO at a U.S. West test site using radio ports (base stations) and radio port control units made by NEC. "Handset" prototypes made by Motorola and Panasonic were trialed. The handsets and ports were brought together for the first time in Boulder. The highly successful trial demonstrated the ease of integrating the subsystems of the low-complexity PACS technology and the overall advantages of PACS from a user's perspective as noted throughout this chapter. Effective vehicular speed operation was demonstrated in these tests. Also, Hughes Network Systems (HNS) has developed and tested many sets of PACS infrastructure technology with different handsets in several settings and has many times demonstrated highly reliable high vehicular speed (in excess of 70 mi/hr) operation and handoff among several radio ports. Motorola also has demonstrated PACS equipment in several settings at vehicular speeds as well as for wireless loop applications. Highly successful demonstrations of PACS prototypes have been conducted from Alaska to Florida, from New York to California, and in China and elsewhere.

A PACS deployment in China using NEC equipment started to provide service in 1998. The U.S. Service Provider, 21st Century Telesis, is poised to begin a PACS deployment in several states in the U.S. using infrastructure equipment from HNS and handsets and switching equipment from different suppliers. Perhaps, with a little more support of these deployments, the public will finally be able to obtain the benefits of low-tier PCS. }

15.5 Comparisons with Other Technologies

15.5.1 Complexity/Coverage Area Comparisons

Experimental research prototypes of radio ports and subscriber sets [64, 66] have been constructed to demonstrate the technical feasibility of the radio link requirements in [7]. These WACS prototypes generally have the characteristics and parameters previously noted, with the exceptions that 1) the portable transmitter power is lower (10 mW average, 100 mW peak), 2) dynamic power control and automatic time slot transfer are not implemented, and 3) a rudimentary automatic link-transfer implementation is based only on received power. The experimental base stations transmit near 2.17 GHz; the experimental subscriber sets transmit near 2.12 GHz. Both operated under a Bellcore experimental license. The experimental prototypes incorporate application-specific, very large-scale integrated circuits[3] fabricated to demonstrate the feasibility of the low-complexity high-performance digital signal-processing techniques [63, 64] for symbol timing and coherent bit detection. These techniques permit the efficient short TDMA bursts having only 100 b that are necessary for low-delay TDMA implementations. Other digital signal-processing functions in the prototypes are implemented in programmable logic devices. All of the digital signal-processing functions combined require about 1/10 of the logic gates that are required for digital signal processing in vehicular digital cellular mobile implementations [42, 62, 63]; that is, this low-complexity PCS implementation having no delay-dispersion-compensating circuits and

[3] Applications specific integrated circuits (ASIC), very large-scale integration (VLSI).

Ford Motor Co.
Exhibit 1030
Page 361

no forward error-correction decoding and is about 1/10 as complex as the digital cellular implementations that include these functions.[4] The 32 kb/s ADPCM speech-encoding in the low-complexity PCS implementation is also about 1/10 as complex as the less than 10-kb/s speech encoding used in digital cellular implementations. This significantly lower complexity will continue to translate into lower power consumption and cost. It is particularly important for low-power pocket personal communicators with power control in which the DC power expended for radio frequency transmitting can be only tens of milliwatts for significant lengths of time.

The experimental radio links have been tested in the laboratory for detection sensitivity [bit error rate (BER) vs SNR] [18, 61, 66] and for performance against cochannel interference [1] and intersymbol interference caused by multipath delay spread [66]. These laboratory tests confirm the performance of the radio-link techniques. In addition to the laboratory tests, qualitative tests have been made in several PCS environments to compare these experimental prototypes with several United States CT-1 cordless telephones at 50 MHz, with CT-2 cordless telephones at 900 MHz, and with DCT-900 cordless telephones at 900 MHz. Some of these comparisons have been reported [8, 71, 84, 85]. In general, depending on the criteria, e.g., either no degradation or limited degradation of circuit quality, these WACS experimental prototypes covered areas inside buildings that ranged from 1.4 to 4 times the areas covered by the other technologies. The coverage areas for the experimental prototypes were always substantially limited in two or three directions by the outside walls of the buildings. These area factors could be expected to be even larger if the coverage were not limited by walls, i.e., once all of a building is covered in one direction, no more area can be covered no matter what the radio link margin. The earlier comparisons [8, 84, 85] were made with only two-branch uplink diversity before subscriber-set transmitting antenna switching was implemented and, with only one radio port before automatic radio-link transfer was implemented. The later tests [71] included these implementations. These reported comparisons agree with similar unreported comparisons made in a Bellcore Laboratory building. Similar coverage comparison results have been noted for a 900-MHz ISM-band cordless telephone compared to the 2-GHz experimental prototype. The area coverage factors (e.g., $\times 1.4$ to $\times 4$) could be expected to be even greater if the cordless technologies had also been operated at 2 GHz since attenuation inside buildings between similar small antennas is about 7 dB greater at 2 GHz than at 900 MHz [35, 36] and the 900 MHz handsets transmitted only 3 dB less average power than the 2-GHz experimental prototypes. The greater area coverage demonstrated for this technology is expected because of the different compromises noted earlier; the following, in particular.

1. Coherent detection of QAM provides more detection sensitivity than noncoherent detection of frequency-shift modulations [17].
2. Antenna diversity mitigates bursts of errors from multipath fading [66].
3. Error detection and blanking of TDMA bursts having errors significantly improves perceived speech quality [72]. (Undetected errors in the most significant bit cause sharp audio pops that seriously degrade perceived speech quality.)
4. Robust symbol timing and burst and frame synchronization reduce the number of frames in error due to imperfect timing and synchronization [66].

[4]Some indication of VLSI complexity can be seen by the number of people required to design the circuits. For the low-complexity TDMA ASIC set, only one person part time plus a student part time were required; the complex CDMA ASIC has six authors on the paper alone.

Ford Motor Co.
Exhibit 1030
Page 362

5. Transmitting more power from the radio port compared to the subscriber set offsets the less sensitive subscriber set receiver compared to the port receiver that results from power and complexity compromises made in a portable set.

Of course, as expected, the low-power (10-mW) radio links cover less area than high-power (0.5-W) cellular mobile pocketphone radio links because of the 17-dB transmitter power difference resulting from the compromises discussed previously. In the case of vehicular mounted sets, even more radio-link advantage accrues to the mobile set because of the higher gain of vehicle-mounted antennas and higher transmitter power (3 W).

{ The power difference between a low-tier PACS handset and a high-tier PCS or cellular pocket handset is not as significant in limiting range as is often portrayed. Other differences in deployment scenarios for low-tier and high-tier systems are as large or even larger factors, e.g., base station antenna height and antenna gain. This can be seen by considering using the same antennas and receiver noise figures at base stations and looking at the range of high-tier and low-tier handsets. High-tier handsets typically transmit a maximum average power of 0.5 watt. The PACS handset average transmit power is 25 milliwatts (peak power is higher for TDMA, but equal comparisons can be made considering average power and equivalent receiver sensitivities). This power ratio of $\times 20$ translates to approximately a range reduction of a factor of about 0.5 for an environment with a distance dependence of $1/(\mathrm{d})^4$ or a factor of about 0.4 for a $1/(\mathrm{d})^{3.5}$ environment. These represent typical values of distance dependence for PCS and cellular environments. Thus, if the high-tier handset would provide a range of 5 miles in some environment, the low-tier handset would provide a range of 2 to 2.5 miles in the same environment, if the base station antennas and receiver noise figures were the same. This difference in range is no greater than the difference in range between high-tier PCS handsets used in 1.9 GHz systems and cellular handsets with the same power used in 800 MHz systems. Range factors are discussed further in the next section. }

15.5.2 {Coverage, Range, Speed, and Environments

Interest has been expressed in having greater range for low-tier PCS technology for low-population-density areas. One should first note that the range of a wireless link is highly dependent on the amount of clutter or obstructions in the environment in which it is operated. For example, radio link calculations that result in a 1400-ft base station (radio-port) separation at 1.9 GHz contain over 50-dB margin for shadowing from obstructions and multipath effects [25, 37]. Thus, in an environment without obstructions, e.g., along a highway, the base station separation can be increased at least by a factor of 4 to over a mile, i.e., 25 dB for an attenuation characteristic of d^{-4}, while providing the same quality of service, without any changes to the base station or subscriber transceivers, and while still allowing over 25-dB margin for multipath and some shadowing. This remaining margin allows for operation of a handset inside an automobile. In such an unobstructed environment, multipath RMS delay spread [21, 33] will still be less than the 0.5 μs in which PACS was designed to operate [28].

Operation at still greater range along unobstructed highways or at a range of a mile along more obstructed streets can be obtained in several ways. Additional link gain of 6 dB can be obtained by going from omnidirectional antennas at base stations to 90° sectored antennas (four sectors). Another 6 dB can be obtained by raising base station antennas by a factor of 2 from 27 ft to 55 ft in height. This additional 12 dB will allow another factor of 2 increase in range to 2-mile base station separation along highways, or to about 3000-ft separation in residential areas. Even higher-gain and taller antennas could be used

Ford Motor Co.
Exhibit 1030
Page 363

to concentrate coverage along highways, particularly in rural areas. Of course, range could be further increased by increasing the power transmitted.

As the range of the low-tier PACS technology is extended in cluttered areas by increasing link gain, increased RMS delay spread is likely to be encountered. This will require increasing complexity in receivers. A factor of 2 in tolerance of delay spread can be obtained by interference-canceling signal combining [76, 77, 78, 79, 80] from two antennas instead of the simpler selection diversity combining originally used in PACS. This will provide adequate delay-spread tolerance for most suburban environments [21, 33].

The PACS downlink contains synchronization words that could be used to train a conventional delay-spread equalizer in subscriber set receivers. Constant-modulus (blind) equalization will provide greater tolerance to delay spread in base station receivers on the uplink [45, 46, 47, 48] than can be obtained by interference-cancellation combining from only two antennas. The use of more base-station antennas and receivers can also help mitigate uplink delay spread. Thus, with some added complexity, the low-tier PACS technology can work effectively in the RMS delay spreads expected in cluttered environments for base station separations of 2 miles or so.

The guard time in the PACS TDMA uplink is adequate for 1-mile range, i.e., 2-mile separation between base station and subscriber transceivers. A separation of up to 3 miles between transceivers could be allowed if some statistical outage were accepted for the few times when adjacent uplink timeslots are occupied by subscribers at the extremes of range (near–far). With some added complexity in assigning timeslots, the assignment of subscribers at very different ranges to adjacent timeslots could be avoided, and the base station separation could be increased to several miles without incurring adjacent slot interference. A simple alternative in low-density (rural) areas, where lower capacity could be acceptable and greater range could be desirable, would be to use every other timeslot to ensure adequate guard time for range differences of many tens of miles. Also, the capability of transmitter time advance has been added to PACS standard in order to increase the range of operation. Such time advance is applied in the cellular TDMA technologies.

The synchronization, carrier recovery, and detection in the low-complexity PACS transceivers will perform well at highway speeds. The two-receiver diversity used in uplink transceivers also will perform well at highway speeds. The performance of the single-receiver selection diversity used in the low-complexity PACS downlink transceivers begins to deteriorate at speeds above about 30 mi/h. However, at any speed, the performance is always at least as good as that of a single transceiver without the low-complexity diversity. Also, fading in the relatively uncluttered environment of a highway is likely to have a less severe Ricean distribution, so diversity will be less needed for mitigating the fading. Cellular handsets do not have diversity. Of course, more complex two-receiver diversity could be added to downlink transceivers to provide two-branch diversity performance at highway speeds. It should be noted that the very short 2.5-ms TDMA frames incorporated into PACS to provide low transmission delay (for high speech quality) also make the technology significantly less sensitive to high-speed fading than the longer-frame-period cellular technologies. The short frame also facilitates the rapid coordination needed to make reliable high-speed handoffs between base stations. Measurements on radio links to potential handoff base stations can be made rapidly, i.e., a measurement on at least one radio link every 2.5 ms. Once a handoff decision is made, signalling exchanges every 2.5 ms ensure that the radio link handoff is completed quickly. In contrast, the long frame periods in the high-tier (cellular) technologies prolong the time it takes to complete a handoff. As noted earlier, high speed handoff has been demonstrated many times with PACS technology and at speeds over 70 mi/hr. }

Ford Motor Co.
Exhibit 1030
Page 364

15.6 Quality, Capacity, and Economic Issues

Although the several trajectories toward low-tier PCS discussed in the preceding section are clear, it does not fit the existing wireless communications paradigms. Thus, low-tier PCS has attracted less attention than the systems and technologies that are compatible with the existing paradigms. Some examples are cited in the following paragraphs.

The need for intense interaction with an intelligent network infrastructure in order to manage mobility is not compatible with the cordless telephone appliance paradigm. In that paradigm, independence of network intelligence and base units that mimic wireline telephones are paramount.

Wireless data systems often do not admit to the dominance of wireless voice communications and, thus, do not take advantage of the economics of sharing network infrastructure and base station equipment. Also, wireless voice systems often do not recognize the importance of data and messaging and, thus, only add them in as bandaids to systems.

The need for a dense collection of many low-complexity, low-cost, low-tier PCS base stations interconnected with inexpensive fixed-network facilities (copper or fiber based) does not fit the cellular high-tier paradigm that expects sparsely distributed $1 million cell sites. Also, the need for high transmission quality to compete with wireline telephones is not compatible with the drive toward maximizing users-per-cell-site and per megahertz to minimize the number of expensive cell sites. These concerns, of course, ignore the hallmark of frequency-reusing cellular systems. That hallmark is the production of almost unlimited overall system capacity by reducing the separation between base stations. The cellular paradigm does not recognize the fact that almost all houses in the U.S. have inexpensive copper wires connecting telephones to the telephone network. The use of low-tier PCS base stations that concentrate individual user services before backhauling in the network will result in less fixed interconnecting facilities than exist now for wireline telephones. Thus, inexpensive techniques for interconnecting many low-tier base stations are already deployed to provide wireline telephones to almost all houses. { The cost of backhaul to many base stations (radio ports) in a low-tier system is often cited as an economic disadvantage that cannot be overcome. However, this perception is based on existing tariffs for T1 digital lines which are excessive considering current digital subscriber line technology. These tariffs were established many years ago when digital subscriber line electronics were very expensive. With modern low-cost high-rate digital subscriber line (HDSL) electronics, the cost of backhaul could be greatly reduced. If efforts were made to revise tariffs for digital line backhaul based on low cost electronics and copper loops like residential loops, the resulting backhaul costs would more nearly approach the cost of residential telephone lines. As it is now, backhaul costs are calculated based on antiquated high T1 line tariffs that were established for "antique" high cost electronics. }

This list could be extended, but the preceding examples are sufficient, along with the earlier sections of the paper, to indicate the many complex interactions among circuit quality, spectrum utilization, complexity (circuit and network), system capacity, and economics that are involved in the design compromises for a large, high-capacity wireless-access system. Unfortunately, the tendency has been to ignore many of the issues and focus on only one, e.g., the focus on cell site capacity that drove the development of digital-cellular high-tier systems in the United States. Interactions among circuit quality, complexity, capacity, and economics are considered in the following sections.

Ford Motor Co.
Exhibit 1030
Page 365

15.6.1 Capacity, Quality, and Complexity

Although capacity comparisons frequently are made without regard to circuit quality, complexity, or cost per base station, such comparisons are not meaningful. An example in Table 15.5 compares capacity factors for U.S. cellular or high-tier PCS technologies with the low-tier PCS technology, PACS/WACS. The mean opinion scores (MOS) (noted in Table 15.5) for speech coding are discussed later. Detection of speech activity and turning off the transmitter during times of no activity is implemented in IS-95. Its impact on MOS also is noted later. A similar technique has been proposed as E-TDMA for use with IS-54 and is discussed with respect to TDMA system in [29]. Note that the use of low-bit-rate speech coding combined with speech activity degrades the high-tier system's quality by nearly one full MOS point on the five-point MOS scale when compared to 32 kb/s ADPCM. Tandem encoding is discussed in a later section. These speech quality degrading factors alone provide a base station capacity increasing factor of $\times 4 \times 2.5 = \times 10$ over the high-speech-quality low-tier system! Speech coding, of course, directly affects base station capacity and, thus, overall system capacity by its effect on the number of speech channels that can fit into a given bandwidth.

TABLE 15.5 Comparison of Cellular (IS-54/IS-95) and Low-Tier PCS (WACS/PACS). Capacity Comparisons Made without Regard to Quality Factors, Complexity, and Cost per Base Station Are not Meaningful

Parameter	Cellular (High-Tier)	Low-Tier PCS	Capacity Factor
Speech Coding, kb/s	8 (MOS 3.4) No tandem coding	32 (MOS 4.1) 3 or 4 tandem	×4
Speech activity	Yes (MOS 3.2)	No (MOS 4.1)	×2.5
Percentage of good areas, %	90	99	×2
Propagation σ, dB	8	10	×1.5
Total: trading quality for capacity			×30

The allowance of extra system margin to provide coverage of 99% of an area for low-tier PCS versus 90% coverage for high-tier is discussed in the previous section and [29]. This additional quality factor costs a capacity factor of ×2. The last item in Table 15.5 does not change the actual system, but only changes the way that frequency reuse is calculated. The additional 2-dB margin in standard deviation σ, allowed for coverage into houses and small buildings for low-tier PCS, costs yet another factor of ×1.5 in calculation only. Frequency reuse factors affect the number of sets of frequencies required and, thus, the bandwidth available for use at each base station. Thus, these factors also affect the base station capacity and the overall system capacity.

For the example in Table 15.5, significant speech and coverage quality has been traded for a factor of ×30 in base station capacity! Whereas base station capacity affects overall system capacity directly, it should be remembered that overall system capacity can be increased arbitrarily by decreasing the spacing between base stations. Thus, if the PACS low-tier PCS technology were to start with a base station capacity of ×0.5 of AMPS cellular[5] (a

[5]Note that the ×0.5 factor is an arbitrary factor taken for illustrating this example. The so-called ×AMPS factors are only with regard to base station capacity, although they are often misused as system capacity.

Ford Motor Co.
Exhibit 1030
Page 366

much lower figure than the $\times 0.8$ sometimes quoted [20]), and then were degraded in quality as described above to yield the $\times 30$ capacity factor, it would have a resulting capacity of $\times 15$ of AMPS! Thus, it is obvious that making such a base station capacity comparison without including quality is not meaningful.

15.6.2 Economics, System Capacity, and Coverage Area Size

Claims are sometimes made that low-tier PCS cannot be provided economically, even though it is what the user wants. These claims are often made based on economic estimates from the cellular paradigm. These include the following.

- Very low estimates of market penetration, much less than cordless telephones, and often even less than cellular.
- High estimates of base station costs more appropriate to high-complexity, high-cost cellular technology than to low-complexity, low-cost, low-tier technology.
- Very low estimates of circuit usage time more appropriate to cellular than to cordless/wireline telephone usage, which is more likely for low-tier PCS.
- { Backhaul costs based on existing T1 line tariffs that are based on "antique" high cost digital loop electronics. (See discussion in fourth paragraph at start of Section 15.6.) }

Such economic estimates are often done by making absolute economic calculations based on very uncertain input data. The resulting estimates for low-tier and high-tier are often closer together than the large uncertainties in the input data. A perhaps more realistic approach for comparing such systems is to vary only one or two parameters while holding all others fixed and then looking at relative economics between high-tier and low-tier systems. This is the approach used in the following examples.

EXAMPLE 15.1:

In the first example (see Table 15.6), the number of channels per megahertz is held constant for cellular and for low-tier PCS. Only the spacing is varied between base stations, e.g., cell sites for cellular and radio ports for low-tier PCS, to account for the differences in transmitter power, antenna height, etc. In this example, overall system capacity varies directly as the square of base station spacing, but base station capacity is the same for both cellular and low-tier PCS. For the typical values in the example, the resulting low-tier system capacity is $\times 400$ greater, only because of the closer base station spacing. If the two systems were to cost the same, the equivalent low-tier PCS base stations would have to cost less than \$2,500.

This cost is well within the range of estimates for such base stations, including equivalent infrastructure. These low-tier PCS base stations are of comparable or lower complexity than cellular vehicular subscriber sets, and large-scale manufacture will be needed to produce the millions that will be required. Also, land, building, antenna tower and legal fees for zoning approval, or rental of expensive space on top of commercial buildings, represent large expenses for cellular cell sites. Low-tier PCS base stations that are mounted on utility poles and sides of buildings will not incur such large additional expenses. Therefore, costs of the order of magnitude indicated seem reasonable in large quantities. Note that, with these estimates, the per-wireless-circuit cost of the low-tier PCS circuits would be only \$14/circuit compared to \$5,555/circuit for the high-tier circuits. Even if there were a factor

Ford Motor Co.
Exhibit 1030
Page 367

TABLE 15.6 System Capacity/Coverage Area Size/Economics

Example 15.1

Assume channels/MHz are the same for cellular and PCS
Cell site: spacing = 20.000 ft cost $ = 1 M
PCS port: spacing = 1,000 ft
PCS system capacity is $(20000/1000)^2 = 400 \times$ cellular capacity

Then, for the system costs to be the same
Port cost = ($ 1 M/400) $2,500 a reasonable figure

If, cell site and port each have 180 channels
Cellular cost/circuit = $ 1 M/180 = $5,555/circuit
PCS cost/circuit = $2500/180 = $14/circuit

Example 15.2

Assume equal cellular and PCS system capacity
Cell site: spacing = 20,000 ft
PCS port: spacing = 1,000 ft

If, a cell site has 180 channels
then, for equal system capacity, a PCS port needs 180/400 < 1 channel/port

Example 15.3

Quality/cost trade
Cell site: Spacing = 20,000 ft cost = $1 M channels = 180
PCS port: Spacing = 1,000 ft cost = $2,500

Cellular to PCS, base station spacing capacity factor = × 400
PCS to cellular quality reduction factors:

32 to 8 kb/s speech	×4
Voice activity (buying)	×2
99–90% good areas	×2
Both in same environment (same σ)	×1
Capacity factor traded	×16

180 ch/16 = 11.25 channels/port then, $2500/11.25 = $222/circuit
and remaining is ×400/16 = ×25 system capacity of PCS over cellular

of 10 error in cost estimates, or a reduction of channels per radio port of a factor of 10, the per-circuit cost of low-tier PCS would still be only $140/circuit, which is still much less than the per-circuit cost of high-tier.

EXAMPLE 15.2:

In the second example (see Table 15.6), the overall system capacity is held constant, and the number of channels/port, i.e., channels/(base station) is varied. In this example, less than 1/2 channel/port is needed, again indicating the tremendous capacity that can be produced with close-spaced low-complexity base stations.

EXAMPLE 15.3:

Since the first two examples are somewhat extreme, the third example (see Table 15.6) uses a more moderate, intermediate approach. In this example, some of the cellular high-tier channels/(base station) are traded to yield higher quality low-tier PCS as in the previous subsection. This reduces the channels/port to 11+, with an accompanying increase in cost/circuit up to $222/circuit, which is still much less than the $5,555/circuit for the high-tier system. Note, also, that the low-tier system still has ×25 the capacity of the high-tier system!

Low-tier base station (Port) cost would have to exceed $62,500 for the low-tier per-circuit cost to exceed that of the high-tier cellular system. Such a high port cost far exceeds any existing realistic estimate of low-tier system costs.

Ford Motor Co.
Exhibit 1030
Page 368

It can be seen from these examples, and particularly Example 15.3, that the circuit economics of low-tier PCS are significantly better than for high-tier PCS, if the user demand and density is sufficient to make use of the large system capacity. Considering the high penetration of cordless telephones, the rapid growth of cellular handsets, and the enormous market projections for wireless PCS noted earlier in this chapter, filling such high capacity in the future would appear to be certain. The major problem is providing rapidly the widespread coverage (buildout) required by the FCC in the United States. If this unrealistic regulatory demand can be overcome, low-tier wireless PCS promises to provide the wireless personal communications that everyone wants.

15.6.3 {Loop Evolution and Economics

It is interesting to note that several wireless loop applications are aimed at reducing cost by replacing parts of wireline or CATV loops with wireless links between transceivers. The economics of these applications are driven by the replacing of labor-intensive wireline and cable technologies with mass-produced solid-state electronics in transceivers.

Consider first a cordless telephone base unit. The cordless base-unit transceiver usually serves one or, at most, two handsets at the end of one wireline loop. Now consider moving such a base unit back along the copper-wire-pair loop end a distance that can be reliably covered by a low-power wireless link [25, 31], i.e., several hundred to a thousand feet or so, and mounting it on a utility pole or a street light pole. This replaces the copper loop end with the wireless link. Many additional copper loop ends to other subscribers will be contained within a circle around the pole having a maximum usable radius of this wireless link. Replace all of the copper loop ends within the circle with cordless base units on the same pole. Note that this process replaces the most expensive parts of these many loops, i.e., the many individual loop ends, with the wireless links from cordless handsets to"equivalent" cordless base units on a pole. Of course, being mounted outside will require somewhat stronger enclosures and means of powering the base units, but these additional costs are considerably more than offset by eliminating the many copper wire drops.

It is instructive to consider how many subscribers could be collected at a pole containing base units. Consider, as an example, a coverage square of 1400 ft on a side (PACS will provide good coverage over this range, i.e., for base unit pole separations of about 1400 ft, at 1.9 GHz). Within this square will be 45 houses for a 1 house/acre density typical of low-housing-density areas, or 180 houses for 4 house/acre density more typical of high-density single-family housing areas. These represent significant concentration of traffic at a pole.

Because of the trunking advantage of the significant number of subscribers concentrated at a pole, they can share a smaller number of base unit, i.e., wireless base unit transceivers, than there are wireless subscriber sets. Therefore, the total cost compared with having a cordless base unit per subscriber also is reduced by the concentration of users.

A single PACS transceiver will support simultaneously eight TDMA channels or circuits at 32 kb/s (or 16 at 16 kb/s or 32 at 8 kb/s) [56]. Of these, one channel is reserved for system control. The cost of such moderate-rate transceivers is relatively insensitive to the number of channels supported; i.e., the cost of such an 8-channel (or 16 or 32) transceiver will be significantly less than twice the cost of a similar one-channel transceiver. Thus, another economic advantage accrues to this wireless loop approach from using time-multiplexed (TDMA) transceivers instead of single-channel-per-transceiver cordless telephone base units.

For an offered traffic of 0.06 Erlang, a typical busy-hour value for a wireline subscriber, a seven-channel transceiver could serve about 40 subscribers at 1% blocking, based on the Erlang B queuing discipline. From the earlier example, such a transceiver could serve most

Ford Motor Co.
Exhibit 1030
Page 369

of the 45 houses within a 1400-ft square. Considering partial penetration, the transceiver capacity is more than adequate for the low-density housing.[6]

Considering the high-density example of 4 houses/acre, a seven-channel transceiver could serve only about 20% of the subscribers within a 1400-ft square. If the penetration became greater than about 20%, either additional transceivers, perhaps those of other service providers, or closer transceiver spacing would be required.

Another advantageous economic factor for wireless loops results when considering time-multiplexed transmission in the fixed distribution facilities. For copper or fiber digital subscriber loop carrier (SLC), e.g., T1 or high-rate digital subscriber line (HDSL), a demultiplexing/multiplexing terminal and drop interface are required at the end of the time-multiplexed SLC line to provide the individual circuits for each subscriber loop-end circuit, i.e., for each drop. The most expensive part of such an SLC terminating unit is the subscriber line cards that provide per-line interfaces for each subscriber drop. Terminating a T1 or HDSL line on a wireless loop transceiver eliminates all per-line interfaces, i.e., all line cards, the most expensive part of a SLC line termination. Thus, the greatly simplified SLC termination can be incorporated within a TDMA wireless loop transceiver, resulting in another cost savings over the conventional copper-wire-pair telephone loop end.

The purpose of the previous discussions is not to give an exact system design or economic analysis, but to illustrate the inherent economic advantages of low-power wireless loops over copper loop ends and over copper loop ends with cordless telephone base units. Some economic analyses have found wireless loop ends to be more economical than copper loop ends when subscribers use low-power wireless handsets. Rizzo and Sollenberger [56] have also discussed the advantageous economics of PACS wireless loop technology in the context of low-tier PCS.

The discussions in this section can be briefly summarized as follows. Replacing copper wire telephone loop ends with low-complexity wireless loop technology like PACS can produce economic benefits in at least four ways. These are.

1. Replacing the most expensive part of a loop, the per-subscriber loop-end, with a wireless link.
2. Taking advantage of trunking in concentrating many wireless subscriber loops into a smaller number of wireless transceiver channels.
3. Reducing the cost of wireless transceivers by time multiplexing (TDMA) a few (7, 15, or 31) wireless loop circuits (channels) in each transceiver.
4. Eliminating per-line interface cards in digital subscriber line terminations by terminating time-multiplexed subscriber lines in the wireless loop transceivers. }

15.7 Other Issues

Several issues in addition to those addressed in the previous two sections continue to be raised with respect to low-tier PCS. These are treated in this section.

[6]The range could be extended by using higher base unit antennas, by using higher-gain directional (sectored) antennas, and/or by increasing the maximum power that can be transmitted.

Ford Motor Co.
Exhibit 1030
Page 370

15.7.1 Improvement of Batteries

Frequently, the suggestion is made that battery technology will improve so that high-power handsets will be able to provide the desired 5 or 6 hours of talk time in addition to 10 or 12 hours of standby time, and still weigh less than one-fourth of the weight of today's smallest cellular handset batteries. This hope does not take into account the maturity of battery technology, and the long history (many decades) of concerted attempts to improve it. Increases in battery capacity have come in small increments, a few percent, and very slowly over many years, and the shortfall is well over a factor of 10. In contrast, integrated electronics and radio frequency devices needed for low-power low-tier PCS continue to improve and to decrease in cost by factors of greater than 2 in time spans on the order of a year or so. It also should be noted that, as the energy density of a battery is increased, the energy release rate per volume must also increase in order to supply the same amount of power. If energy storage density and release rate are increased significantly, the difference between a battery and a bomb become indistinguishable! The likelihood of a $\times 10$ improvement in battery capacity appears to be essentially zero. If even a modest improvement in battery capacity were possible, many people would be driving electric vehicles.

{ As noted in the addition to the "Reality Check" section, new lithium batteries have become the batteries of choice for the smallest cellular/high-tier PCS handsets. While these lithium batteries have higher energy density than earlier nickel cadmium batteries, they still fall far short of the factor of 10 improvement that was needed to make long talk time, small size, and low weight possible. With the much larger advance in electronics, the battery is even more dominant in the size and weight of the newest cellular handsets. The introduction of these batteries incurred considerable startup pain because of the greater fire and explosive hazard associated with lithium materials, i.e., closer approach to a bomb. Further attempts in this direction will be even more hazardous. }

15.7.2 People Only Want One Handset

This issue is often raised in support of high-tier cellular handsets over low-tier handsets. Whereas the statement is likely true, the assumption that the handset must work with high-tier cellular is not. Such a statement follows from the current large usage of cellular handsets; but such usage results because that is the only form of widespread wireless service currently available, not because it is what people want. The statement assumes inadequate coverage of a region by low-tier PCS, and that low-tier handsets will not work in vehicles. The only way that high-tier handsets could serve the desires of people discussed earlier would be for an unlikely breakthrough in battery technology to occur. A low-tier system, however, can cover economically any large region having some people in it. (It will not cover rural or isolated areas but, by definition, there is essentially no one there to want communications anyway.)

Low-tier handsets will work in vehicles on village and city streets at speeds up to 30 or 40 mi/h, and the required handoffs make use of computer technology that is rapidly becoming inexpensive. { As noted earlier, vehicular speed handoff is readily accomplished with PACS. Reliable handoff has been demonstrated for PACS at speeds in excess of 70 mi/hr. } Highways between populated areas, and also streets within them, will need to be covered by high-tier cellular PCS, but users are likely to use vehicular sets in these cellular systems. Frequently the vehicular mobile user will want a different communications device anyway, e.g., a hands-free phone. The use of hands-free phones in vehicles is becoming a legal requirement in some places now and is likely to become a requirement in many more places in the future. Thus, handsets may not be legally usable in vehicles anyway.

Ford Motor Co.
Exhibit 1030
Page 371

With widespread deployment of low-tier PCS systems, the one handset of choice will be the low-power, low-tier PCS pocket handset or voice/data communicator.

{ As discussed in earlier sections, it is quite feasible economically to cover highways between cities with low-tier systems, if the low-tier base stations have antennas with the same height and gain as used for cellular and high-tier PCS systems. (The range penalty for the lower power was noted earlier to be only on the order of 1/2, or about the same as the range penalty in going from 800 MHz cellular to 1.9 GHz high-tier PCS.) }

There are approaches for integrating low-tier pocket phones or pocket communicators with high-tier vehicular cellular mobile telephones. The user's identity could be contained either in memory in the low-tier set or in a small smart card inserted into the set, as is a feature of the European GSM system. When entering an automobile, the small low-tier communicator or card could be inserted into a receptacle in a high-tier vehicular cellular set installed in the automobile.[7] The user's identity would then be transferred to the mobile set. { "Car adapters" that have a cradle for a small cellular handset providing battery charging and connection to an outside antenna are quite common — e.g., in Sweden use of such adapters is commonplace. Thus, this concept has already evolved significantly, even for the disadvantaged cellular handsets when they are used in vehicles. } The mobile set could then initiate a data exchange with the high-tier system, indicating that the user could now receive calls at that mobile set. This information about the user's location would then be exchanged between the network intelligence so that calls to the user could be correctly routed.[8] In this approach the radio sets are optimized for their specific environments, high-power, high-tier vehicular or low-power, low-tier pedestrian, as discussed earlier, and the network access and call routing is coordinated by the interworking of network intelligence. This approach does not compromise the design of either radio set or radio system. It places the burden on network intelligence technology that benefits from the large and rapid advances in computer technology.

The approach of using different communications devices for pedestrians than for vehicles is consistent with what has actually happened in other applications of technology in similarly different environments. For example, consider the case of audio cassette tape players. Pedestrians often carry and listen to small portable tape players with lightweight headsets (e.g., a Walkman).[9] When one of these people enters an automobile, he or she often removes the tape from the Walkman and inserts it into a tape player installed in the automobile. The automobile player has speakers that fill the car with sound. The Walkman is optimized for a pedestrian, whereas the vehicular-mounted player is optimized for an automobile. Both use the same tape, but they have separate tape heads, tape transports, audio preamps, etc. They do not attempt to share electronics. In this example, the tape cassette is the information-carrying entity similar to the user identification in the personal communications example discussed earlier. The main points are that the information is shared among different devices but that the devices are optimized for their environments and do not share electronics.

Similarly, a high-tier vehicular-cellular set does not need to share oscillators, synthesizers, signal processing, or even frequency bands or protocols with a low-tier pocket-size communicator. Only the information identifying the user and where he or she can be reached

[7] Inserting the small personal communicator in the vehicular set would also facilitate charging the personal communicator's battery.

[8] This is a feature proposed for FPLMTS in CCIR Rec. 687.

[9] Walkman is a registered trademark of Sony Corporation.

Ford Motor Co.
Exhibit 1030
Page 372

needs to be shared among the intelligence elements, e.g., routing logic, databases, and common channel signalling [26, 29] of the infrastructure networks. This information exchange between network intelligence functions can be standardized and coordinated among infrastructure subnetworks owned and operated by different business entities (e.g., vehicular cellular mobile radio networks and intelligent low-tier PCS networks). Such standardization and coordination are the same as are required today to pass intelligence among local exchange networks and interexchange carrier networks.

15.7.3 Other Environments

Low-tier personal communications can be provided to occupants of airplanes, trains, and buses by installing compatible low-tier radio access ports inside these vehicles. The ports can be connected to high-power, high-tier vehicular cellular mobile sets or to special air-ground or satellite-based mobile communications sets. Intelligence between the internal ports and mobile sets could interact with cellular mobile, air-ground, or satellite networks in one direction, using protocols and spectrum allocated for that purpose, and with low-tier personal communicators in the other direction to exchange user identification and route calls to and from users inside these large vehicles. Radio isolation between the low-power units inside the large metal vehicles and low-power systems outside the vehicles can be ensured by using windows that are opaque to the radio frequencies. Such an approach also has been considered for automobiles, i.e., a radio port for low-tier personal communications connected to a cellular mobile set in a vehicle so that the low-tier personal communicator can access a high-tier cellular network. (This could be done in the United States using unlicensed PCS frequencies within the vehicle.)

15.7.4 Speech Quality Issues

All of the PCS and cordless telephone technologies that use CCITT standardized 32-kb/s ADPCM speech encoding can provide similar error-free speech distortion quality. This quality often is rated on a five-point subjective mean opinion score (MOS) with 5 excellent, 4 good, 3 fair, 2 poor, and 1 very poor. The error-free MOS of 32-kb/s ADPCM is about 4.1 and degrades very slightly with tandem encodings. Tandem encodings could be expected in going from a digital-radio PCS access link, through a network using analog transmission or 64-kb/s PCM, and back to another digital-radio PCS access link on the other end of the circuit. In contrast, a low-bit-rate (<10-kb/s) vocoder proposed for a digital cellular system was recently reported [54] to yield an MOS of 3.4 on an error-free link without speech-activity detection. This score dropped to 3.2 when speech-activity detection was implemented to increase system capacity. This nearly one full point decrease on the five-point MOS score indicates significant degradation below accepted CCITT wireline speech distortion quality. Either almost half of the population must have rated it as poor or most of the population must have rated it as only fair. It should also be noted that these MOS scores may not reflect additional degradation that may occur in low-bit-rate speech encoding when the speech being encoded is combined with acoustical noise in a mobile environment, e.g., tire, wind, and engine noise in automobiles and street noise, background talking, etc., in handheld phone environments along streets and in buildings. Comments from actual users of low-bit-rate speech technology in acoustically noisy environments suggest that the MOS scores just quoted are significantly degraded in these real world environments. Waveform coders, e.g., ADPCM are not likely to experience degradation from such background noise. In addition, the low-bit-rate speech encoding is not at all tolerant of the tandem speech encodings that will inevitably occur for PCS for many years. That is, when low-bit-rate

Ford Motor Co.
Exhibit 1030
Page 373

speech is transcoded to a different encoding format, e.g., to 64 kb/s as is used in many networks or from an IS-54 phone on one end to a GSM or IS-95 phone on the other end, the speech quality deteriorates precipitously. Although this may not be a serious issue for a vehicular mobile user who has no choice other than not to communicate at all, it is likely to be a serious issue in an environment where a wireline telephone is available as an alternative. It is also less serious when there are few mobile-to-mobile calls through the network, but as wireless usage increases and digital mobile-to-mobile calls become commonplace, the marginal transcoded speech quality is likely to become a serious issue. These comments in this paragraph are generally applicable to speech encoding at rates of 13 kb/s or less. { The predictions above about the seriousness of the speech coding quality issues have been proven to be true in virtually all cellular technologies [2, 5]. The much touted speech quality of the original IS-95 CDMA 8 kb/s QCELP speech coder proved to be unacceptable to users.

The 8 kb/s speech coder was replaced with a 13 kb/s speech coder when IS-95 CDMA technology was deployed for high-tier PCS at 1.9 GHz. Many of the 800 MHz CDMA systems are converting or considering converting to the higher bit rate coder to improve their speech quality. This change was accompanied by a base station capacity reduction of 1/3, but the CDMA advocates seldom acknowledge this fact. The GSM effort to create a "half-rate" speech coder was redirected to develop an improved speech quality coder at 13 kb/s that would be less sensitive to acoustic background noise, bit errors, tandem encoding, etc. The quality of the new coder is better under some conditions. The 8 kb/s North American TDMA/IS-136 speech coder was redesigned to improve speech quality. Some improvement has been achieved. In all these revised coder cases, the issues discussed above in the original article still apply. }

In the arena of transmission delay, the short-frame (2-ms) FDMA/TDD and TDMA technologies (e.g., CT-2 and WACS noted earlier) can readily provide single-radio-link round-trip delays of <10 ms and, perhaps, even <5 ms. The longer frame (10 ms and greater) cordless-phone TDMA technologies, e.g., DCT-900/CT-3/DECT and some ISM-band implementations, inherently have a single-link round-trip delay of at least 20 ms and can range 30-40 ms or more in some implementations. As mentioned earlier, the digital vehicular-cellular technologies with low-bit-rate speech encoding, bit interleaving, forward error-correction decoding, and relatively long frame time (~16–20 ms) result in single-link round-trip delays on the order of 200 ms, well over an order of magnitude greater than the short-frame technologies, and on the same order of magnitude as a single-direction synchronous satellite link. { Even with the somewhat improved low bit rate speech coders noted previously, the transmission delay remains excessive for all of the cellular/high-tier PCS digital technologies. } It should be noted that almost all United States domestic long-distance telephone circuits have been removed from such satellite links, and many international satellite links also are being replaced by undersea fiber links. These long-distance-circuit technology changes are made partially to reduce the perceptual impairment of long transmission delay.

15.7.5 New Technology

New technology, e.g., spread spectrum or CDMA, is sometimes offered as a solution to both the higher-tier cell site capacity and transmitter power issues. As these new technologies are pursued vigorously, however, it becomes increasingly evident that the early projections were considerably overoptimistic, that the base station capacity will be about the same as other technologies [29], and that the high complexity will result in more, not less, power consumption.

Ford Motor Co.
Exhibit 1030
Page 374

With the continuing problems and delays in initial deployments, there is increasing concern throughout the industry as to whether CDMA is a viable technology for high-capacity cellular applications. With the passage of time, it is becoming more obvious that Viterbi was correct in his 1985 paper in which he questioned the use of spread spectrum for commercial communications [73].

The IS-95 proposal is considerably more technically sophisticated than earlier spread spectrum proposals. It includes fast feedback control of mobile transmitter power, heavy forward error correction, speech detection and speech-encoding rate adjustment to take advantage of speech inactivity, and multiple receiver correlators to latch onto and track resolvable multipath maxima [57]. The spreading sequence rate is 1.23 MHz.

The near–far problem is addressed directly and elegantly on the uplink by a combination of the fast-feedback power control and a technique called soft handoff that permits the instantaneous selection of the best paths between a mobile and two cell sites. Path selection is done on a frame-by-frame basis when paths between a mobile and the two cell sites are within a specified average level (perhaps 6–10 dB) for each other. This soft handoff provides a form of macroscopic diversity [10] between pairs of cell sites when it is advantageous. Increasing capacity by soft handoff requires precise time synchronization (on the order of a microsecond) among all cell sites in a system. An advantage of this proposal is that frequency coordination is not needed among cell sites since all sites can share a frequency channel. Coordination of the absolute time delays of spreading sequences among cell sites is required, however, since these sequence delays are used to distinguish different cell sites for initial access and for soft handoff. Also, handoff from one frequency to another is complicated.

Initially, the projected cell site capacity of this CDMA system, determined by mathematical analysis and computer simulation of simplified versions of the system, was $\times 20$ to $\times 40$ that of the analog AMPS, with a coverage criterion of 99% of the covered area [11]. However, some other early estimates [83] suggested that the factors were more likely to be $\times 6$ to $\times 8$ of AMPS.

A limited experiment was run in San Diego, California, during the fourth quarter of 1991 under the observation of cellular equipment vendors and service providers. This experiment had 42–62 mobile units in fewer than that many vehicles,[10] and four or five cell sites, one with three sectors. Well over half of the mobiles needed to provide the interference environment for system capacity tests were simulated by hardware noise simulation by a method not yet revealed for technical assessment. Estimates of the cell site capacity from this CDMA experiment center around $\times 10$ that of AMPS [12, 54][11] with coverage criteria $<99\%$, perhaps 90-95%, and with other capacity estimates ranging between $\times 8$ and $\times 15$.

This experiment did not exercise several potential capacity-reducing factors; the following, for example.

1. Only four cells participated in capacity tests. The test mobiles were all located in a relatively limited area and had limited choices of cell sites with which to communicate for soft handoffs. This excludes the effects of selecting a strong cell-site downlink for soft handoff that does not have the lowest uplink attenuation because of uncorrelated uplink and downlink multipath fading at slow vehicle speeds [3].

[10] Some vehicles contained more than one mobile unit.

[11] AT&T stated that the San Diego data supported a capacity improvement over analog cellular of at least $\times 10$.

Ford Motor Co.
Exhibit 1030
Page 375

2. The distribution of time-dispersed energy in hilly environments like San Diego usually is more concentrated around one or two delays than is the dispersed energy scattered about in heavily built-up urban areas like downtown Manhattan [13, 14], or Chicago. Energy concentrated at one or two delays is more fully captured by the limited number of receiver correlators than is energy more evenly dispersed in time.
3. Network delay in setting up soft handoff channels can result in stronger paths to other cell sites than to the one controlling uplink transmitter power. This effect can be more pronounced when coming out of shadows of tall buildings at intersections in heavily built-up areas. The effect will not occur as frequently in a system with four or five cell sites as it will in a large, many cell site system.

All of these effects and others [4] will increase the interference in a large system, similarly to the increase in interference that results from additional mobiles and, thus, will decrease cell site capacity significantly below that estimated in the San Diego trial. Factors like these have been shown to reduce the San Diego estimate of $\times 10$ to an expected CDMA capacity of $\times 5$ or $\times 6$ of analog AMPS [3, 4, 13]. This further reduction in going from a limited experiment to a large-scale system in a large metropolitan area is consistent with the reduction already experienced in going from simplified theoretical estimates to the limited experiment in a restricted environment. { There are many "small" effects that decrease the performance and, thus, the base station capacity of actual mobile CDMA systems compared to calculations of performance based on ideal assumptions — e.g., stationary channel and additive white Gaussian noise analysis. Some of these effects, in addition to those noted above, included inaccuracy in assigning correlators ("fingers") to time varying statistically nonstationary multiple "paths" at different delays when many "paths" exist, and lack of diversity because many environments have delay spreads less than the resolution of correlators. }

The San Diego trial also indicated a higher rate of soft(er) handoffs [54] between antenna sectors at a single cell site than expected for sectors well isolated by antenna patterns. This result suggests a lower realizable sectorization gain because of reflected energy than would be expected from more idealized antennas and locations. This could further reduce the estimated cell site capacity of a large-scale system.

Even considering the aforementioned factors, capacity increases of $\times 5$ or $\times 6$ are significant. These estimates, however, are consistent with the factor of $\times 3$ obtained from low-bit-rate (<10-kb/s) speech coding and the $\times 2$ to $\times 2.5$ obtained by taking advantage of speech pauses. These factors result in an expected increase of $\times 6$–7.5, with none of these speech-processing-related contributions being directly attributable to the spread-spectrum processing in CDMA. These results are consistent with the factor of $\times 6$–8 estimate made earlier [83] and are not far from the factor of $\times 8$ quoted recently [86].

Thus, it is clear that new high-complexity, high-tier technology will not be a substitute for low-complexity, low-power, low-tier PCS.

{ Since the writing for the first edition of this handbook, IS-95 based CDMA systems have been deployed and are providing service in several parts of the world, e.g., the U.S.A., Korea, and Hong Kong. The "high priests" and the "disciples" of CDMA continue to espouse their "belief" in the superiority of their CDMA technology. Good hard data on system performance is still scarce and of questionable reliability. However, there is sufficient evidence from some "usually reliable sources" and from "conversations in the halls" to confirm that, not only does CDMA not perform as well as was originally promised, but also in many places it is likely to be providing performance that is inferior to the other high-tier digital technologies. In spite of the continuing dismal failures of CDMA technology to

Ford Motor Co.
Exhibit 1030
Page 376

meet the early claims for it, a significant fraction of the wireless industry appears to be committed to pursuing variations of CDMA for future applications. At this time it is not obvious how many more redesign and reprogramming failures, or perhaps even failures of manufacturing companies and service providers will be needed to overcome the religiously fanatic attachment of a part of the industry to CDMA. It appears that some are heavily committed again to "sweep on to the grand fallacy." Viterbi may even have been over optimistic about CDMA in 1985 [73], but he was correct in the way it attracts advocates.

IS-95 CDMA has failed to satisfy any of the major claims made for it in the late 1980s and early 1990s. The listing in Table 15.7 and the following discussion cites specific claims that have NOT been met now that reality has set in.

TABLE 15.7 CDMA Scoresheet

Early Claim	Reality
Easy to install and expand (no frequency planning)	Many "parameters" to adjust; adjustment is very difficult; many software revisions required
"Capacity" of AMPS ×20 (with high expectation of AMPS ×40)	AMPS ×3 or 4; AMPS × <3? (at unspecified but likely high call dropping rate) (continuously revised downward over time)
No more dropped calls	More dropped calls than other technologies (some reports say up to 40% dropped calls when system is loaded!)
No problem with interference (CDMA "loves" interference)	Can't live with interference from AMPS or other sources
Greater range and coverage than other technologies	System coverage "breathes" as loading increases; range often is less than other technologies
"Toll quality" 8 kb/s QCELP speech	Original 8 kb/s QCELP speech quality worse than VSELP in TDMA; required change to 13 kb/s coder (change reduced "capacity")

Much of the original CDMA hype was focused on undefined cell site capacity (see text from original Handbook, particularly this section and the section on "Capacity, Quality, and Complexity.") Information from many sources (e.g., [9], Benjamin and Lusignan, 1997 — Benjamin is with Lucent Technologies; Lusignan with Stanford University; [40, 49, 69], etc.) suggest that with 13 kb/s speech coding, "fully loaded" systems may provide base station capacities on the order of AMPS ×3 or perhaps ×4. However, the meaning of fully loaded is not well defined. If it is taken to be at a very high call dropping rate (sometimes said to be as high as 40%!), perhaps even these low factors are again unrealistically inflated. A usable system loading must provide an acceptable call dropping rate, perhaps 1% or less. If loading is reduced to provide sufficient operating margin to prevent too many dropped calls, perhaps the usable base station capacity may be even significantly less than AMPS ×3. (Note: Publicly presented CDMA claims have been consistently inflated in the past. Why should we be any more inclined to believe them now?) The capacity situation is so bad that the CDMA crowd doesn't even want to discuss this topic anymore. They now attempt to shift attention to the fact that there are operating CDMA systems, and try to brush aside the fact CDMA performance is so very poor. Low capacity has a significant negative impact on the economics of systems; high call dropping rates have a significant negative impact on service quality.

Another big early claim for CDMA was that it was easy to install and expand, that is, that it needed little engineering and planning for installation because it didn't require frequency reuse and power coordination among base stations. In reality, many parameters

Ford Motor Co.
Exhibit 1030
Page 377

require adjustment and critical thresholds require setting for a system to provide even the low capacity numbers cited above. Many software revisions have been made in attempts to cope with fundamental problems that keep appearing and reappearing in new deployments and expansions. Deployments and service starts have been consistently late by many months and even years. [40, 69, and others]. A partial list of late CDMA deployments that required many months or even years of "parameter adjustment" and/or "software revisions" and/or "redesign" include those in Seattle, WA, U.S.A. (800 MHz); Los Angeles, CA, U.S.A. (800 MHz); Hong Kong, China (800 MHz); Trenton, NJ, U.S.A.; Seoul, Korea (800 MHz); Houston and Dallas, TX, U.S.A. (1.9 GHz); etc. As these systems were deployed a consistent pattern developed. Announcements would concentrate on a projected future service date, e.g., Seattle, 1Q 1994. When that date approached, news focused on yet another service date elsewhere, e.g., after Seattle was Los Angeles, to divert attention from the deployment problems causing large slips in earlier announcements. Sometimes when service date slips became excessive, sizes of deployed areas were revised downward, and "trials" were substituted for service starts, e.g., Los Angeles. This pattern continued through the early deployments and service eventually was started on systems, even though their performance was significantly below earlier claims. Excuses for slips were often vague, for example: software problems (e.g., Seattle, Hong Kong, and Korea); adjusting parameters (e.g., Seattle, Houston, and Dallas); too much interference from AMPS (e.g., Los Angeles); not enough multipath! (e.g., Seattle). These sound like descriptions of attempts to overcome some of the fundamental problems in mobile CDMA systems. In any event it is obvious that CDMA is not easy to install and is likely more difficult to adjust than other technologies.

The CDMA speech quality issue was discussed earlier in the "Speech Quality Issues" section. It is just another example of unrealistic initial claims.

Throughout the CDMA debate over almost 10 years, when unrealistic claims were questioned and even when sound technical analysis was cited by more conservative members of the wireless technical community, including this author, these members were immediately branded as heretics and their concerns were dismissed with rhetoric by the high priests and their disciples. (In the technical world, "never have so many been misled for so long by so few!") However, as deployments have proceeded and reality set in, even the more conservative initial assessments have, unfortunately, turned out to be somewhat optimistic. Thus, even the conservative heretics in the community now appear to be in the unfortunate position of having also been overoptimistic in their initial estimates of the performance of IS-95 CDMA. }

Statistical Multiplexing, Speech Activity, CDMA, and TDMA

Factors of $\times 2$–2.5 have been projected for the capacity increase possible by taking advantage of pauses in speech. It has been suggested that implementing statistical multiplexing is easier for CDMA systems because it is sometimes thought to be time consuming to negotiate channels for speech spurts for implementation in TDMA systems. The most negative quality-impacting factor in implementing statistical multiplexing for speech, however, is not in obtaining a channel when needed but is in the detection of the onset of speech, particularly in an acoustically noisy environment. The effect of clipping at the onset of speech is evident in the MOS scores noted for the speech-activity implementation in the United States cellular CDMA proposal discussed earlier (i.e., an MOS of 3.4 without statistical multiplexing and of 3.2 with it). The degradation in MOS can be expected to be even greater for encoding that starts with a higher MOS, e.g., 32-kb/s ADPCM.

Ford Motor Co.
Exhibit 1030
Page 378

It was noted earlier that the proposed cellular CDMA implementation was ×10 as complex as the proposed WACS wireless access for personal communications TDMA implementation. From earlier discussion, the CDMA round-trip delay approaches 200 ms, whereas the short 2-ms-frame TDMA delay is <10-ms round trip. It should be noted that the TDMA architecture could permit negotiation for time slots when speech activity is detected. Since the TDMA frames already have capability for exchange of signalling data, added complexity for statistical multiplexing of voice could readily be added within less than 200 ms of delay and less than ×10 in complexity. That TDMA implementation supports 8 circuits at 32 kb/s or 16 circuits at 16 kb/s for each frequency. These are enough circuits to gain benefit from statistical multiplexing. Even more gain could be obtained at radio ports that support two or three frequencies and, thus, have 16–48 circuits over which to multiplex.

A statistical multiplexing protocol for speech and data has been researched at Rutgers WINLAB [39]. The Rutgers packet reservation multiple access (PRMA) protocol has been used to demonstrate the feasibility of increasing capacity on TDMA radio links. These PRMA TDMA radio links are equivalent to slotted ALOHA packet-data networks. Transmission delays of less than 50 ms are realizable. The capacity increase achievable depends on the acceptable packet-dropping ratio. This increase is soft in that a small increase in users causes a small increase in packet-dropping ratio. This is analogous to the soft capacity claimed for CDMA.

Thus, for similar complexity and speech quality, there appears to be no inherent advantage of either CDMA or TDMA for the incorporation of statistical multiplexing. It is not included in the personal communications proposal but is included in cellular proposals because of the different speech-quality/complexity design compromises discussed throughout this chapter, not because of any inherent ease of incorporating it in any particular access technology.

15.7.6 High-Tier to Low-Tier or Low-Tier to High-Tier Dual Mode

Industry and the FCC in the United States appear willing to embrace multimode handsets for operating in very different high-tier cellular systems, e.g., analog FM AMPS, TDMA IS-54, and CDMA IS-95. Such sets incur significant penalties for dual mode operation with dissimilar air interface standards and, of course, incur the high-tier complexity penalties.

It has been suggested that multimode high-tier and low-tier handsets could be built around one air-interface standard, for example, TDMA IS-54 or GSM. When closely spaced low-power base stations were available, the handset could turn off unneeded power-consuming circuitry, e.g., the multipath equalizer. The problem with this approach is that the handset is still encumbered with power-consuming and quality-reducing signal processing inherent in the high-tier technology, e.g., error correction decoding and low-bit-rate speech encoding and decoding.

{ With widespread and successful deployment of GSM, perhaps an initial dual-mode strategy of GSM and PACS could be desirable. Of all the high-tier and low-tier technologies, GSM and PACS appear most compatible for dual-mode handset and base station implementation. They are considerably more compatible than the dual-mode combinations, e.g., AMPS/IS-54/IS-136 and AMPS/IS-95, already implemented. }

An alternative dual-mode low-tier, high-tier system based on a common air-interface standard can be configured around the low-tier PACS/WACS system, if such a dual-mode system is deemed desirable in spite of the discussion in this chapter. The range of PACS can readily be extended by increasing transmitter power and/or the height and gain of base station antennas. With increased range, the multipath delay spread will be more severe in some locations [21, 22, 33]. Two different solutions to the increased delay spread can be employed, one for the downlink and another for the uplink. The PACS radio-link

Ford Motor Co.
Exhibit 1030
Page 379

architecture has a specified bit sequence, i.e., a unique word, between each data word on the TDM downlink [7, 28]. This unique word can be used as a training sequence for setting the tap weights of a conventional equalizer added to subscriber sets for use in a high-tier PACS mode. Since received data can be stored digitally [62, 65], tap weights can be trimmed, if necessary, by additional passes through an adaptive equalizer algorithm, e.g., a decision feedback equalizer algorithm.

The PACS TDMA uplink has no unique word. The high-tier uplink, however, will terminate on a base station that can support greater complexity but still be no more complex than the high-tier cellular technologies. Research at Stanford University has indicated that blind equalization, using constant-modulus algorithms (CMA), [58, 70], can be effective for equalizing the PACS uplink. Techniques have been developed for converging the CMA equalizer on the short TDMA data burst.

{ See earlier added section on "Coverage, Range, Speed and Environments." }

The advantages of building a dual-mode high-tier, low-tier PCS system around the low-tier PACS air-interface standard follow.

1. The interface can still support small low-complexity, low-power, high-speech-quality low-tier handsets.
2. Both data and voice can be supported in a PACS personal communicator.
3. In high-tier low-tier dual mode PACS sets, circuits used for low-tier operation will also be used for high-tier operation, with additional circuits being activated only for high-tier operation.
4. The flexibility built into the PACS radio link to handle different data rates from 8 kb/s to several hundred kb/s will be available to both modes of operation.

15.8 Infrastructure Networks

It is beyond the scope of this chapter to consider the details of PCS network infrastructures. There are, however, perhaps as many network issues as there are wireless access issues discussed herein [26, 32, 43, 81]. With the possible exception of the self-organizing WLANS, wireless PCS technologies serve as access technologies to large integrated intelligent fixed communications infrastructure networks.

These infrastructure networks must incorporate intelligence, i.e., database storage, signalling, processing and protocols, to handle both small-scale mobility, i.e., handoff from base station to base station as users move, and large-scale mobility, i.e., providing service to users who roam over large distances, and perhaps from one network to another. The fixed infrastructure networks also must provide the interconnection among base stations and other network entities, e.g., switches, databases, and control processors. Of course, existing cellular mobile networks now contain or are incorporating these infrastructure network capabilities. Existing cellular networks, however, are small compared to the expected size of future high-tier and low-tier PCS networks, e.g., 20 million cellular users in the United States compared with perhaps 100 million users or more each in the future for high-tier and low-tier PCS.

Several other existing networks have some of the capabilities needed to serve as access networks for PCS. Existing networks that could provide fixed base station interconnection include:

- Local exchange networks that could provide interconnection using copper or glass-fiber distribution facilities

Ford Motor Co.
Exhibit 1030
Page 380

- Cable TV networks that could provide interconnection using new glass-fiber and coaxial-cable distribution facilities
- Metropolitan fiber digital networks that could provide interconnection in some cities in which they are being deployed

Networks that contain intelligence, e.g., databases, control processors, and signalling that is suitable or could be readily adapted to support PCS access include:

- Local exchange networks that are equipped with signalling system 7 common channel signalling (SS7 CCS), databases, and digital control processors
- Interexchange networks that are similarly equipped

Data networks, e.g., the internet, could perhaps be adapted to provide the needed intelligence for wireless data access, but they do not have the capacity needed to support large voice/data wireless low-tier PCS access.

Many entities and standards bodies worldwide are working on the access network aspects of wireless PCS. The signalling, control processing, and database interactions required for wireless access PCS are considerably greater than those required for fixed place-to-place networks, but that fact must be accepted when considering such networks.

Low-tier PCS, when viewed from a cellular high-tier paradigm, requires much greater fixed interconnection for the much closer spaced base stations. When viewed from a cordless telephone paradigm of a base unit for every handset and, perhaps, several base units per wireline, however, the requirement is much less fixed interconnection because of the concentration of users and trunking that occurs at the multiuser base stations. One should remember that there are economical fixed wireline connections to almost all houses and business offices in the United States now. If wireless access displaces some of the wireline connections, as expected, the overall need for fixed interconnection could decrease!

15.9 Conclusion

Wireless personal communications embraces about seven relatively distinct groups of tetherless voice and data applications or services having different degrees of mobility for operation in different environments. Many different technologies and systems are evolving to provide the different perceived needs of different groups. Different design compromises are evident in the different technologies and systems. The evidence suggests that the evolutionary trajectories are aimed toward at least three large groups of applications or services, namely, high-tier PCS (current cellular radio), high-speed wireless local-area networks (WLANS), and low-tier PCS (an evolution from several of the current groups). It is not clear to what extent several groups, e.g., cordless telephones, paging, and wide-area data, will remain after some merging with the three large groups. Major considerations that separate current cellular technologies from evolving low-tier low-power PCS technologies are speech quality, complexity, flexibility of radio-link architecture, economics for serving high-user-density or low-user-density areas, and power consumption in pocket carried handsets or communicators. High-tier technologies make use of large complex expensive cell sites and have attempted to increase capacity and reduce circuit costs by increasing the capacity of the expensive cell sites. Low-tier technologies increase capacity by reducing the spacing between base stations, and achieve low circuit cost by using low-complexity, low-cost base stations. The differences between these approaches result in significantly different compromises in circuit quality and power consumption in pocket-sized handsets or communicators. These

Ford Motor Co.
Exhibit 1030
Page 381

kinds of differences also can be seen in evolving wireless systems optimized for data. Advantages of the low-tier PACS/WACS technology are reviewed in the chapter, along with techniques for using that technology in high-tier PCS systems.

References

[1] Afrashteh, A., Sollenberg, N.R., and Chukurov, D.D., Signal to interference performance for a TDMA portable radio link. *IEEE VTC'91,* St. Louis, MO, 19–22, May, 1991.

[2] {Andrews, E.L., "When Digital Doesn't Always Mean Clear," *New York Times,* p. C1, 4, Jun. 26, 1995.}

[3] Ariyavisitakul, S., SIR-based power control in a CDMA system. *IEEE GLOBECOM'92,* Orlando, FL, Paper 26.3, Dec. 6–9, to be published in *IEEE Trans. Comm.,* 1992.

[4] Ariyavisitakul, S., et al., private communications.

[5] {Baugh, C.R., Wireless quality: a ticking time bomb, *Telephony,* 232(9), 50–58, Mar. 3, 1997.}

[6] {Baugh, C.R., Laborde, E., Pandey, V., and Varma, V., Personal access communications system: fixed wireless local loop and mobile configurations and services, *Proc. IEEE,* 1498–1506, July, 1998. }

[7] Bellcore Technical Advisories. Generic framework criteria for universal digital personal communications systems (PCS). FA-TSY-001013(1) March and FA-NWT-001013(2) Dec., and Tech. Ref. 1993. Generic criteria for version 0.1 wireless access communications systems (WACS). (1) Oct. rev. 1, Jun. 1994.

[8] BellSouth Services Inc., Quarterly progress report number 3 for experimental licenses KF2XFO and KF2XFN. To the Federal Communications Commission, Nov. 25, 1991.

[9] { Benjamin, B. and Lusignan, B., CDMA vs. TDMA, Stanford Univ. Center for Telecom. Symposium, The Acceleration of World Wide Wireless Communication, May 21, 1997.}

[10] Bernhardt, R.C., Macroscopic diversity in frequency reuse radio systems. *IEEE J. Sel. Areas in Comm.,* SAC-5, 862–870, Jun. 1987.

[11] Cellular Telecommunications Industries Association. CDMA digital cellular technology open forum. Jun. 6, 1989.

[12] Cellular Telecommunications Industries Association. CTIA Meeting, Washington, D.C., 5–6, Dec. 1991.

[13] Chang, L.F., Dispersive fading effects in CDMA radio systems. *Elect. Lett.,* 28(19), 1801–1802, 1992.

[14] Chang, L.F. and Ariyavisitakul, S., Performance of a CDMA radio communications system with feed-back power control and multipath dispersion. *IEEE GLOBECOM'91,* Phoenix, AZ, 1017–1021, Dec. 1991.

[15] Chuang, J.C.-I., Performance issues and algorithms for dynamic channel assignment. *IEEE JSAC,* Aug. 1993.

[16] Chuang, J.C.-I., Performance limitations of TDD wireless personal communications with asynchronous radio ports. *Electron. Lett.,* 28, 532–533, Mar. 1992.

[17] Chuang, J.C.-I., Comparison of coherent and differential detection of BPSK and QPSK in a quasistatic fading channel. *IEEE Trans. Comm.,* 565–576, May 1990.

[18] Chuang, J.C.-I. and Sollenberger, N.R., Burst coherent detection with robust frequency and timing estimation for portable radio communications. *IEEE GLOBECOM'88,* Hollywood, FL, 28–30, Nov. 1988.

[19] Chuang, J.C.-I., Sollenberger, N.R., and Cox, D.C., A pilot based dynamic channel assignment scheme for wireless access TDMA/FDMA systems. *Proc. IEEE ICUPC'93,* Ottawa, Canada, Oct. 12–15, 706–712, 1994. *Intl. J. Wireless Inf. Net.,* 1(1), 37–48, 1993.

[20] Cook, C.I., Development of air interface standards for PCS. *IEEE Personal Comm.,* 4th Quarter, 30–34, 1994.

Ford Motor Co.
Exhibit 1030
Page 382

[21] Cox, D.C., Delay-Doppler characteristics of multipath propagation at 910 MHz in a suburban mobile radio environment. *IEEE Trans. on Antennas and Propagation,* 625–635, Sep. 1972.

[22] Cox, D.C., Multipath delay spread and path loss correlation for 910 MHz urban mobile radio propagation. *IEEE Trans. on Veh. Tech.,* 340–344, Nov. 1977.

[23] Cox, D.C., Universal portable radio communications. *IEEE Trans. on Veh. Tech.,* 117–121, Aug. 1985.

[24] Cox, D.C., Research toward a wireless digital loop. *Bellcore Exchange,* 2, 2–7, Nov./Dec. 1986.

[25] Cox, D.C., Universal digital portable radio communications. *Proc. IEEE,* 75, 436–477, Apr. 1987.

[26] Cox, D.C., Portable digital radio communications—An approach to tetherless access. *IEEE Comm. Mag.,* 30–40, Jul. 1989.

[27] Cox, D.C., Personal communications—A viewpoint. *IEEE Comm. Mag.,* 8–20, Nov. 1990.

[28] Cox, D.C., A radio system proposal for widespread low-power tetherless communications. *IEEE Trans. on Comm.,* 324–335, Feb. 1991.

[29] Cox, D.C., Wireless network access for personal communications. *IEEE Comm. Mag.,* 96–115, Dec. 1992.

[30] {Cox, D.C., Wireless loops: what are they. *Intl. J. Wireless Inf. Net.,* 3(3), Plenum Press, 1996.}

[31] Cox, D.C., Arnold, H.W., and Porter, P.T., Universal digital portable communications—a system perspective. *IEEE JSAC,* JSAC-5, 764–773, Jun. 1987.

[32] Cox, D.C., Gifford, W.G., and Sherry, H., Low-power digital radio as a ubiquitous subscriber loop. *IEEE Comm. Mag.,* 92–95, Mar. 1991.

[33] {Devasirvatham, D.M.J., Radio propagation studies in a small city for universal portable communications. *IEEE VTC'88,* Conference Record. Philadelphia, PA, 100–104, Jun. 15–17, 1988.}

[34] {Devasirvatham, D.M.J., Radio propagation studies in a small city for universal portable communications, *IEEE VTC'88,* Philadelphia, PA, 100–104, Jun. 15–17, 1988.}

[35] Devasirvatham, D.M.J., et al., Radio propagation measurements at 850 MHz, 1.7 GHz and 4 GHz inside two dissimilar office buildings. *Elect. Lett.,* 26(7), 445–447, 1990a.

[36] Devasirvatham, D.M.J., et al., Multi-frequency radiowave propagation measurements in the portable radio environment. *IEEE ICC'90,* 1334–1340, Apr. 1990b.

[37] {Devasirvatham, D.M.J., Murray, R.R., Arnold, H.W., and Cox, D.C., Four-frequency CW measurements in residential environments for personal communications, symposium on personal indoor and mobile radio communications (PIMRC'93), Yokohama, Japan, 201–205, Sep. 9–11, 1993.}

[38] {Goldhirsh, J. and Vogel, W., Mobile satellite system fade and statistics for shadowing and multipath from roadside trees at UFH and L-band, *IEEE Transactions on Antennas and Propagation,* 37, Apr. 1989.}

[39] Goodman, D.J., Trends in cellular and cordless communications. *IEEE Comm. Mag.,* 31–40, Jun. 1991.

[40] {Hardy, Q., Jacob's Patter, The Wall Street Journal, Friday, Sep. 6, 1996.}

[41] {Hattori, T., PHS system deployment success story, Stanford Univ. Center for Telecom. Symposium, May 21, 1997.}

[42] Hinderling, J., et al., CDMA mobile station modem ASIC. *IEEE CICC 92,* Boston, MA, May. *Intl. J. Wireless Info. Net.,* 1(1), 37–48, Jan. 1994.

[43] Jabbari, B., et al., Network issues for wireless personal communications. *IEEE Comm. Mag.,* 88–98, Jan. 1995.

Ford Motor Co.
Exhibit 1030
Page 383

[44] Katz, R.H., Adaptation and mobility in wireless information systems. *IEEE Personal Comm.*, 1st Quarter, 6–17, 1994.

[45] {Kim, B.-J. and Cox, D.C., Blind equalization for short burst communications over frequency selective wireless channels, 1997 *IEEE VTC'97*, Phoenix, AZ, 2, 544–548, May 4–7, 1997a.}

[46] {Kim, B.-J. and Cox, D.C., Blind sequence estimation for short burst communications over time-varying wireless channels, *IEEE ICUPC'97*, San Diego, CA, 2, 713–717, Oct. 12–16, 1997b.}

[47] {Kim, B.-J. and Cox, D.C., Blind diversity combining equalization for short burst wireless communications, *IEEE Globecom'97*, Phoenix, AZ, 3, 1163–1167, Nov. 3–8, 1997c.}

[48] {Kim, B.-J., Blind equalization for short burst wireless communications, Stanford University Ph.D. Thesis, Jan. 1998.}

[49] {Korea Mobile Telecom, Article on CDMA, Mobile Communications International, Feb. 1997.}

[50] Padgett, J.E., Hattori, T., and Gunther, C., Overview of wireless personal communications. *IEEE Comm. Mag.*, 28–41, Jan. 1995.

[51] {Patel, V., Personal Access Communications System, Stanford Univ. Center for Telecom. Symposium, May 21, 1997.}

[52] {Phillips, B.W., Broadband in the Local Loop, Telecommunications, 37–42, Nov. 1994.}

[53] {PHS, World Congress, Singapore, Nov. 18–20, 1996.}

[54] Qualcomm Technology Forum, Open meeting on status of CDMA technology and review of San Diego experiment. San Diego, CA, Jan. 16–17, 1992.

[55] {Ranade, A., Local access radio interference due to building reflections, *IEEE Transaction on Communications*, 70–74, Jan. 1989.}

[56] {Rizzo, J.F. and Sollenberger, N.R., Multiuser Wireless Access, *IEEE Personal Communications Magazine*, 18–30, Jun. 1995.}

[57] Salmasi, A. and Gilhousen, K.S., On the system design aspects of code division multiple access (CDMA) applied to digital cellular and personal communications networks. *IEEE VTC'91*, St. Louis, MO, 57–62, May 1991.

[58] Sato, Y., A method of self-recovering equalization for multilevel amplitude modulation systems. *IEEE Trans. on Comm.*, 679–682, Jun. 1975.

[59] Schneideman, R., Spread spectrum gains wireless applications. *Microwaves and RF*, 31–42, May 1992.

[60] {Seidel, S.Y. and Arnold, H.W., Propagation measurements of 28 GHz to investigate the performance of local multipoint distribution services (LMDS), *IEEE Globecom'95*, 1995.}

[61] Sollenberger, N.R. and Chuang, J.C-I., Low overhead symbol timing and carrier recovery for TDMA portable radio systems. Third nordic sem. digital land mobile radio comm., Copenhagen, Denmark, 13–15, Sep. 1988.

[62] Sollenberger, N.R., An experimental VLSI implementation of low-overhead symbol timing and frequency offset estimation for TDMA portable radio applications. *IEEE Globecom'90*, San Diego, CA, 1701–1711, Dec. 1990.

[63] Sollenberger, N.R., An experimental TDMA modulation/demodulation CMOS VLSI chip-set. *IEEE CICC'91*, San Diego, CA, May 12–15, 1991.

[64] Sollenberger, N.R. and Afrashteh, A., An experimental low-delay TDMA portable radio link. *Wireless'91*, Calgary, Canada, Jul. 8–10, 1991.

[65] Sollenberger, N.R. and Chuang, J.C.-I., Low-overhead symbol timing and carrier recovery for TDMA portable radio systems. *IEEE Trans. on Comm.*, 1886–1892, Oct. 1990.

[66] Sollenberger, N.R., et al., Architecture and implementation of an efficient and robust TDMA frame structure for digital portable communications. *IEEE Trans. Veh. Tech.*, 40, 250–260, Feb. 1991.

[67] Steele, R., Deploying personal communications networks. *IEEE Comm. Mag.*, 12–15, Sep. 1990.

Ford Motor Co.
Exhibit 1030
Page 384

[68] Steer, D.G., Coexistence and access etiquette in the United States unlicensed PCS band. *IEEE Personal Comm. Mag.*, 4th Quarter, 36–43, 1994.

[69] {Titch, S., Blind Faith, Telephony, Sep. 8, 1997 and Global Telephony, Intertec Publishing Corp., Oct. 30, 1997.}

[70] Treichler, J.R. and Agee, B.G., A new approach to multipath correction of constant modulus signals, *IEEE Trans. on Acoustics, Speech and Signal Processing*, 459–472, Apr. 1983.

[71] Tuthill, J.P., Granger, B.S., and Wurtz, J.L., Request for a pioneer's preference. Before the Federal Communications Commission, Pacific Bell submission for FCC General Docket No. 90–314, RM-7140, and RM-7175, May 4, 1992.

[72] Varma, V.K., et al., Performance of sub-band and RPE coders in the portable communication environment. *Fourth Intl. Conf. on Land Mobile Radio*, Coventry, UK, 221–227, Dec. 14–17, 1987.

[73] Viterbi, A.J., When not to spread spectrum—A sequel. *IEEE Comm. Mag.*, 12–17, Apr. 1985.

[74] {Weseloh, C., Solutions for Wireless Communications, Stanford University Center for Telecommunications, Symposium, Stanford, CA, May 16, 1995.}

[75] Wong, A., Regulating public wireless networks. Workshop on lightwave, wireless and networking technologies, Chinese Univ. of Hong Kong, Hong Kong, Aug. 24, 1994.

[76] {Wong, P., Low-power, low-complexity diversity combining algorithm and VLSI circuit chip for hand-held PCS receivers, Stanford University Ph.D. Thesis, Apr. 1997.}

[77] {Wong, P. and Cox, D.C., Low-complexity cochannel interference cancellation and macroscopic diversity for high capacity PCS, *IEEE ICC'95 Digest*, Seattle, WA, 852–857, Jun. 19–22, 1995.}

[78] {Wong, P. and Cox, D.C., Low-complexity diversity combining algorithms and circuit architectures for cochannel interference cancellation and frequency selective fading mitigation, *IEEE ICC'96 Digest*, Dallas TX, Jun. 23–27, 1996a.}

[79] {Wong, P. and Cox, D.C., Low-complexity diversity combining algorithms and circuit architectures for cochannel interference cancellation and frequency-selective fading mitigation, *IEEE Trans. on Comm.*, 44(9), 1107–1116, Sep. 1996b.}

[80] {Wong, P. and Cox, D.C., Low-complexity interference cancellation and macroscopic diversity for high capacity PCS, *IEEE Trans. on Veh. Tech.*, 124–132, Feb. 1998.}

[81] Zaid, M., Personal mobility in PCS. *IEEE Personal Comm.*, 4th Quarter, 12–16, 1994.

[82] Special issue on advanced mobile phone service (AMPS). *Bell System Tech. J.*, (BSTS). 58, Jan. 1979.

[83] CDMA capacity seen as less than advertised. *Adv. Wireless Comm.*, (AWC). 1–2, Feb. 6, 1991.

[84] Bellcore's Wireless Prototype. *Microcell News*, (MN). 5–6, Jan. 25, 1992.

[85] Bellcore PCS phone shines in Bellsouth test; others fall short. *Adv. Wireless Comm.*, (AWC). 2–3, Feb. 19, 1992a.

[86] TDMA accelerates, but CDMA could still be second standard. *Adv. Wireless Comm.*, (AWC). 6, Mar. 4, 1992b.

[87] Special issue on wireless personal communications. *IEEE Comm. Mag.*, Jan. 1995.

Ford Motor Co.
Exhibit 1030
Page 385

16

Modulation Methods

16.1 Introduction 16-1
16.2 Basic Description of Modulated Signals 16-2
16.3 Analog Frequency Modulation 16-2
16.4 Phase Shift Keying (PSK) and $\pi/4$-QPSK 16-3
16.5 Continuous Phase Modulation (CPM) and MSK 16-7
16.6 Gaussian Minimum Shift Keying 16-9
16.7 Orthogonal Frequency Division Multiplexing (OFDM) 16-11
16.8 Conclusions 16-13
Defining Terms 16-13
References 16-14
Further Information 16-14

Gordon L. Stüber
Georgia Institute of Technology

16.1 Introduction

Modulation is the process where the message information is added to the radio carrier. Most first generation cellular systems such as the advanced mobile telephone system (AMPS) use analog **frequency modulation** (**FM**), because analog technology was very mature when these systems were first introduced. Digital modulation schemes, however, are the obvious choice for future wireless systems, especially if data services such as wireless multimedia are to be supported. Digital modulation can also improve spectral efficiency, because digital signals are more robust against channel impairments. Spectral efficiency is a key attribute of wireless systems that must operate in a crowded radio frequency spectrum.

To achieve high spectral efficiency, modulation schemes must be selected that have a high **bandwidth efficiency** as measured in units of bits per second per Hertz of bandwidth. Many wireless communication systems, such as cellular telephones, operate on the principle of frequency reuse, where the carrier frequencies are reused at geographically separated locations. The link quality in these systems is limited by cochannel interference. Hence, modulation schemes must be identified that are both bandwidth efficient and capable of tolerating high levels of cochannel interference. More specifically, digital modulation techniques are chosen for wireless systems that satisfy the following properties.

Compact Power Density Spectrum: To minimize the effect of adjacent channel interference, it is desirable that the power radiated into the adjacent channel be 60–80 dB below that in the desired channel. Hence, modulation techniques with a narrow main lobe and fast rolloff of sidelobes are desirable.

0-8493-8597-0/99/$0.00+$.50

© 1999 by CRC Press LLC

Ford Motor Co.
Exhibit 1030
Page 386

Good Bit-Error-Rate Performance: A low-bit-error probability should be achieved in the presence of cochannel interference, adjacent channel interference, thermal noise, and other channel impairments, such as fading and intersymbol interference.

Envelope Properties: Portable and mobile applications typically employ nonlinear (class C) power amplifiers to minimize battery drain. Nonlinear amplification may degrade the bit-error-rate performance of modulation schemes that transmit information in the amplitude of the carrier. Also, spectral shaping is usually performed prior to up-conversion and nonlinear amplification. To prevent the regrowth of spectral sidelobes during nonlinear amplification, the input signal must have a relatively constant envelope.

A variety of digital modulation techniques are currently being used in wireless communication systems. Two of the more widely used digital modulation techniques for cellular mobile radio are $\pi/4$ phase-shifted quadrature **phase shift keying** ($\pi/4$-QPSK) and **Gaussian minimum shift keying** (**GMSK**). The former is used in the North American IS-54 digital cellular system and Japanese Personal Digital Cellular (PDC), whereas the latter is used in the global system for mobile communications (GSM system). This chapter provides a discussion of these and other modulation techniques that are employed in wireless communication systems.

16.2 Basic Description of Modulated Signals

With any modulation technique, the bandpass signal can be expressed in the form

$$s(t) = \mathrm{Re}\left\{v(t)e^{j2\pi f_c t}\right\} \tag{16.1}$$

where $v(t)$ is the complex envelope, f_c is the carrier frequency, and $\mathrm{Re}\{z\}$ denotes the real part of z. For digital modulation schemes $v(t)$ has the general form

$$v(t) = A\sum_k b\left(t - kT, \boldsymbol{x}_k\right) \tag{16.2}$$

where A is the amplitude of the carrier $\boldsymbol{x}_k = (x_k, x_{k-1}, \ldots, x_{k-K})$ is the data sequence, T is the symbol or baud duration, and $b(t, \boldsymbol{x}_i)$ is an equivalent shaping function usually of duration T. The precise form of $b(t, \boldsymbol{x}_i)$ and the memory length K depends on the type of modulation that is employed. Several examples are provided in this chapter where information is transmitted in the amplitude, phase, or frequency of the bandpass signal.

The **power spectral density** of the bandpass signal $S_{ss}(f)$ is related to the power spectral density of the complex envelope $S_{vv}(f)$ by

$$S_{ss}(f) = \frac{1}{2}\left[S_{vv}\left(f - f_c\right) + S_{vv}\left(f + f_c\right)\right] \tag{16.3}$$

The power density spectrum of the complex envelope for a digital modulation scheme has the general form

$$S_{vv}(f) = \frac{A^2}{T}\sum_m S_{b,m}(f)e^{-j2\pi fmT} \tag{16.4}$$

where

$$S_{b,m}(f) = \frac{1}{2}E\left[B\left(f, \boldsymbol{x}_m\right)B^*\left(f, \boldsymbol{x}_0\right)\right] \tag{16.5}$$

$B(f, \boldsymbol{x}_m)$ is the Fourier transform of $b(t, \boldsymbol{x}_m)$, and $E[\cdot]$ denotes the expectation operator. Usually symmetric signal sets are chosen so that the complex envelope has zero mean, i.e.,

Ford Motor Co.
Exhibit 1030
Page 387

$E[b(t, \boldsymbol{x}_0)] = 0$. This implies that the power density spectrum has no discrete components. If, in addition, $\boldsymbol{x}_m$ and $\boldsymbol{x}_0$ are independent for $|m| > K$, then

$$S_{vv}(f) = \frac{A^2}{T} \sum_{|m|<K} S_{b,m}(f) e^{-j2\pi fmT} \tag{16.6}$$

16.3 Analog Frequency Modulation

With analog frequency modulation the complex envelope is

$$v(t) = A \exp\left[j2\pi k_f \int_0^t m(\tau)\, \mathrm{d}\tau\right] \tag{16.7}$$

where $m(t)$ is the modulating waveform and k_f in Hz/v is the frequency sensitivity of the FM modulator. The bandpass signal is

$$s(t) = A \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(t)\, \mathrm{d}t\right] . \tag{16.8}$$

The instantaneous frequency of the carrier $f_i(t) = f_c + k_f m(t)$ varies linearly with the waveform $m(t)$, hence, the name frequency modulation. Notice that FM has a constant envelope making it suitable for nonlinear amplification. However, the complex envelope is a nonlinear function of the modulating waveform $m(t)$ and, therefore, the spectral characteristics of $v(t)$ cannot be obtained directly from the spectral characteristics of $m(t)$.

With the sinusoidal modulating waveform $m(t) = A_m \cos(2\pi f_m t)$ the instantaneous carrier frequency is

$$f_i(t) = f_c + \Delta_f \cos(2\pi f_m t) \tag{16.9}$$

where $\Delta_f = k_f A_m$ is the peak frequency deviation. The complex envelope becomes

$$\begin{aligned} v(t) &= \exp\left[2\pi \int_0^t f_i(t)\, \mathrm{d}t\right] \\ &= \exp\left[2\pi f_c t + \beta \sin(2\pi f_m t)\right] \end{aligned} \tag{16.10}$$

where $\beta = \Delta_f / f_m$ is called the modulation index. The bandwidth of $v(t)$ depends on the value of β. If $\beta < 1$, then narrowband FM is generated, where the spectral widths of $v(t)$ and $m(t)$ are about the same, i.e., $2f_m$. If $\beta \gg 1$, then wideband FM is generated, where the spectral occupancy of $v(t)$ is slightly greater than $2\Delta_f$. In general, the approximate bandwidth of an FM signal is

$$W \approx 2\Delta_f + 2f_m = 2\Delta_f \left(1 + \frac{1}{\beta}\right) \tag{16.11}$$

which is a relation known as Carson's rule. Unfortunately, typical analog cellular radio systems use a modulation index in the range $1 \lesssim \beta \lesssim 3$ where Carson's rule is not accurate. Furthermore, the message waveform $m(t)$ is not a pure sinusoid so that Carson's rule does not directly apply.

In analog cellular systems the waveform $m(t)$ is obtained by first companding the speech waveform and then hard limiting the resulting signal. The purpose of the limiter is to control the peak frequency deviation Δ_f. The limiter introduces high-frequency components that

Ford Motor Co.
Exhibit 1030
Page 388

must be removed with a low-pass filter prior to modulation. To estimate the bandwidth occupancy, we first determine the ratio of the frequency deviation Δ_f corresponding to the maximum amplitude of $m(t)$, and the highest frequency component B that is present in $m(t)$. These two conditions are the most extreme cases, and the resulting ratio, $D = \Delta_f/B$, is called the *deviation ratio*. Then replace β by D and f_m by B in Carson's rule, giving

$$W \approx 2\Delta_f + 2B = 2\Delta_f \left(1 + \frac{1}{D}\right) \tag{16.12}$$

This approximation will overestimate the bandwidth requirements. A more accurate estimate of the bandwidth requirements must be obtained from simulation or measurements.

16.4 Phase Shift Keying (PSK) and $\pi/4$-QPSK

With **phase shift keying** (PSK), the equivalent shaping function in Eq. (16.2) has the form

$$b(t, \boldsymbol{x}_k) = \psi_T(t)\exp\left[j\frac{\pi}{M}x_k h_s(t)\right], \qquad \boldsymbol{x}_k = x_k \tag{16.13}$$

where $h_s(t)$ is a phase shaping pulse, $\psi_T(t)$ an amplitude shaping pulse, and M the size of the modulation alphabet. Notice that the phase varies linearly with the symbol sequence $\{x_k\}$, hence the name phase shift keying. For a modulation alphabet size of M, $x_k \in \{\pm 1, \pm 3, \ldots, \pm(M-1)\}$. Each symbol x_k is mapped onto $\log_2 M$ source bits. A QPSK signal is obtained by using $M = 4$, resulting in a transmission rate of 2 b/symbol.

Usually, the phase shaping pulse is chosen to be the rectangular pulse $h_s(t) = u_T(t) \triangleq u(t) - u(t-T)$, where $u(t)$ is the unit step function. The amplitude shaping pulse is very often chosen to be a square root raised cosine pulse, where the Fourier transform of $\psi_T(t)$ is

$$\Psi_T(f) = \begin{cases} \sqrt{T} & 0 \le |f| \le (1-\beta)/2T \\ \sqrt{\dfrac{T}{2}\left[1 - \sin\dfrac{\pi T}{\beta}\left(f - \dfrac{1}{2T}\right)\right]} & (1-\beta)/2T \le |f| \le (1+\beta)/2T \end{cases} \tag{16.14}$$

The receiver implements the same filter $\Psi_R(f) = \Psi_T(f)$ so that the overall pulse has the raised cosine spectrum $\Psi(f) = \Psi_R(f)\Psi_T(f) = |\Psi_T(f)|^2$. If the channel is affected by flat fading and additive white Gaussian noise, then this partitioning of the filtering operations between the transmitter and receiver will optimize the signal to noise ratio at the output of the receiver filter at the sampling instants. The rolloff factor β usually lies between 0 and 1 and defines the **excess bandwidth** $100\beta\%$. Using a smaller β results in a more compact power density spectrum, but the link performance becomes more sensitive to errors in the symbol timing. The IS-54 system uses $\beta = 0.35$, while PDC uses $\beta = 0.5$.

The time domain pulse corresponding to Eq. (16.14) can be obtained by taking the inverse Fourier transform, resulting in

$$\psi_T(t) = 4\beta\frac{\cos\left[(1+\beta)\,\pi t/T\right] + \sin\left[(1-\beta)\,\pi t/T\right](4\beta t/T)^{-1}}{\pi\sqrt{T}\left[1 - 16\beta^2 t^2/T^2\right]} \tag{16.15}$$

A typical square root raised cosine pulse with a rolloff factor of $\beta = 0.5$ is shown in Fig. 16.1. Strictly speaking the pulse $\psi_T(t)$ is noncausal, but in practice a truncated time domain pulse is used. For example, in Fig. 16.1 the pulse is truncated to $6T$ and time shifted by $3T$ to yield a causal pulse.

Ford Motor Co.
Exhibit 1030
Page 389

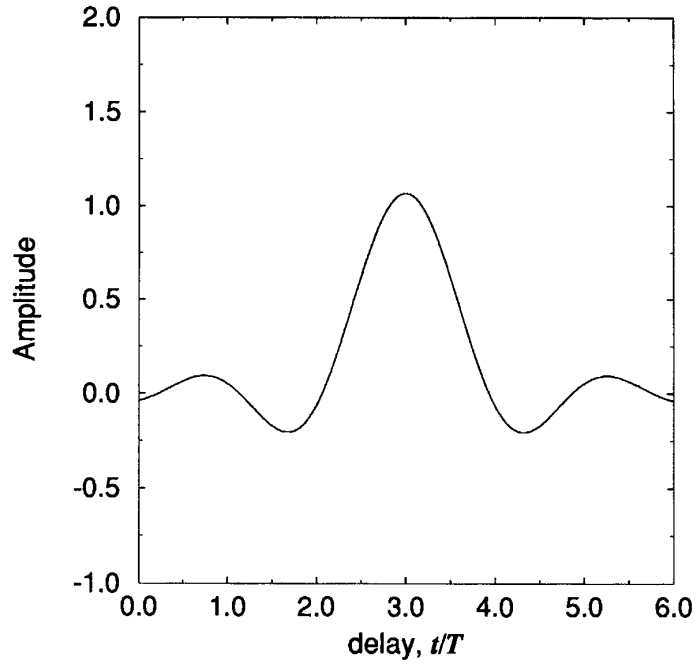


FIGURE 16.1 Square root raised cosine pulse with rolloff factor $\beta = 0.5$.

Unlike conventional QPSK that has four possible transmitted phases, $\pi/4$-QPSK has eight possible transmitted phases. Let $\theta(n)$ be the transmitted carrier phase for the nth epoch, and let $\Delta\theta(n) = \theta(n) - \theta(n-1)$ be the differential carrier phase between epochs n and $n-1$. With $\pi/4$-QPSK, the transmission rate is 2 b/symbol and the differential phase is related to the symbol sequence $\{x_n\}$ through the mapping

$$\Delta\theta(n) = \begin{cases} -3\pi/4, & x_n = -3 \\ -\pi/4, & x_n = -1 \\ \pi/4, & x_n = +1 \\ 3\pi/4, & x_n = +3 \end{cases} \tag{16.16}$$

Since the symbol sequence $\{x_n\}$ is random, the mapping in Eq. (16.16) is arbitrary, except that the phase differences must be $\pm\pi/4$ and $\pm 3\pi/4$. The phase difference with the given mapping can be written in the convenient algebraic form

$$\Delta\theta(n) = x_n \frac{\pi}{4} \tag{16.17}$$

which allows us to write the equivalent shaping function of the $\pi/4$-QPSK signal as

$$\begin{aligned} b\,(t, \underline{\boldsymbol{x}}_k) &= \psi(t)\exp\left\{j\left[\theta(k-1) + x_k\frac{\pi}{4}\right]\right\} \\ &= \psi_T(t)\exp\left[j\frac{\pi}{4}\left(\sum_{n=-\infty}^{k-1} x_n + x_k\right)\right] \end{aligned} \tag{16.18}$$

The summation in the exponent represents the accumulated carrier phase, whereas the last term is the phase change due to the kth symbol. Observe that the phase shaping function

Ford Motor Co.
Exhibit 1030
Page 390

is the rectangular pulse $u_T(t)$. The amplitude shaping function $\psi_T(t)$ is usually the square root raised cosine pulse in Eq. (16.15).

The phase states of QPSK and $\pi/4$-QPSK signals can be summarized by the signal space diagram in Fig. 16.2 that shows the phase states and allowable transitions between the phase states. However, it does not describe the actual phase trajectories. A typical diagram showing phase trajectories with square root raised cosine pulse shaping is shown in Fig. 16.3. Note that the phase trajectories do not pass through the origin. This reduces the envelope fluctuations of the signal making it less susceptible to amplifier nonlinearities and reduces the dynamic range required of the power amplifier.

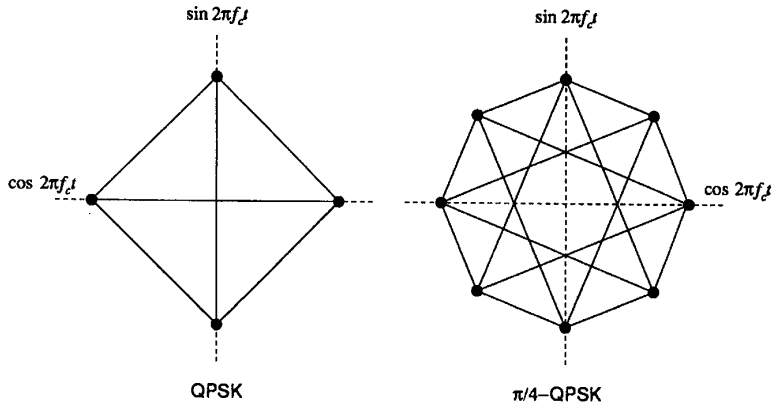


FIGURE 16.2 Signal-space constellations for QPSK and $\pi/4$-DQPSK.

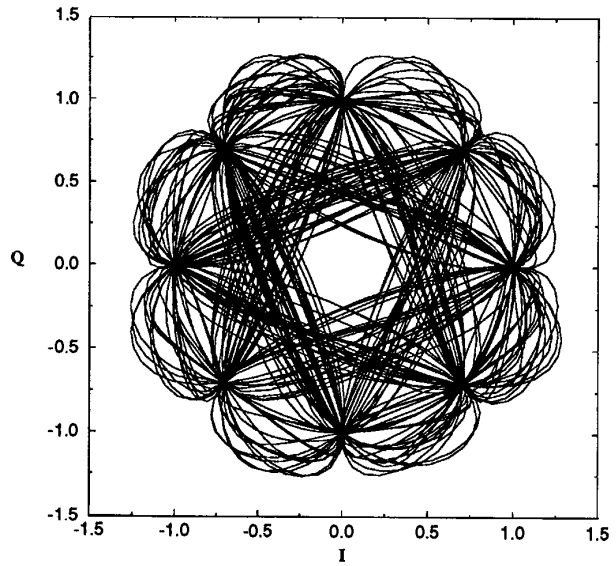


FIGURE 16.3 Phase diagram of $\pi/4$-QPSK with square root raised cosine pulse; $\beta = 0.5$.

Ford Motor Co.
Exhibit 1030
Page 391

The power density spectrum of QPSK and $\pi/4$-QPSK depends on both the amplitude and phase shaping pulses. For the rectangular phase shaping pulse $h_s(t) = u_T(t)$, the power density spectrum of the complex envelope is

$$S_{vv}(f) = \frac{A^2}{T} |\Psi_T(f)|^2 \tag{16.19}$$

With square root raised cosine pulse shaping, $\Psi_T(f)$ has the form defined in Eq. (16.14). The power density spectrum of a pulse $\tilde{\psi}_T(t)$ that is obtained by truncating $\psi_T(t)$ to length τ can be obtained by writing $\tilde{\psi}_T(t) = \psi_T(t)\text{rect}(t/\tau)$. Then $\tilde{\Psi}_T(f) = \Psi_T(f) * \tau\text{sinc}(f\tau)$, where $*$ denotes the operation of convolution, and the power density spectrum is again obtained by applying Eq. (16.19). Truncation of the pulse will regenerate some side lobes, thus causing adjacent channel interference. Figure 16.4 illustrates the power density spectrum of a truncated square root raised cosine pulse for various truncation lengths τ.

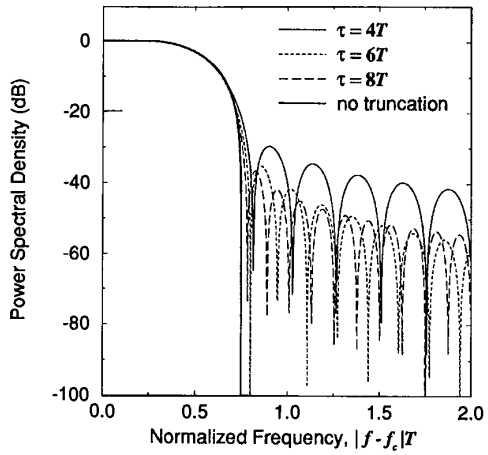


FIGURE 16.4 Power density spectrum of truncated square root raised cosine pulse with various truncation lengths; $\beta = 0.5$.

16.5 Continuous Phase Modulation (CPM) and MSK

Continuous phase modulation (CPM) refers to a broad class of frequency modulation techniques where the carrier phase varies in a continuous manner. A comprehensive treatment of CPM is provided in [1]. CPM schemes are attractive because they have constant envelope and excellent spectral characteristics. The complex envelope of any CPM signal is

$$v(t) = A \exp\left[j2\pi k_f \int_{-\infty}^{t} \sum_n x_n h_s(\tau - nT)\, \mathrm{d}\tau \right] \tag{16.20}$$

The instantaneous frequency deviation from the carrier is

$$f_{\text{dev}}(t) = k_f \sum_n x_n h_s(t - nT) \tag{16.21}$$

where k_f is the peak frequency deviation. If the frequency shaping pulse $h_s(t)$ has duration T, then the equivalent shaping function in Eq. (16.2) has the form

$$b(t, \boldsymbol{x}_k) = \exp\left\{ j \left[\beta(T) \sum_{n=-\infty}^{k-1} x_n + x_k \beta(t) \right] \right\} u_T(t) \tag{16.22}$$

where

$$\beta(t) = \begin{cases} 0, & t < 0 \\ \dfrac{\pi h}{\int_0^T h_s(\tau)\,\mathrm{d}\tau} \displaystyle\int_0^t h_s(\tau)\,\mathrm{d}\tau, & 0 \le t \le T \\ \pi h, & t \ge T \end{cases} \tag{16.23}$$

is the phase shaping pulse, and $h = \beta(T)/\pi$ is called the modulation index.

Minimum shift keying (MSK) is a special form of binary CPM ($x_k \in \{-1, +1\}$) that is defined by a rectangular frequency shaping pulse $h_s(t) = u_T(t)$, and a modulation index $h = 1/2$ so that

$$\beta(t) = \begin{cases} 0, & t < 0 \\ \pi t/2T, & 0 \le t \le T \\ \pi/2, & t \ge T \end{cases} \tag{16.24}$$

Therefore, the complex envelope is

$$v(t) = A \exp\left(j \frac{\pi}{2} \sum_{n=-\infty}^{k-1} x_n + \frac{\pi}{2} x_k \frac{t - kT}{T} \right) \tag{16.25}$$

A MSK signal can be described by the phase trellis diagram shown in Fig. 16.5 which plots the time behavior of the phase

$$\theta(t) = \frac{\pi}{2} \sum_{n=-\infty}^{k-1} x_n + \frac{\pi}{2} x_k \frac{t - kT}{T} \tag{16.26}$$

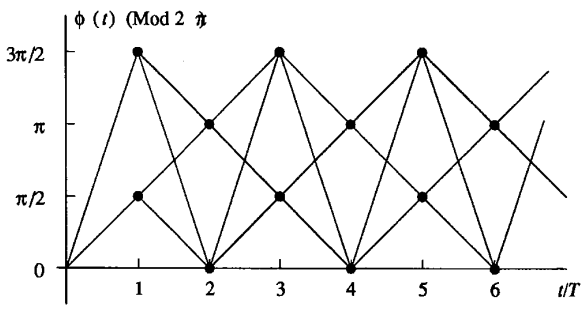


FIGURE 16.5 Phase-trellis diagram for MSK.

The MSK bandpass signal is

$$\begin{aligned} s(t) &= A \cos\left(2\pi f_c t + \frac{\pi}{2} \sum_{n=-\infty}^{k-1} x_n + \frac{\pi}{2} x_k \frac{t - kT}{T} \right) \\ &= A \cos\left[2\pi \left(f_c + \frac{x_k}{4T} \right) t - \frac{k\pi}{2} x_k + \frac{\pi}{2} \sum_{n=-\infty}^{k-1} x_n \right] \quad kT \le t \le (k+1)T \end{aligned} \tag{16.27}$$

From Eq. (16.27) we observe that the MSK signal has one of two possible frequencies $f_L = f_c - 1/4T$ or $f_U = f_c + 1/4T$ during each symbol interval. The difference between these frequencies is $f_U - f_L = 1/2T$. This is the minimum frequency difference between two sinusoids of duration T that will ensure orthogonality with coherent demodulation [7], hence, the name minimum shift keying. By applying various trigonometric identities to Eq. (16.27) we can write

$$s(t) = A\left[x_k^I \psi(t-k2T)\cos(2\pi f_c t) - x_k^Q \psi(t-k2T-T)\sin(2\pi f_c t)\right], \quad kT \le t \le (k+1)T \tag{16.28}$$

where

$$\begin{aligned} x_k^I &= -x_{k-1}^Q x_{2k-1} \\ x_k^Q &= x_k^I x_{2k} \\ \psi(t) &= \cos\left(\frac{\pi t}{2T}\right), \qquad -T \le t \le T \end{aligned}$$

Note that the x_k^I and x_k^Q are independent binary symbols that take on elements from the set $\{-1,+1\}$, and the half-sinusoid amplitude shaping pulse $\psi(t)$ has duration $2T$ and $\psi(t-T) = \sin(\pi t/2T), 0 \le t \le 2T$. Therefore, MSK is equivalent to offset quadrature amplitude shift keying (OQASK) with a half-sinusoid amplitude shaping pulse.

To obtain the power density spectrum of MSK, we observe from Eq. (16.28) that the equivalent shaping function of MSK has the form

$$b(t, \boldsymbol{x}_k) = x_k^I \psi(t) + j x_k^Q \psi(t-T) \tag{16.29}$$

The Fourier transform of Eq. (16.29) is

$$B(f, \boldsymbol{x}_k) = \left(x_k^I + j x_k^Q e^{-j2\pi fT}\right)\Psi(f) \tag{16.30}$$

Since the symbols x_k^I and x_k^Q are independent and zero mean, it follows from Eqs. (16.5) and (16.6) that

$$S_{vv}(f) = \frac{A^2|\Psi(f)|^2}{2T} \tag{16.31}$$

Therefore, the power density spectrum of MSK is determined solely by the Fourier transform of the half-sinusoid amplitude shaping pulse $\psi(t)$, resulting in

$$S_{vv}(f) = \frac{16A^2T}{\pi^2}\left[\frac{\cos 2\pi fT}{1-16f^2T^2}\right]^2 \tag{16.32}$$

The power spectral density of MSK is plotted in Fig. 16.8. Observe that an MSK signal has fairly large sidelobes compared to $\pi/4$-QPSK with a truncated square root raised cosine pulse (c.f., Fig. 16.4).

16.6 Gaussian Minimum Shift Keying

MSK signals have all of the desirable attributes for mobile radio, except for a compact power density spectrum. This can be alleviated by filtering the modulating signal $x(t) = \Sigma_n x_n u_T(t-nT)$ with a low-pass filter prior to frequency modulation, as shown in Fig. 16.6.

Ford Motor Co.
Exhibit 1030
Page 394

Such filtering removes the higher frequency components in $x(t)$ and, therefore, yields a more compact spectrum. The low-pass filter is chosen to have 1) narrow bandwidth and a sharp transition band, 2) low-overshoot impulse response, and 3) preservation of the output pulse area to ensure a phase shift of $\pi/2$.

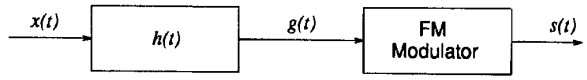


FIGURE 16.6 Premodulation filtered MSK.

GMSK uses a low-pass filter with the following transfer function:

$$H(f) = A \exp\left\{-\left(\frac{f}{B}\right)^2 \frac{\ln 2}{2}\right\} \tag{16.33}$$

where B is the 3-dB bandwidth of the filter and A a constant. It is apparent that $H(f)$ is bell shaped about $f = 0$, hence the name Gaussian MSK. A rectangular pulse $\text{rect}(t/T) = u_T(t + T/2)$ transmitted through this filter yields the frequency shaping pulse

$$h_s(t) = A\sqrt{\frac{2\pi}{\ln 2}}(BT) \int_{t/T-1/2}^{t/T+1/2} \exp\left\{-\frac{2\pi^2(BT)^2x^2}{\ln 2}\right\} dx \tag{16.34}$$

The phase change over the time interval from $-T/2 \leq t \leq T/2$ is

$$\theta\left(\frac{T}{2}\right) - \theta\left(\frac{-T}{2}\right) = x_0\beta_0(T) + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} x_n\beta_n(T) \tag{16.35}$$

where

$$\beta_n(T) = \frac{\pi h}{\int_{-\infty}^{\infty} h_s(\nu)\, d\nu} \int_{-T/2-nT}^{T/2-nT} h_s(\nu)\, d\nu \tag{16.36}$$

The first term in Eq. (16.35) is the desired term, and the second term is the intersymbol interference (ISI) introduced by the premodulation filter. Once again, with GMSK $h = 1/2$ so that a total phase shift of $\pi/2$ is maintained.

Notice that the pulse $h_s(t)$ is noncausal so that a truncated pulse must be used in practice. Figure 16.7 plots a GMSK frequency shaping pulse that is truncated to $\tau = 5T$ and time shifted by $2.5T$, for various normalized filter bandwidths BT. Notice that the frequency shaping pulse has a duration greater than T so that ISI is introduced. As BT decreases, the induced ISI is increased. Thus, whereas a smaller value of BT results in a more compact power density spectrum, the induced ISI will degrade the bit-error-rate performance. Hence, there is a tradeoff in the choice of BT. Some studies have indicated that $BT = 0.25$ is a good choice for cellular radio systems [6].

The power density spectrum of GMSK is quite difficult to obtain, but can be computed by using published methods [3]. Figure 16.8 plots the power density spectrum for $BT = 0.2$, 0.25, and 0.3, obtained from Wesolowski, [8]. Observe that the spectral sidelobes are greatly reduced by the Gaussian low-pass filter.

Ford Motor Co.
Exhibit 1030
Page 395

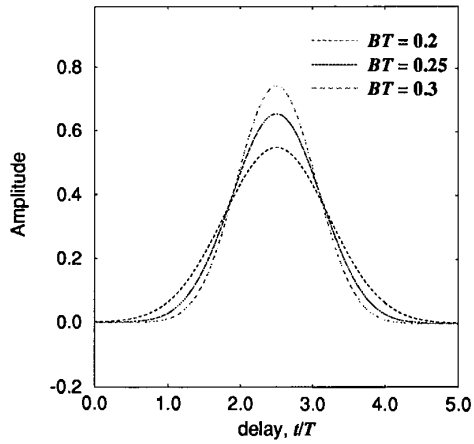


FIGURE 16.7 GMSK frequency shaping pulse for various normalized filter bandwidths BT.

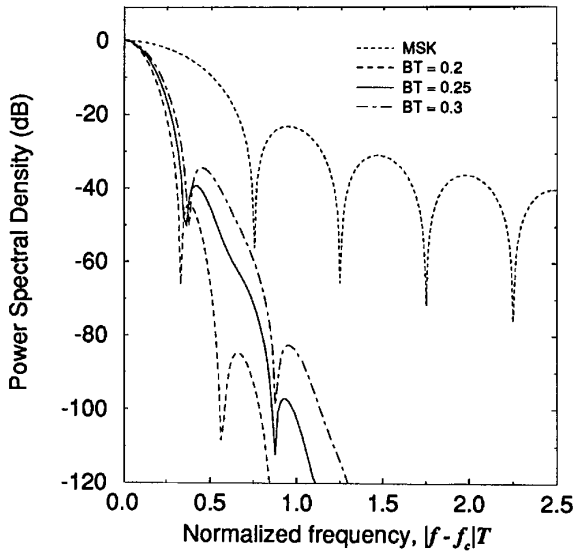


FIGURE 16.8 Power density spectrum of MSK and GMSK.

16.7 Orthogonal Frequency Division Multiplexing (OFDM)

Orthogonal frequency division multiplexing (OFDM) is a modulation technique that has been recently suggested for use in cellular radio [2], digital audio broadcasting [4], and digital video broadcasting. The basic idea of OFDM is to transmit blocks of symbols in parallel by employing a (large) number of orthogonal subcarriers. With block transmission, N serial source symbols each with period T_s are converted into a block of N parallel modulated symbols each with period $T = NT_s$. The block length N is chosen so that $NT_s \gg \sigma_\tau$, where σ_τ is the rms delay spread of the channel. Since the symbol rate on each subcarrier is much less than the serial source rate, the effects of delay spread are greatly reduced. This has practical advantages because it may reduce or even eliminate the need for equalization.

Ford Motor Co.
Exhibit 1030
Page 396

Although the block length N is chosen so that $NT_s \gg \sigma_\tau$, the channel dispersion will still cause consecutive blocks to overlap. This results in some residual ISI that will degrade the performance. This residual ISI can be eliminated at the expense of channel capacity by using guard intervals between the blocks that are at least as long as the effective channel impulse response.

The complex envelope of an OFDM signal is described by

$$v(t) = A \sum_k \sum_{n=0}^{N-1} x_{k,n} \phi_n(t - kT) \tag{16.37}$$

where

$$\phi_n(t) = \exp\left\{ j \frac{2\pi \left(n - \dfrac{N-1}{2} \right) t}{T} \right\} U_T(t), \qquad n = 0, 1, \ldots, N-1 \tag{16.38}$$

are orthogonal waveforms and $U_T(t)$ is a rectangular shaping function. The frequency separation of the subcarriers, $1/T$, ensures that the subcarriers are orthogonal and phase continuity is maintained from one symbol to the next, but is twice the minimum required for orthogonality with coherent detection. At epoch k, N-data symbols are transmitted by using the N distinct pulses. The data symbols $x_{k,n}$ are often chosen from an M-ary **quadrature amplitude modulation** (M-QAM) constellation, where $x_{k,n} = x_{k,n}^I + j x_{k,n}^Q$ with $x_{k,n}^I, x_{k,n}^Q \in \{\pm 1, \pm 3, \ldots, \pm(N-1)\}$ and $N = \sqrt{M}$.

A key advantage of using OFDM is that the modulation can be achieved in the discrete domain by using either an inverse discrete Fourier transform (IDFT) or the more computationally efficient inverse fast Fourier transform (IFFT). Considering the data block at epoch $k = 0$ and ignoring the frequency offset $\exp\{-j[2\pi(N-1)t/2T]\}$, the complex low-pass OFDM signal has the form

$$v(t) = \sum_{n=0}^{N-1} x_{0,n} \exp\left\{ \frac{j2\pi nt}{NT_s} \right\}, \qquad 0 \le t \le T \tag{16.39}$$

If this signal is sampled at epochs $t = kT_s$, then

$$v^k = v(kT_s) = \sum_{n=0}^{N-1} x_{0,n} \exp\left\{ \frac{j2\pi nk}{N} \right\}, \qquad k = 0, 1, \ldots, N-1 \tag{16.40}$$

Observe that the sampled OFDM signal has duration N and the samples $v^0, v^1, \ldots, v^{N-1}$ are just the IDFT of the data block $x_{0,0}, x_{0,1}, \ldots, x_{0,N-1}$. A block diagram of an OFDM transmitter is shown in Fig. 16.9.

The power spectral density of an OFDM signal can be obtained by treating OFDM as independent modulation on subcarriers that are separated in frequency by $1/T$. Because the subcarriers are only separated by $1/T$, significant spectral overlap results. Because the subcarriers are orthogonal, however, the overlap improves the spectral efficiency of the scheme. For a signal constellation with zero mean and the waveforms in Eq. (16.38), the power density spectrum of the complex envelope is

$$S_{vv}(f) = \frac{A^2}{T} \sigma_x^2 \sum_{n=0}^{N-1} \left| \operatorname{sinc}\left[fT - \left(n - \frac{N-1}{2} \right) \right] \right|^2 \tag{16.41}$$

where $\sigma_x^2 = \frac{1}{2} E[|x_{k,n}|^2]$ is the variance of the signal constellation. For example, the complex envelope power spectrum of OFDM with $N = 32$ subcarriers is shown in Fig. 16.10.

Ford Motor Co.
Exhibit 1030
Page 397

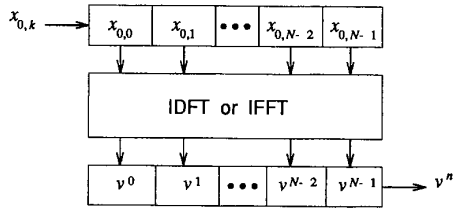


FIGURE 16.9 Block diagram of OFDM transmitter using IDFT or IFFT.

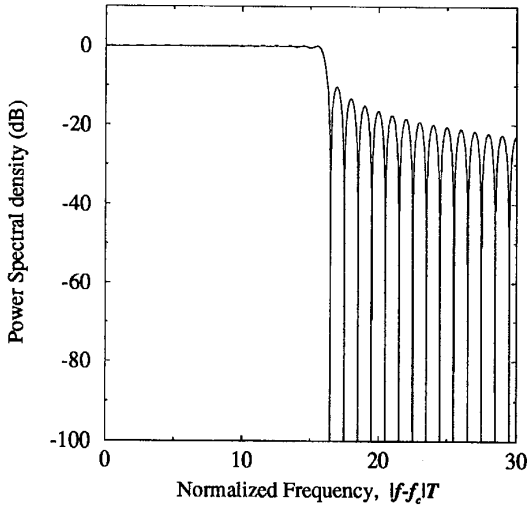


FIGURE 16.10 Power density spectrum of OFDM with $N = 32$.

16.8 Conclusions

A variety of modulation schemes are employed in wireless communication systems. Wireless modulation schemes must have a compact power density spectrum, while at the same time providing a good bit-error-rate performance in the presence of channel impairments such as cochannel interference and fading. The most popular digital modulation techniques employed in wireless systems are GMSK in the European GSM system, $\pi/4$-QPSK in the North American IS-54 and Japanese PDC systems, and OFDM in digital audio broadcasting systems.

Defining Terms

Bandwidth efficiency: Transmission efficiency of a digital modulation scheme measured in units of bits per second per Hertz of bandwidth.

Continuous phase modulation: Frequency modulation where the phase varies in a continuous manner.

Excess bandwidth: Percentage of bandwidth that is in excess of the minimum of $1/2T$ (T is the baud or symbol duration) required for data communication.

Frequency modulation: Modulation where the instantaneous frequency of the carrier varies linearly with the data signal.

Ford Motor Co.
Exhibit 1030
Page 398

Gaussian minimum shift keying: MSK where the data signal is prefiltered with a Gaussian filter prior to frequency modulation.

Minimum shift keying: A special form of continuous phase modulation having linear phase trajectories and a modulation index of 1/2.

Orthogonal frequency division multiplexing: Modulation by using a collection of low-bit-rate orthogonal subcarriers.

Phase shift keying: Modulation where the instantaneous phase of the carrier varies linearly with the data signal.

Power spectral density: Relative power in a modulated signal as a function of frequency.

Quadrature amplitude modulation: Modulation where information is transmitted in the amplitude of the cosine and sine components of the carrier.

References

[1] Anderson, J.B., Aulin, T., and Sundberg, C.-E., *Digital Phase Modulation,* Plenum Press, New York, 1986.

[2] Birchler, M.A. and Jasper, S.C., A 64 kbps digital land mobile radio system employing M-16QAM. *Proc. 5th Nordic Sem. Dig. Mobile Radio Commun.*, 237–241, Dec. 1992.

[3] Garrison, G.J., A power spectral density analysis for digital FM, *IEEE Trans. Commun.*, COM-23, 1228–1243, Nov. 1975.

[4] Le Floch, B., Halbert-Lassalle, R., and Castelain, D., Digital sound broadcasting to mobile receivers, *IEEE Trans. Consum. Elec.*, 35, Aug. 1989.

[5] Murota, K. and Hirade, K., GMSK modulation for digital mobile radio telephony, *IEEE Trans. Commun.*, COM-29, 1044–1050, Jul. 1981.

[6] Murota, K., Kinoshita, K., and Hirade, K., Spectral efficiency of GMSK land mobile radio. *Proc. ICC'81,* 23.8.1, Jun. 1981.

[7] Proakis, J.G., *Digital Communications,* 2nd ed., McGraw-Hill, New York, 1989.

[8] Wesolowski, K., Private Communication, 1994.

Further Information

A good discussion of digital modem techniques is presented in *Advanced Digital Communications,* edited by K. Feher, Prentice-Hall, 1987.

Proceedings of various IEEE conferences such as the Vehicular Technology Conference, International Conference on Communications, and Global Telecommunications Conference, document the lastest development in the field of wireless communications each year.

Journals such as the *IEEE Transactions on Communications* and *IEEE Transactions on Vehicular Technology* report advances in wireless modulation.

Ford Motor Co.
Exhibit 1030
Page 399