An Introduction to Digital Communication Theory:

Notes for EE230B

Michael P. Fitz
The University of California Los Angeles
fitz@ee.ucla.edu

Winter 2005

Note to Students. This text is an evolving entity. Please help make an UCLA education more valuable by providing me feedback on this work. Small things like catching typos or big things like highlighting sections that are not clear are both important.

My goal in teaching communications (and in authoring these notes) is to provide students with

1. the required theory,
2. an insight into the required tradeoffs between spectral efficiency, performance, and complexity that are required for a communication system design,
3. demonstration of the utility and applicability of the theory in the homework problems and projects,
4. a logical progression in thinking about communication theory.

Consequently this textbook will be more mathematical than most and does not discuss a host of examples of communication systems. Matlab is used extensively to illustrate the concepts of communication theory as it is a great visualization tool. To me the beauty of communication theory is the logical flow of ideas. I have tried to capture this progression in this text.

Contents

1 Introduction **9**
- 1.1 Historical Perspectives on Communication Theory . . . 9
- 1.2 Goals For This Text . . . 10
- 1.3 Modern Communication System Engineering . . . 11
- 1.4 Technology's Impact on this Text . . . 12
- 1.5 The Text Relation to Communication Curriculum . . . 13
- 1.6 Book Overview . . . 15
  - 1.6.1 Mathematical Foundations . . . 15
  - 1.6.2 Analog Communication . . . 15
  - 1.6.3 Noise in Communication Systems . . . 16
  - 1.6.4 Digital Basics . . . 17
  - 1.6.5 Advanced Topics in Digital Communications . . . 17

I Mathematical Foundations **19**

2 Complex Baseband Representation of Bandpass Signals **21**
- 2.1 Introduction . . . 21
- 2.2 Baseband Representation of Bandpass Signals . . . 22
- 2.3 Visualization of Complex Envelopes . . . 25
- 2.4 Spectral Characteristics of the Complex Envelope . . . 27
  - 2.4.1 Basics . . . 27
  - 2.4.2 Bandwidth of Bandpass Signals . . . 30
- 2.5 Linear Systems and Bandpass Signals . . . 31
- 2.6 Conclusions . . . 34
- 2.7 Homework Problems . . . 35
- 2.8 Example Solutions . . . 44
- 2.9 Mini-Projects . . . 46

3 Random Processes **49**
- 3.1 Basic Definitions . . . 50
- 3.2 Gaussian Random Processes . . . 51
- 3.3 Stationary Random Processes . . . 52
  - 3.3.1 Basics . . . 53
  - 3.3.2 Gaussian Processes . . . 53
  - 3.3.3 Frequency Domain Representation . . . 55
- 3.4 Thermal Noise . . . 57
- 3.5 Linear Systems and Random Processes . . . 58
- 3.6 The Solution of the Canonical Problem . . . 61

3.7 Homework Problems . . . 64
3.8 Example Solutions . . . 73
3.9 Mini-Projects . . . 73

4 Noise in Bandpass Communication Systems 75
4.1 Bandpass Random Processes . . . 78
4.2 Characteristics of the Complex Envelope . . . 80
4.2.1 Three Important Results . . . 80
4.2.2 Important Corollaries . . . 82
4.3 Spectral Characteristics . . . 84
4.4 The Solution of the Canonical Bandpass Problem . . . 86
4.5 Homework Problems . . . 88
4.6 Example Solutions . . . 93
4.7 Mini-Projects . . . 94

5 Bandpass Gaussian Random Processes 97
5.1 Preliminaries . . . 97
5.2 Amplitude and Phase of Noise . . . 99
5.3 The Amplitude of Signal Plus Noise . . . 99
5.4 The Phase of Signal Plus Noise . . . 101
5.5 Complex Gaussian Random Processes . . . 105
5.5.1 Complex Gaussian Distribution . . . 105
5.5.2 Complex Additive White Gaussian Noise . . . 105
5.6 Quadratic Forms of Complex Gaussians . . . 106
5.7 Homework Problems . . . 106

II Digital Communication Basics 109

6 Digital Communication Basics 111
6.1 Digital Transmission . . . 111
6.1.1 Digital Modulation . . . 112
6.1.2 Digital Demodulation . . . 112
6.2 Performance Metrics for Digital Communication . . . 113
6.2.1 Performance . . . 113
6.2.2 Complexity . . . 114
6.2.3 Bandwidth Efficiency . . . 114
6.2.4 Other Important Characteristics . . . 115
6.3 Some Limits on Performance of Digital Communication Systems . . . 115
6.4 Homework Problems . . . 116

7 Optimal Single Bit Demodulation Structures 119
7.1 Introduction . . . 119
7.1.1 Statistical Hypothesis Testing . . . 120
7.1.2 Statistical Hypothesis Testing in Digital Communications . . . 122
7.1.3 Digital Communications Design Problem . . . 122
7.2 Minimum Probability of Error Bit Demodulation . . . 123
7.2.1 Characterizing the Filter Output . . . 125
7.2.2 Uniform A Priori Probability . . . 126
7.3 Performance Analysis . . . 129

7.3.1 Erf Function . . . 131
7.3.2 Uniform A Priori Probability . . . 131
7.4 Filter Design . . . 133
7.4.1 Maximizing Effective SNR . . . 133
7.4.2 The Matched Filter . . . 134
7.4.3 MLBD with the Matched Filter . . . 135
7.4.4 More Insights on the Matched Filter . . . 137
7.5 Signal Design . . . 138
7.6 Spectral Characteristics . . . 140
7.7 Examples . . . 141
7.7.1 Frequency Shift Keying . . . 141
7.7.2 Phase Shift Keying . . . 145
7.7.3 Discussion . . . 149
7.8 Homework Problems . . . 149
7.9 Projects . . . 161
7.9.1 Project 1 . . . 161

8 Transmitting More Than One Bit 163
8.1 A Reformulation For One Bit Transmission . . . 163
8.2 Optimum Demodulation . . . 164
8.2.1 Optimum Word Demodulation Receivers . . . 165
8.2.2 Performance Analysis . . . 167
8.2.3 Union Bound . . . 170
8.2.4 Signal Design . . . 173
8.3 Optimum Bit Demodulation . . . 174
8.3.1 Optimum Bit Demodulation Receivers . . . 174
8.3.2 Comparison with Optimum Word Demodulation . . . 176
8.4 Examples . . . 177
8.4.1 M-ary FSK . . . 177
8.4.2 M-ary PSK . . . 181
8.4.3 Discussion . . . 186
8.5 Homework Problems . . . 186
8.6 Example Solutions . . . 195

9 Managing the Complexity of Optimum Demodulation 199
9.1 Linear Modulations . . . 199
9.1.1 MLWD for Linear Modulation . . . 200
9.1.2 Performance Evaluation for Linear Modulation . . . 201
9.1.3 Spectral Characteristics of Linear Modulation . . . 206
9.1.4 Summary of Linear Modulation . . . 206
9.2 Low Complexity Optimal Bit Decisions . . . 206
9.2.1 Optimum Bit Demodulation . . . 207
9.2.2 Orthogonal Modulations . . . 208
9.3 Orthogonal Modulation Examples . . . 210
9.3.1 Orthogonal Frequency Division Multiplexing . . . 210
9.3.2 Orthogonal Code Division Multiplexing . . . 214
9.3.3 Binary Stream Modulation . . . 218
9.4 Conclusions . . . 221
9.5 Homework Problems . . . 222

10 Spectrally Efficient Data Transmission **231**
10.1 Spectral Containment . . . 231
10.2 Squared Cosine Pulse . . . 232
10.3 Spectral Shaping in OFDM . . . 234
10.4 Spectral Shaping in Linear Stream Modulations . . . 235
10.5 Testing Orthogonal Modulations . . . 238
10.5.1 The Scatter Plot . . . 239
10.5.2 Stream Modulations . . . 240
10.6 Conclusions . . . 242
10.7 Homework Problems . . . 242
10.8 Example Solutions . . . 244
10.9 Mini-Projects . . . 244
10.9.1 Project 1 . . . 245
10.9.2 Project 2 . . . 246

III Advanced Digital Communication **249**

11 Demodulation in Frequency Selective Channels **251**
11.1 Current Status . . . 251
11.2 General M-ary Modulations . . . 253
11.3 Frequency Selectivity and OCDM . . . 254
11.3.1 MLWD for OCDM in Frequency Selective Channels . . . 255
11.3.2 Suboptimal OCDM Demodulation . . . 261
11.4 Frequency Selectivity and OFDM . . . 266
11.4.1 Simplified Demodulator for OFDM . . . 269
11.5 Conclusion . . . 274
11.6 Homework Problems . . . 275
11.7 Projects . . . 282

12 Frequency Selectivity and Stream Modulations **283**
12.1 Current Status . . . 283
12.2 MLWD for Stream Modulations . . . 287
12.2.1 Forward Step–Time 1 . . . 290
12.2.2 Survivor Selection–Time 1 . . . 292
12.2.3 Forward Step – General . . . 293
12.2.4 Survivor Selection – General . . . 297
12.2.5 Termination . . . 301
12.2.6 Ungerboeck MLWD Summary . . . 301
12.2.7 MLWD Performance . . . 303
12.3 Equalization . . . 303
12.3.1 Zero Forcing Equalizer . . . 305
12.3.2 Linear MMSE Equalizer . . . 307
12.3.3 Decision Feedback Equalizer . . . 309
12.4 Conclusion . . . 313
12.5 Homework Problems . . . 313

13 Orthogonal Modulations with Memory **319**
13.1 Canonical Problems . . . 319
13.2 Orthogonal Modulations with Memory . . . 320
13.2.1 MLWD for Orthogonal Modulations with Memory . . . 321
13.2.2 An Example OMWM Providing Better Performance . . . 322
13.2.3 Discussion . . . 323
13.3 Spectral Characteristics of Stream OMWM . . . 325
13.3.1 An Example OMWM with a Modified Spectrum . . . 325
13.3.2 Spectrum of OMWM for Large K_b . . . 326
13.4 Varying Transmission Rates with OMWM . . . 330
13.4.1 Spectrum of Stream Modulations with $N_m > 1$ . . . 330
13.4.2 Example $R < 1$: Convolutional Codes . . . 332
13.4.3 Example $R > 1$: Trellis Codes . . . 335
13.4.4 Example for Spectral Shaping: Miller Code . . . 336
13.5 Forney Equivalence . . . 339
13.6 Conclusions . . . 343
13.7 Homework Problems . . . 344
13.8 Projects . . . 351

14 Maximum Likelihood Sequence Demodulation **353**
14.1 Maximum Likelihood Sequence Demodulation . . . 354
14.1.1 The Viterbi Algorithm: Enabling Observations . . . 357
14.1.2 The Viterbi Algorithm: Formal Derivation . . . 359
14.1.3 Initialization and Termination . . . 363
14.2 Performance of MLWD for OMWM . . . 363
14.2.1 Simple Error Events . . . 364
14.2.2 The Modified Trellis . . . 366
14.3 Approximations for Large Frame Size . . . 369
14.3.1 Full Union Bound . . . 375
14.3.2 State Reduction Techniques . . . 378
14.4 Other Methods for MLWD for OMWM . . . 380
14.5 Homework Problems . . . 380
14.6 Mini-Projects . . . 383
14.6.1 Project 1 . . . 383
14.6.2 Project 2 . . . 384

A Detection Theory **391**
A.1 Binary Hypothesis Testing . . . 391
A.2 Karhunen Loève Expansion . . . 393
A.2.1 Characteristics of the KL Expansion . . . 395
A.3 Coherent Detection . . . 397
A.3.1 Coherent Detection in AWGN . . . 401
A.4 Noncoherent Detection . . . 401
A.4.1 Noncoherent Detection In AWGN . . . 402

B Minimum Mean Square Error Estimation **403**

C Notation **407**

D Acronyms **409**

Chapter 1

Introduction

The students that read this book have grown up with pervasive communications. A vast majority have listened to broadcast radio and television, used a mobile phone, surfed the world wide web, and played a compact or video disk recording. Hence there is no need for this book to motivate the student about the utility of communication technology. They use it everyday. This chapter will consequently be focused on the engineering aspects of communication technology that are not apparent from a user perspective.

1.1 Historical Perspectives on Communication Theory

Getting information from point A to point B using electricity or magnetism is the subject of this book. This field was born in the mid-1800's with the telegraph and continues today in a vast number of applications. Humans have needed communications since prehistoric times for the building of wealth and the waging of war. These social forces with the aid, at various points in time, of government sponsored monopolies have continuously pushed forward the performance of communications. It is perhaps interesting to note that the first electronic communications (telegraphy) was sending digital data (words were turned into a series of electronic dashes and dots). As the invention of the telephone took hold (1870s) communication became more focused on analog communication as voice was the information source of most interest to convey. The first World War led to great advances in wireless technology and television and radio broadcast soon followed. Again the transmitted information sources were analog. The digital revolution was spawned by the need for the telephone network to multiplex and automatically switch a variety of phone calls. A further technology boost was given during the second World War in wireless communications and system theory. The cold war led to rapid advances in satellite communications and system theory as the race for space gripped the world's major technology innovation centers. The invention of the semiconductor transistor and the march of Moore's law has spurred the march of innovation since the early 1980s. The evolving power of the microprocessor, the embedded computer and the signal processor has enabled algorithms that were considered preposterous at their formulation to see cost effective implementation. Distilling this 150 plus years of innovation into a small part of an engineering curriculum is a challenge but one this book arrogantly attempts.

The relative growth rate of electronic communications is phenomenal. Consider for example transatlantic transmission of information using under sea cables. This system has gone from from roughly 10 bits/sec in 1866 to roughly 10^{12} bits/sec in the year 2000 [Huu03]. The world community has gone in a very short period of time from accepting message delivery delays of weeks to down to seconds. The period from 1850-1900 was one filled with a remarkable advances in technology. Remarkably the advances in communications prior to 1900 can almost all be attributed to a single individual or invention. This started to change as technology became more complex in the 1900s. Large corporations and research labs began to be formed to support the large and complex systems that were evolving.

The evolution of these technologies and the personalities involved in their development are simply fascinating. Several books that are worth some reading if you are interested in the history of the field are [Huu03, Bur04, SW49, Bra95, Les69]. It is a rare invention that has an uncontested claim to ownership. These intellectual property disputes have existed from the telegraph up until modern times, but the tide of human innovation seems to be ever rising in spite of who gets credit for all the advances.

The ability to communicate has been markedly pushed by advances in technology but this book is not about technology. From the invention of the microphone, the electric motor, the electronic tube, the transistor, to the laser, engineers and physicist have made great technology leaps forward. These technological leaps have made great advances in communications possible. As technology has advanced the job of an engineer has become multifaceted and specialized over time. What once was a field where non-experts could contribute[1] prior to 1900 became a field where great specialization was needed in the post 1900 era. Two areas of specialization formed through the 1900s; the devices engineer and the systems engineer. The devices engineer is focused on designing technology to complete certain tasks. Devices engineers for example build antennas, amplifiers and/or oscillators and are heavily involved with current technology. Systems engineers try to put devices together in a way that will work as a system to achieve an overall goal. System engineers try to form mathematical models for how systems operate and use these models to design and specify systems. This text is written with a **systems engineering** perspective. Systems engineering in communications did not really come to be a formalized field until the early 1900s hence few of the references in this book were published before 1900. Some interesting historical system engineering references are [Car26, Nyq28, Arm35, Har28, Ric45, Sha48, Wie49]. This systems level perspective is very useful for education in that while technology will change greatly during an engineer's career the theory will be reasonably stable.

1.2 Goals For This Text

What this text is attempting to do is to show the mathematical and engineering underpinings to communication systems and systems engineering. While most students have used communication technology, few realize that the technology is built upon a strong core of engineering principles and over 100 years of hard work by a large group of talented people. Without a talented engineering workforce who understood the fundamental theory and put this theory into practice, humans would not have been able to deploy the pervasive communications society experiences. The goal for this text is to have some small part in the education of the workforce that will implement the next 50 years of progress. To reach this goal this text will focus on teaching the fundamentals of communication theory by

1. demonstrating that the mathematical tools the students have learned in their undergraduate education are useful in engineering practice,

2. showing that with modern integrated circuits the theory is directly reflected in engineering practice,

3. detailing how engineering tradeoffs in a communication system are ever evolving and that these tradeoffs involve performance, bandwidth efficiency, and complexity of the implementation.

Hopefully in addition to these professional goals, the reader of this text will come away with

1. A historical perspective on the hard work that has led to the current state of the art.

2. A sense of how fundamental engineering tools have real impact on system design.

[1] For example, Samuel Morse of Morse code and telegraph systems fame in the United States was a professor in the liberal arts.

3. A realization that fundamental engineering tools have changed little even as the technology to implement designs has evolved at a whithering pace.

4. An understanding that communications engineering is a growing and evolving entity and that continued education will be an important part of a career as a communication engineer.

1.3 Modern Communication System Engineering

Modern communication systems are very complex systems and no one engineer can be an expert in all the areas of the system. The initial communication systems were very simple point-to-point communication systems (telegraphy) or broadcast systems (commercial radio). As these systems were simple the engineering expertise could be common. As systems started to get more sophisticated (public telephony) a bifurcation of the needed expertise to address problems became apparent. There was a need to have an engineer who understood the details of the physical channel and how the information was transmitted and decoded. In addition there was a need to have an engineer who abstracted the problem at a higher level. This "higher level" engineer needed to think about switching architectures, supporting multiple users, scalability of networks, fault tolerance, and supporting applications. As the amount of information, system design options, and technology to implement these options grew further subdisciplines grew within the communications engineering field.

Modern communication systems are typically designed in layers to compartmentalize the different expertise and ease the interfacing of these multitude of expertises. In a modern system the communication system has a high level network architecture specification. This high level architecture is typically broken down into layers for implementation. The advantage of the layered architecture in the design process is that in designing a system for a particular layer the next lower layer can be dealt with as an abstract entity and the higher layer functions do not impact the design. Another advantage to the layered design is that components can be reused at each layer. This allows services and systems to be developed much more quickly in that designs can reuse layers from previous designs when appropriate. This layered design eliminated monolithic communications systems and allowed incremental changes much more readily.

An example of this layered architecture is the *open systems interconnection* (OSI) model. The OSI model was developed by the International Organization for Standardization (ISO) and has found significant utilization in practice. The OSI reference model is shown in Fig. 1.1. Each layer of abstraction communicates logically with entities at the same layer but produces this communication by calling the next lower layer in the stack. Using this model for instance it is possible to develope different applications (e.g., email versus web browsing) on the same base architecture (e.g., public phone system) as well as provide a method to insert new technology at any layer of the stack without impacting the rest of the system performance (e.g., replace a telephone modem with a cable modem). This concept of a layered architecture has allowed communications to take great advantage of prior advances and leap–frog technology along at a phenomenal pace.

This text is entirely focussed on what is known as **physical layer** communications. The physical layer of communications refers to the direct transfer of physical messages (analog waveforms or digital data bits) over a communications channel. The model for a physical layer communication abstraction is shown in Fig. 1.2. Examples of physical communication channels include, copper wire pairs (telephony), coaxial cables, radio channels (mobile telephony), or optical fibers. The engineering tools, the technology, and design paradigms are significantly different at the physical layer than at the higher layers in the stack. Consequently system engineering expertise in practice tends to have the greatest divide at the boundary to the physical layer. Engineering education has followed that trend and typically course work in telecommunications at both the undergraduate level and the graduate level tends to be bifurcated along these lines. To reflect the trend in both education and in industrial practice this book will

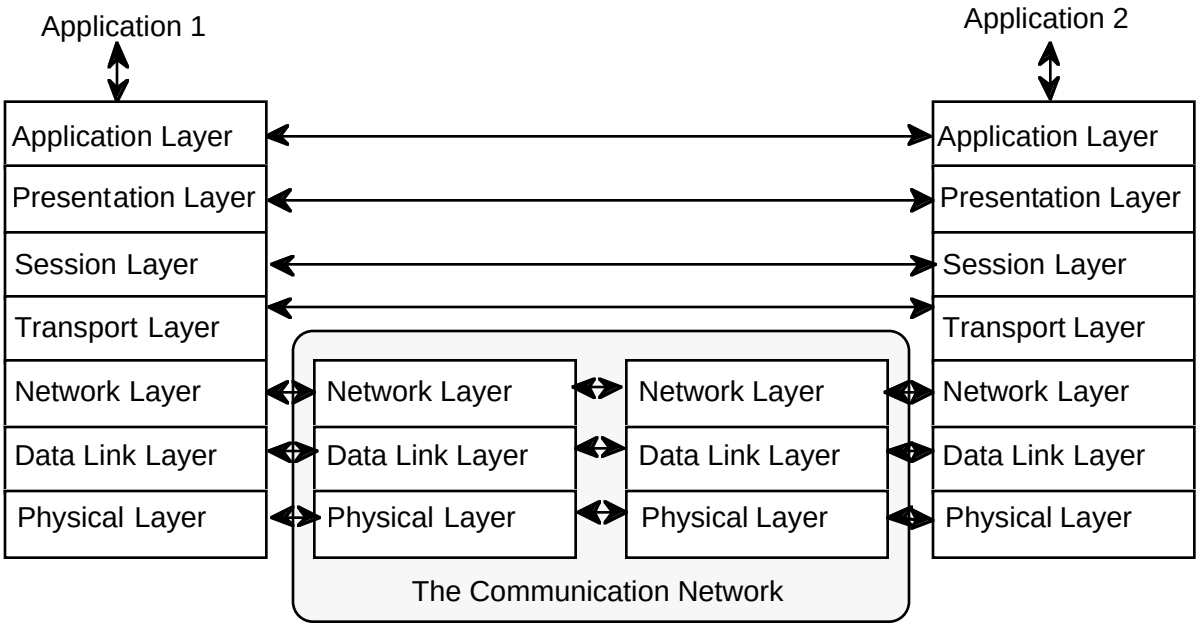


Figure 1.1: The OSI reference model.

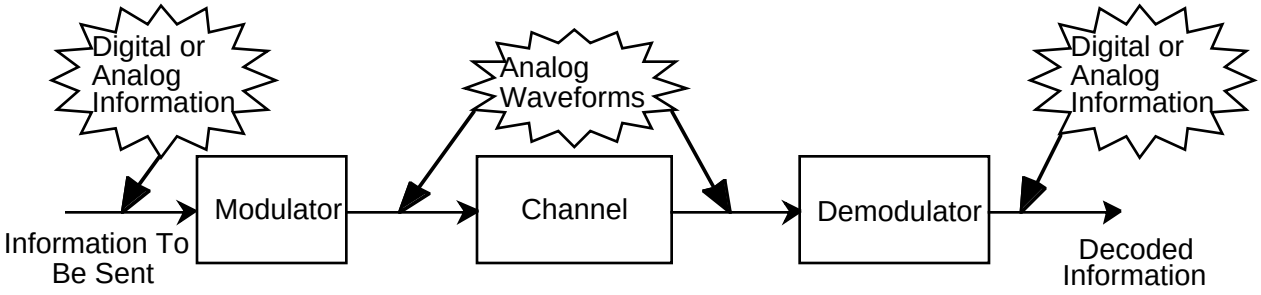


Figure 1.2: The physical layer model.

only try to educate in the area of physical layer communication systems. To reflect this abstraction the perspective in this text will be on point-to-point communications. Certainly multiple access communications is very important in practice but it will not be considered in this text to maintain a consistent focus. Students who want a focus more on the upper layers of the communication stack should refer to [KR04, LGW00, Tan02].

1.4 Technology's Impact on this Text

This text has been heavily influenced by the relatively recent trends in the field of physical layer communications system engineering

- Advanced communication theory finding utility in practice,
- Baseband processing power increasing at a rate predicted by Moore's law,
- Iterative processing as realized in turbo or low density parity check codes gaining wide spread

acceptance in engineering practice.

Physical layer communications (with the perhaps partial exception of fiber optics) is filled with examples of sophisticated communication **theory** being directly placed into **practice**. Examples include wireless digital communications and high speed cable communications. A communication engineer should truly feel lucky to live in a time when theory and practice are linked so closely. It allows people to work on very complex and sophisticated algorithms and have the algorithms almost immediately be put into practice. Because of this reason it is not surprising that many of the prominent communication theorist have also been very successful entrepreneurs (Andrew Viterbi [VO79] and Irwin Jacobs [WJ65] being two obvious examples). In this text we will attempt to feature the underlying theory as this theory is so important in practical systems from mobile phones to television receivers.

The reason that theory is migrating to practice so quickly is the rapid advance of **baseband processing power**. Moores law is now almost outstripping the ability of communication theory to use the available processing. In fact a great paradigm shift occurred in the industry (in my humble opinion) when Qualcomm, in championing the cellular standard IS-95, started the design philosophy of designing a system that was too complicated for the current technology with the knowledge that Moore's law would soon enable the design to be implemented in a cost effective fashion. Because of this shift in the design philosophy, future engineers are going to be exploring ways to better utilize this ever increasingly cost efficient processing power. Since the future engineer will be using baseband processing power to implement their algorithms, this text is written with a focus on the baseband signal processing. To reflect this focus, this text starts immediately with the complex envelope representation of carrier signals and uses this representation throughout the entire text. This is a significant deviation from most of the prior teaching texts but directly in line with the notation used in research and in industry.

Iterative processing has fundamentally changed how communication engineers view problems. In 1948 Claude Shannon laid down the foundations of digital communications [Sha48]. In his seminal work Shannon identified the fundamental upper limits for rates at which reliable communication can be achieved on some important channels. This upper limit is known as channel capacity. Before the 1990's the digital communication engineer searched for ways to approach Shannon capacity at a reasonable complexity. These pre-1990 techniques usually involve identifying optimal or slightly suboptimal demodulation algorithms. The complexity of these algorithms usually grew exponentially as the performance approached the capacity. At a particular time the complexity of the modulation was pushed as far as the technology would allow. Since Moore's law only allows a doubling of the complexity every 18 months, performance was approaching capacity slowly. In 1993 a paper was published [BGT93] that changed the course of the communications engineering philosophy. This paper forced communications engineers to explore the ideas of message passing and suboptimal iterative decoding and led to the rediscovery of some long latent work [Gal62] on powerful Shannon capacity approaching codes. This paradigm shift led communication engineers to successfully adopt this suboptimal iterative decoding approach to a wide variety of problems. Because of this new found tool it is important to teach the fundamentals of digital communications in such a way that this tool is a natural consequence of the pedagogy. This book is my attempt to integrate the fundamental ideas of message passing and probabilistic inference into the core of digital communications in a way that a text written before 1993 would not attempt. This approach will hopefully give a much better backgound to the modern communication engineer.

1.5 The Text Relation to Communication Curriculum

This book is written for the modern communications curriculum. Most modern communications curriculum at the undergraduate level have a networking course hence no coverage is given for networking in this book. Example texts that might be used in a networking course in a modern undergraduate

curriculum include [KR04, LGW00, Tan02]. The course objectives that can be taught from this text are (along with their ABET criteria)

- Students learn the bandpass representation for carrier modulated signals. (Criterion 3(a))
- Students engage in engineering design of communications system components. (Criteria 3(c),(k))
- Students learn to analyze the performance, spectral efficiency and complexity of the various options for transmitting analog and digital message signals. (Criteria 3(e),(k))
- Students learn to characterize noise in communication systems. (Criterion 3(a))

Prerequisites for the undergraduate course that could be taught out of this book are probability and random variables and a signal and systems course.

This text is written to compliment the digital communications curriculum at the graduate level. At the graduate level this text is composed with the assumption that the following courses and *example* texts complement the course:

1. Probability and random processes (Yates and Goodman) (pre-requisite)
2. Information theory and source coding (Cover and Thomas)
3. Error control coding (Algebraic, trellis and concatenated coding) (Lin and Costello)
4. Estimation and detection (Poor) (co-requisite)
5. Wireless communications (Stuber)

The material on random processes is covered but only to the level exactly needed for this text. The abovementioned set of courses would provide an appropriate coverage of physical layer communications at the graduate level and reflects what is taught in most graduate programs.

This book is constructed to align with the quote by Albert Einstein

> *Everything should be made as simple as possible, but not one bit simpler.*

Consequently text will be void of advanced topics in communication theory that I did not see as fundamental in an introductory communication theory book. Examples here include details of information theory and source coding. While these topics are critical to the training of a communications engineer it is not necessary to the understanding of analog and digital information transmission. The goal is essentially not to lose the proverbial forest for the trees. Many interesting advanced issues and systems are pursued in the homework problems and projects. The text is written to build up a tool set in students that allow them to flourish in their profession over a full career. Readers looking for a buzzword level treatment of communications will not find the text satisfactory. Since the focus of this text is the tools that will be important in the future, many ideas are not discussed in detail that traditionally were prominent in communication texts (e.g., pulse modulations). While a communication text can often take the form of an encyclopedia I have purposely avoided this format for a more focused tool oriented version. Writing this paragraph I feel a little like my Mom telling me to "eat my vegtables" but as I grow in age (and hopefully wisdom) I more fully appreciate the wisdom of my parents and of learning fundamental tools in physical layer communication engineering.

1.6 Book Overview

The book consists of five parts

I Mathematical Foundations

II Analog Communication

III Noise in Communications Systems

IV Fundamentals of Digital Communications

V Advanced Topics in Digital Communications

This organization allows a slow logical build up from a base knowledge in Fourier transforms, linear systems, and probability to an understanding of the fundamental concepts in communication theory. A significant effort has been made to make the development logical and to cover the important concepts.

1.6.1 Mathematical Foundations

This part of the book consists of three chapters that provide the mathematical foundations of communication theory. There are three pieces of test equipment that are critical for a communication engineer to be able to use to understand and troubleshoot communication systems: the oscilloscope, the spectrum analyzer, and the vector signal analyzer. Most communications laboratories contain this equipment and examples of this equipment are

1. Digital oscilloscope and logic analyzer – Agilent 54622D
2. Spectrum analyzer – Agilent E4402B
3. Vector signal analyzer – Agilent 89600

The first chapter of this part of the book covers material that is included in the core undergraduate electrical engineering curriculum that provides the theory for the operation for the oscilloscope (time domain characterizations of signals) and the theory for the operation of the spectrum analyzer (frequency domain characterization). The fundamental difference between a communication engineer and a technician is that engineers will consider noise to provide a complete characterization of the system tradeoffs. The tools used to charcterize noise are based on probability theory. The second chapter reviews this material. Courses covering the material on signal and systems and probability theory are taken before a course in communication theory. These two chapters are included strictly as review and to establish notation for the remainder of the book. The chapter on bandpass signals and the complex envelope notation is where the student steps into modern communication systems theory. This chapter will teach students how to characterize bandpass signals in the time domain and in the frequency domain. This chapter also introduces the vector diagram which is the theoretical basis of the vector signal analyzer which is a tool that is frequently used by the modern communication engineer. These mathematical foundations will provide the basis for communications engineering.

1.6.2 Analog Communication

This part of the book consists of four chapters that introduce the theory of bandpass analog communication. The approach taken here is to introduce analog communications before the concepts of random processes. Consequently the message signal is treated as a known deterministic waveform in the discussion of analog communications. The downside of this approach is that many of the powerful results

on the power spectrum of analog communication waveforms cannot be introduced. The advantage of this approach is that since the message signals are known and deterministic, the tools of Fourier series, Fourier transforms, and signals and systems can be applied to the understanding of analog modulation systems. This approach allows students to ease into the world of communications by building upon their prior knowledge in deterministic signal and system analysis. By tying these tools from early undergraduate courses into the process of assessing the spectral efficiency and complexity of analog communications, the student will see the efficacy of an education in the fundamentals of electrical engineering. The first chapter of this part of the book presents the performance metrics in communications: performance, complexity, and spectral efficiency. This three level tradeoff is a recurring theme in the book. The following chapters introduce the classic methods of communicating analog information; amplitude and angle modulation. Amplitude modulation is a modulation format where the message signal is impressed upon the amplitude of the bandpass signal. Similarly angle modulation is a modulation format where the message signal is impressed in some way in the phase of the bandpass signal. Finally, the important ideas in analog communications of multiplexing and the phase locked loop are presented.

Many of my professional colleagues have made the suggestion that analog modulation concepts should be removed from the modern undergraduate curriculum. Comments such as "We do not teach about vacuum tubes so why should we teach about analog modulations?" are frequently heard. I heartily disagree with this opinion but not because I have a fondness for analog modulation but because analog modulation concepts are so important in modern communication systems. The theory and notation for analog signals learned in this text is a solid foundation for further explorations into modern communication systems because modern digital communications use analog waveforms. For example in the testing of modern communication systems and subsystems analog modulation and demodulation concepts are used extensively. In fact most of my good problems for the analog communication chapters have come as a result of my work in experimental wireless communications even though my research work has always been focused on digital communication systems! Another example of the utility of analog communications is that I am unaware of a synthesized signal generator that does not have an option to produce amplitude modulated (AM) and frequency modulated (FM) test signals. While modern communication engineers do not often design analog communication systems, the theory is still a useful tool. Consequently this part of the book focuses on analog communications but using a modern perspective that will provide students the tools to flourish in their careers.

1.6.3 Noise in Communication Systems

Understanding the effects of noise and interference on communication systems what makes a communication system engineer uniquely trained. I have always been struck by the fact the engineering technicians have a training in time domain analysis, Fourier analysis, modulation techniques and demodulation techniques. The main thing a technician does not understand is how to characterize how noise impacts the tradeoffs that must be made in system design. On the other hand the understanding of noise is often a frustrating subject for students as the level of mathematics and abstraction can often seem not worth the gains in useful skills. The approach taken in this text is to introduce the minimal amount of abstraction necessary to get useful results for engineering practice.

There are four topics/chapters that are presented in this section to introduce the techniques to analyze the impact of noise and interference. The first chapter focuses on the characterization of Gaussian stationary random processes and how linear filters impact this characterization. This material builds heavily on probability and random variable concepts. This text offers little new insights than has been available since the 1950s [DR87, Pap84] other than a reordering of topic presentation. The next chapter generalizes the concepts of random processes to the case of noise in bandpass communication receivers. The impact of filters in the receiver on the noise characteristics is explored. After these preliminary tools are in place, a revisiting of all the forms of analog communications in the presence of noise is

completed. At this point an understanding of the tradeoffs associated with analog communications can be completed and interpreted. A final short chapter introduces some concepts in random processes that are useful for various topics in digital communications.

1.6.4 Digital Basics

This part of the book consists of five chapters that introduce the theory of memoryless digital communication. The first chapter of this part of the book again presents the performance metrics in communications: performance, complexity, and spectral efficiency and reinterprets these metrics for digital communications. This re-emphasizes the tradeoff first introduced with analog modulation. The following chapter introduces the classic methods of communicating one bit of information. This is developed by going through a five step design process. The digital communication design problem is generalized for the transmission of multiple bits of information in the following chapter. Here both optimum structures for decoding the entire word and the individual bits are considered. The unfortunate situation exists with multiple bit transmission that the complexity of the optimum decoder grows exponentially with the number of bits to be transmitted. To address this exponential growth in complexity signal structures that offer reduced complexity optimum demodulation structures are introduced. Several examples of these reduced complexity modulations are used in engineering practice and these examples are considered in detail. Finally the last chapter of this section considers techniques to greatly improve the spectral efficiency of digital communications. The final section of this part of the book considers the tools used to test digital communication systems in practice. The ideas and uses of the vector diagram and the eye diagram in the testing of digital communication systems are explained and motivated.

1.6.5 Advanced Topics in Digital Communications

Not included this edition.

Part I

Mathematical Foundations

Chapter 2

Complex Baseband Representation of Bandpass Signals

2.1 Introduction

A majority of communication systems operate by modulating an information bearing waveform onto a sinusoidal carrier. As examples, Table 2.1 lists the carrier frequencies of various methods of electronic communication.

Type of Transmission	Center Frequency of Transmission
Telephone Modems	1600-1800 Hz
AM radio	530-1600 kHz
CB radio	27 MHz
FM radio	88-108 MHz
VHF TV	178-216 MHz
Cellular radio	850 MHz, 1.8GHz
Indoor Wireless Networks	2.4GHz
Commercial Satellite Downlink	3.7-4.2 GHz
Commercial Satellite Uplink	5.9-6.4 GHz
Fiber Optics	2×10^{14} Hz

Table 2.1: Carrier frequency assignments for different methods of information transmission.

One can see by examining Table 2.1 that the carrier frequency of the transmitted signal is not the component which contains the information. Instead it is the signal modulated on the carrier which contains the information. Hence a method of characterizing a communication signal which is independent of the carrier frequency is desired. This has led communication system engineers to use a **complex baseband representation** of communication signals to simplify their job. All of the communication systems mentioned in Table 2.1 can be and typically are analyzed with this complex baseband representation. This chapter develops the complex baseband representation for deterministic signals. Other references that develope these topics well are [Pro89, PS94, Hay83, BB99]. One advantage of the complex baseband representation is simplicity. All signals are lowpass signals and the fundamental ideas behind modulation and communication signal processing are easily developed. Also any receiver that processes the received waveform digitally uses the complex baseband representation to develop the baseband processing algorithms. In fact complex baseband representation is so prevalent in engineering systems that the most widely used tool, *Matlab*, has been configured by default to process all variables

in a program as complex signals. Hopefully by the time you are done with this course the utility of this view will be apparent.

2.2 Baseband Representation of Bandpass Signals

The first step in the development of a complex baseband representation is to define a bandpass signal.

Definition 2.1 *A bandpass signal, $x_c(t)$, is a signal whose one-sided energy spectrum is both: 1) centered at a non-zero frequency, f_C, and 2) does not extend in frequency to zero (DC).*

The two sided transmission bandwidth of a signal is typically denoted by B_T Hertz so that the one-sided spectrum of the bandpass signal is zero except in $[f_C - B_T/2, f_C + B_T/2]$. This implies that a bandpass signal satisfies the following constraint: $B_T/2 < f_C$. Fig. 2.1 shows a typical bandpass energy spectrum. Since a bandpass signal, $x_c(t)$, is a physically realizable signal it is real valued and consequently the energy spectrum will always be even symmetric around $f = 0$. The relative sizes of B_T and f_C are not important, only that the spectrum takes negligible values around DC. In telephone modem communications this region of negligible spectral values is only about 300Hz while in satellite communications it can be many Gigahertz.

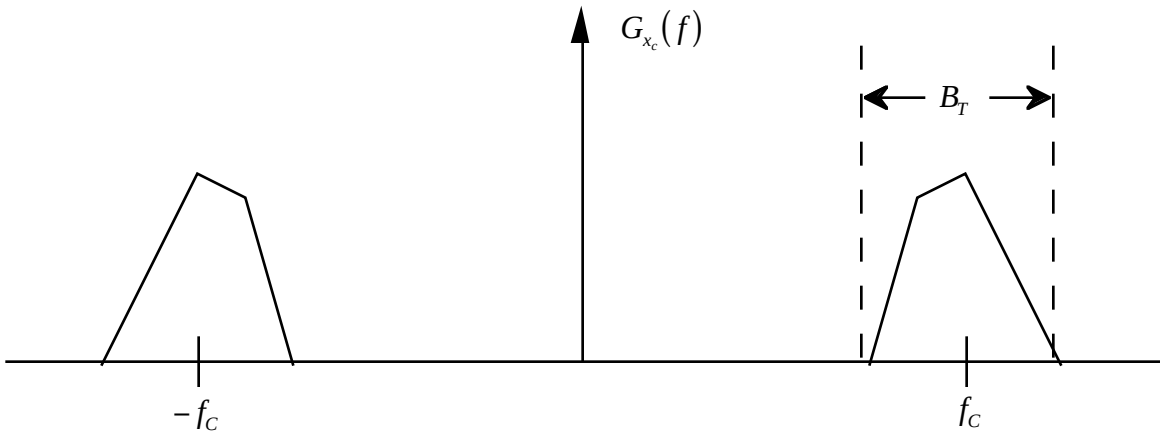


Figure 2.1: Energy spectrum of a bandpass signal.

A bandpass signal has a representation of

$$\begin{aligned} x_c(t) &= x_I(t)\sqrt{2}\cos(2\pi f_c t) - x_Q(t)\sqrt{2}\sin(2\pi f_c t) && (2.1)\\ &= x_A(t)\sqrt{2}\cos\left(2\pi f_c t + x_P(t)\right) && (2.2)\end{aligned}$$

where f_c is denoted the carrier frequency with $f_C - B_T/2 \leq f_c \leq f_C + B_T/2$. The signal $x_I(t)$ in (2.1) is normally referred to as the **in-phase (I)** component of the signal and the signal $x_Q(t)$ is normally referred to as the **quadrature (Q)** component of the bandpass signal. $x_I(t)$ and $x_Q(t)$ are real valued lowpass signals with a one-sided non–negligible energy spectrum no larger than B_T Hertz. Two items should be noted

- The center frequency of the bandpass signal, f_C, (see Fig. 2.1) and the carrier frequency, f_c are not always the same. While f_c can theoretically take a continuum of values, in most applications an obvious value of f_c will give the simplest representation[1].

[1]This idea will become more obvious in Chapter **??**.

- The $\sqrt{2}$ term is included in the definition of the bandpass signal to ensure that the bandpass signal and the baseband signal have the same power/energy. This will become apparent in Section. 2.4.

The carrier signal is normally thought of as the cosine term, hence the I component is in-phase with the carrier. Likewise the sine term is 90° out-of-phase (in quadrature) with the cosine or carrier term, hence the Q component is quadrature to the carrier. Equation (2.1) is known as the canonical form of a bandpass signal. Equation (2.2) is the amplitude and phase form of the bandpass signal, where $x_A(t)$ is the **amplitude** of the signal and $x_P(t)$ is the **phase** of the signal. A bandpass signal has two degrees of freedom and the I/Q or the amplitude and phase representations are equivalent. The transformations between the two representations are given by

$$x_A(t) = \sqrt{x_I(t)^2 + x_Q(t)^2} \qquad x_P(t) = \tan^{-1}\left[x_Q(t), x_I(t)\right] \tag{2.3}$$

and

$$x_I(t) = x_A(t)\cos\left(x_P(t)\right) \qquad x_Q(t) = x_A(t)\sin\left(x_P(t)\right). \tag{2.4}$$

Note that the inverse tangent function in (2.3) has a range of $[-\pi, \pi]$ (i.e., both the sign $x_I(t)$ and $x_Q(t)$ and the ratio of $x_I(t)$ and $x_Q(t)$ are needed to evaluate the function). This inverse tangent function is different than the single argument function that is on most calculators. The particulars of the communication design analysis determine which form for the bandpass signal is most applicable.

A complex valued signal, denoted the **complex envelope**, is defined as

$$x_z(t) = x_I(t) + jx_Q(t) = x_A(t)\exp\left[jx_P(t)\right].$$

The original bandpass signal can be obtained from the complex envelope by

$$x_c(t) = \sqrt{2}\Re\left[x_z(t)\exp\left[j2\pi f_c t\right]\right]. \tag{2.5}$$

Since the complex exponential only determines the carrier frequency, the complex signal $x_z(t)$ contains all the information in $x_c(t)$. Using this complex baseband representation of bandpass signals greatly simplifies the notation for communication system analysis. As the quarter goes along hopefully the additional simplicity will become very evident.

Example 2.1: Consider the bandpass signal

$$x_c(t) = 2\cos(2\pi f_m t)\sqrt{2}\cos(2\pi f_c t) - \sin(2\pi f_m t)\sqrt{2}\sin(2\pi f_c t)$$

where $f_m < f_c$. A plot of this bandpass signal is seen in Fig. 2.2 with $f_c = 10f_m$. Obviously we have

$$x_I(t) = 2\cos(2\pi f_m t) \qquad x_Q(t) = \sin(2\pi f_m t)$$

and

$$x_z(t) = 2\cos(2\pi f_m t) + j\sin(2\pi f_m t).$$

The amplitude and phase can be computed as

$$x_A(t) = \sqrt{1 + 3\cos^2(2\pi f_m t)} \qquad x_P(t) = \tan^{-1}\left[\sin(2\pi f_m t), 2\cos(2\pi f_m t)\right].$$

A plot of the amplitude and phase of this signal is seen in Fig. 2.3.

The next item to consider is methods to translate between a bandpass signal and a complex envelope signal. Basically a bandpass signal is generated from its I and Q components in a straightforward fashion

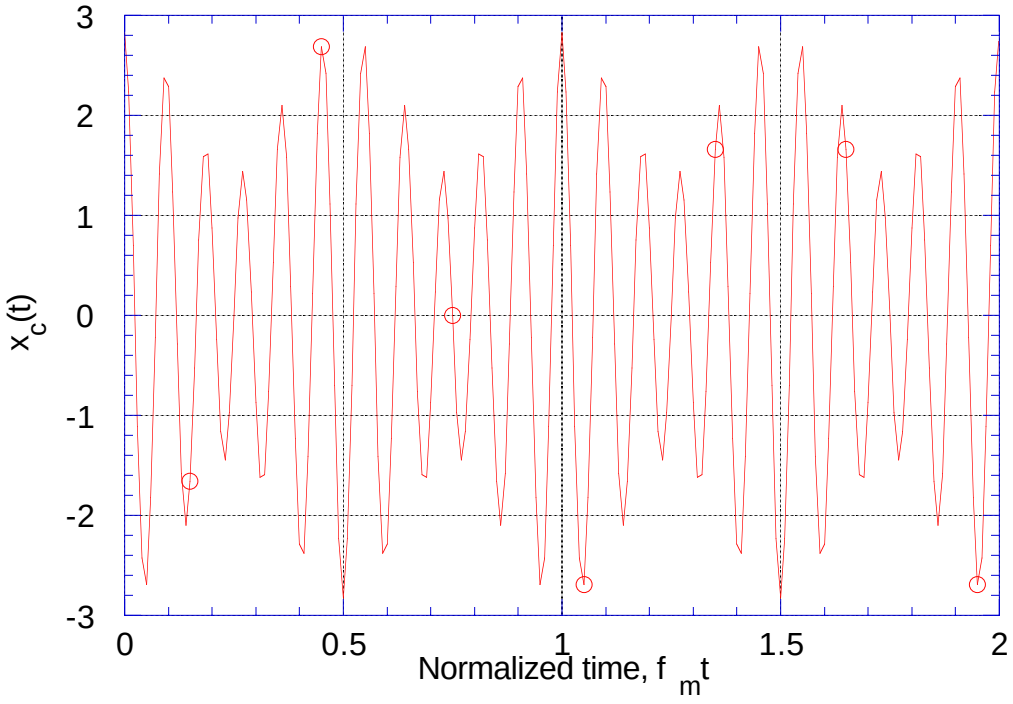


Figure 2.2: Plot of the bandpass signal for Example 2.1.

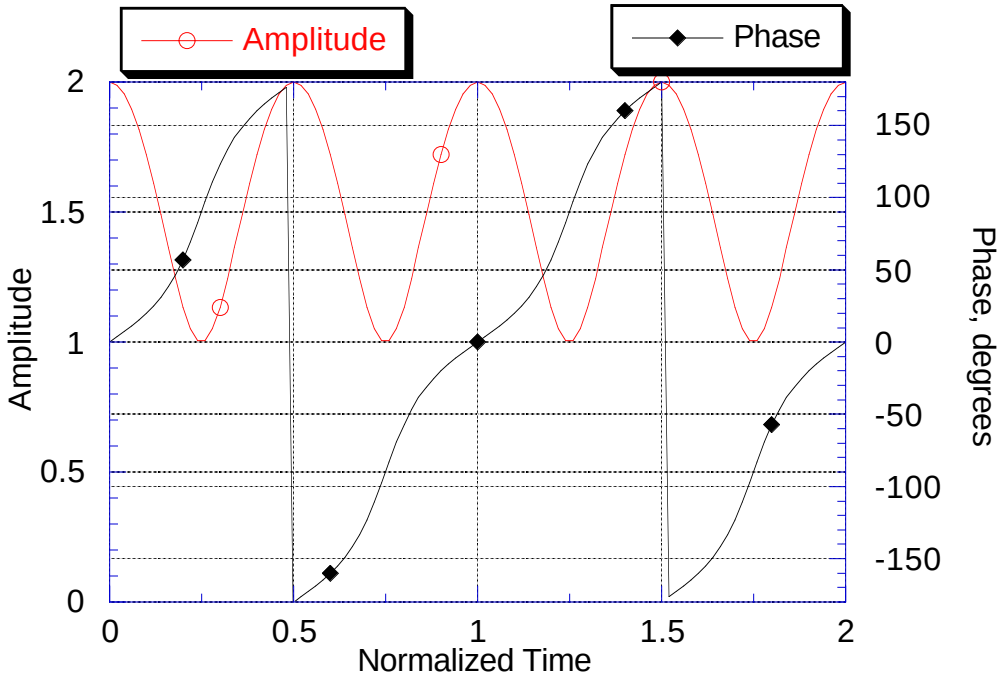


Figure 2.3: Plot of the amplitude and phase for Example 2.1.

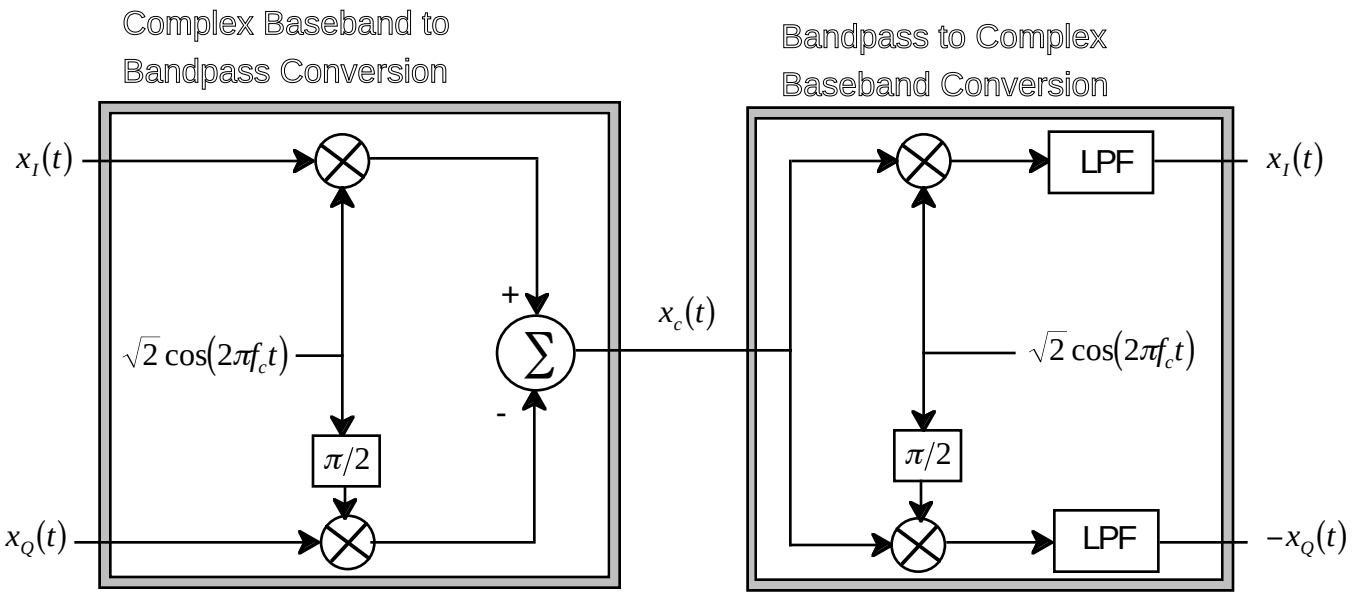


Figure 2.4: Schemes for converting between complex baseband and bandpass representations. Note that the LPF simply removes the double frequency term associated with the down conversion.

corresponding to (2.1). Likewise a complex envelope signal is generated from the bandpass signal with a similar architecture. The idea behind bandpass to baseband downconversion can be understood by using trigonometric identities to give

$$\begin{aligned} x_1(t) = x_c(t)\sqrt{2}\cos(2\pi f_c t) &= x_I(t) + x_I(t)\cos(4\pi f_c t) - x_Q(t)\sin(4\pi f_c t) \\ x_2(t) = x_c(t)\sqrt{2}\sin(2\pi f_c t) &= -x_Q(t) + x_Q(t)\cos(4\pi f_c t) + x_I(t)\sin(4\pi f_c t). \end{aligned} \quad (2.6)$$

In Fig. 2.4 the lowpass filters remove the $2f_c$ terms in (2.6). Note in Fig. 2.4 the boxes with $\pi/2$ are phase shifters (i.e., $\cos(\theta - \pi/2) = \sin(\theta)$) typically implemented with delay elements. The structure in Fig. 2.4 is fundamental to the study of all carrier modulation techniques.

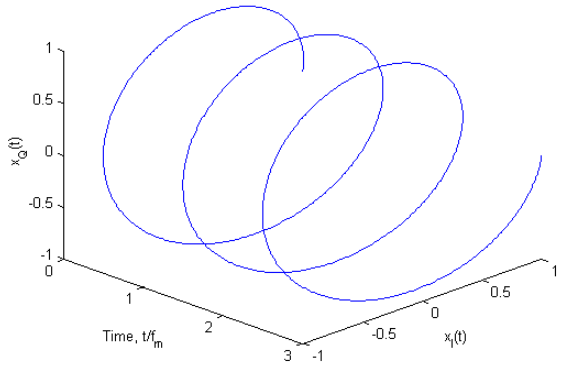
2.3 Visualization of Complex Envelopes

The complex envelope is a signal that is a complex function of time. Consequently the complex envelope needs to be characterized in three dimensions (time, in-phase, and quadrature). For example, the complex envelope given as

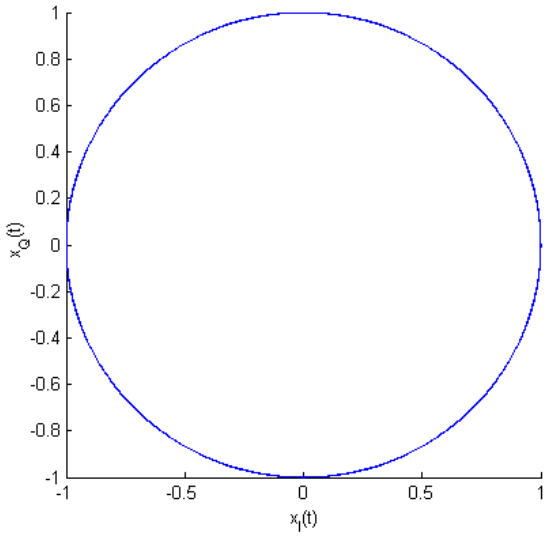
$$x_z(t) = \exp[j2\pi f_m t] \quad (2.7)$$

can be represented as a three dimensional plot as shown in Fig. 2.5-a). It is often difficult to comprehend all that is going on in a three dimensional plot when the complex envelope is a typical communication signal hence communication engineers often project these signals into two dimensions. Examples of this two dimensional projection are shown in Fig. 2.5-a) ($x_I(t)$ versus $x_Q(t)$), Fig. 2.6-a) (t versus $x_I(t)$), and Fig. 2.6-b) (t versus $x_Q(t)$). All of these methods of viewing and visualizing a complex envelope signal are used in engineering practice.

The vector diagram is the two dimensional projection where $x_I(t)$ is plotted versus $x_Q(t)$. The vector diagram represents the time trajectory of the complex envelope, $x_z(t)$ in the complex plane. This

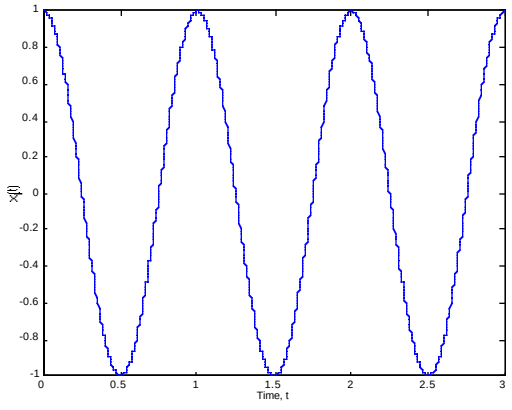


a) The three dimensional plot.

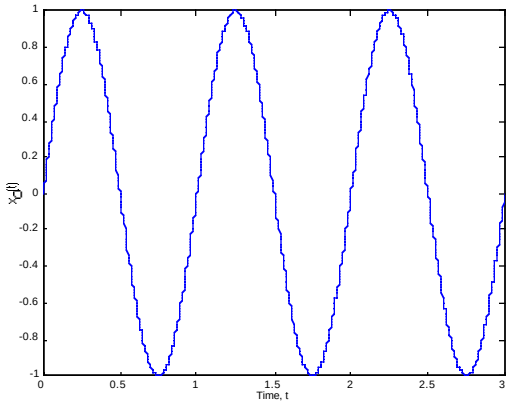


b) The projection onto the $x_I(t)$, $x_Q(t)$ plane.

Figure 2.5: Plots of a complex exponential.



a) The projection onto the t, $x_I(t)$ plane.



b) The projection onto the t, $x_Q(t)$ plane.

Figure 2.6: Plots of the two time functions.

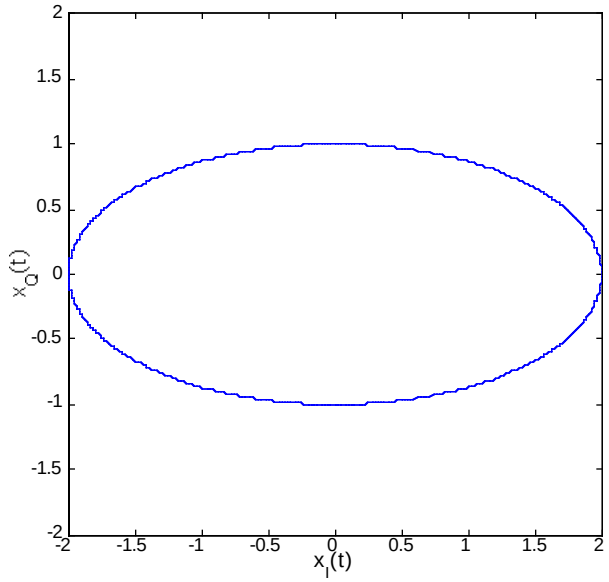


Figure 2.7: The vector diagram for the bandpass signal in *Example 2.1*.

vector diagram often gives significant insight into the performance or characteristics of a communication system and it will be used often in future sections of this text. The vector diagram first gained utility in the days when the standard tool for examining the time domain characteristics of a communication signal was the dual channel analog oscilloscope. To produce a vector diagram with a dual channel oscilloscope one puts $x_I(t)$ into one channel and $x_Q(t)$ into the second channel and configure the scope to plot channel one versus channel two. Since this visualization was simply with the early test instruments, it has found significant utility in engineering practice.

Example 2.2: The vector diagram for the bandpass signal given in *Example 2.1* is shown in Fig. 2.7. It should be noted that the point $x_z(t) = (2, 0)$ corresponds to time $t = n/f_m$ where n is an integer, e.g., $2\cos(2\pi n) = 2$ and $\sin(2\pi n) = 0$. Likewise $t = n/f_m + 1/(4f_m)$ corresponds to the point $x_z(t) = (0, 1)$.

2.4 Spectral Characteristics of the Complex Envelope

2.4.1 Basics

It is of interest to derive the spectral representation of the complex baseband signal, $x_z(t)$, and compare it to the spectral representation of the bandpass signal, $x_c(t)$. Assuming $x_z(t)$ is an energy signal, the Fourier transform of $x_z(t)$ is given by

$$X_z(f) = X_I(f) + jX_Q(f) \tag{2.8}$$

where $X_I(f)$ and $X_Q(f)$ are the Fourier transform of $x_I(t)$ and $x_Q(t)$, respectively, and the energy spectrum is given by

$$G_{x_z}(f) = |X_z(f)|^2 = G_{x_I}(f) + G_{x_Q}(f) + 2\Im\left[X_I(f)X_Q^*(f)\right] \tag{2.9}$$

where $G_{x_I}(f)$ and $G_{x_Q}(f)$ are the energy spectrum of $x_I(t)$ and $x_Q(t)$, respectively. The signals $x_I(t)$ and $x_Q(t)$ are lowpass signals with a one-sided bandwidth of less than $B_T/2$ so consequently $X_z(f)$ and $G_{x_z}(f)$ can only take nonzero values for $|f| < B_T/2$.

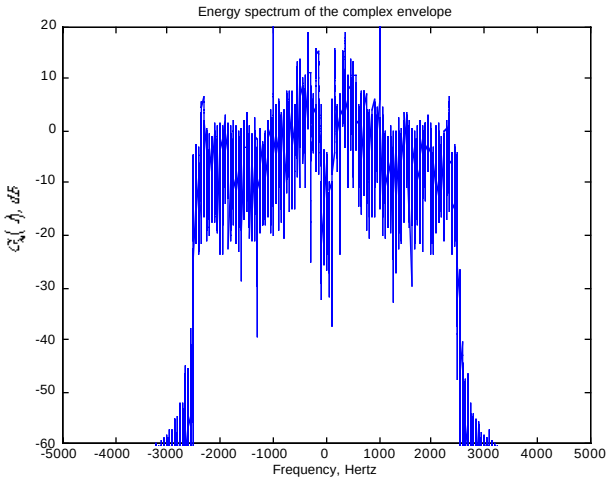


Figure 2.8: The energy spectrum resulting from $x_I(t)$ being a computer generated voice signal and $x_Q(t)$ being a sinusoid.

Example 2.3: Consider the case when $x_I(t)$ is set to be the message signal from Example 1.0(computer voice saying "bingo") and $x_Q(t) = \cos(2000\pi t)$. $X_I(f)$ will be a lowpass spectrum with a bandwidth of 2500Hz while $X_Q(f)$ will have two impulses located at $\pm$1000Hz. Fig. 2.8 show the measured complex envelope energy spectrum for these lowpass signals. The complex envelope energy spectrum has a relation to the voice spectrum and the sinusoidal spectrum exactly as predicted in (2.8). Note here $B_T = 5000$Hz.

Eq. (2.8) gives a simple way to transform between the lowpass signal spectrums to the complex envelope spectrum. A similar simple formula exists for the opposite transformation. Note that $x_I(t)$ and $x_Q(t)$ are both real signals so that $X_I(f)$ and $X_Q(f)$ are Hermitian symmetric functions of frequency and it is straightforward to show

$$\begin{aligned} X_z(-f) &= X_I^*(f) + jX_Q^*(f) \\ X_z^*(-f) &= X_I(f) - jX_Q(f). \end{aligned} \tag{2.10}$$

This leads directly to

$$\begin{aligned} X_I(f) &= \frac{X_z(f) + X_z^*(-f)}{2} \\ X_Q(f) &= \frac{X_z(f) - X_z^*(-f)}{j2}. \end{aligned} \tag{2.11}$$

Since $x_z(t)$ is a complex signal, in general, the energy spectrum, $G_{x_z}(f)$, has none of the usual properties of real signal spectra which have a spectral magnitude is an even function of f and a spectral phase that is an odd function of f.

An analogous derivation produces the spectral characteristics of the bandpass signal. Examining (2.1) and using the Frequency Translation Theorem of the Fourier transform, the Fourier transform of the bandpass signal, $x_c(t)$, is expressed as

$$X_c(f) = \left[\frac{1}{\sqrt{2}} X_I(f - f_c) + \frac{1}{\sqrt{2}} X_I(f + f_c)\right] - \left[\frac{1}{\sqrt{2}j} X_Q(f - f_c) - \frac{1}{\sqrt{2}j} X_Q(f + f_c)\right].$$

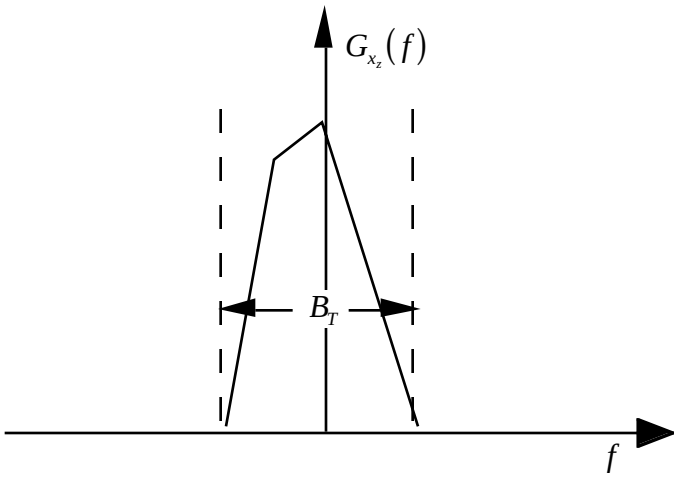


Figure 2.9: The complex envelope energy spectrum of the bandpass signal in Fig. 2.1 with $f_c = f_C$.

This can be rearranged to give

$$X_c(f) = \left[\frac{X_I(f-f_c) + jX_Q(f-f_c)}{\sqrt{2}}\right] + \left[\frac{X_I(f+f_c) - jX_Q(f+f_c)}{\sqrt{2}}\right]. \tag{2.12}$$

Using (2.10) in (2.12) gives

$$X_c(f) = \frac{1}{\sqrt{2}}X_z(f-f_c) + \frac{1}{\sqrt{2}}X_z^*(-f-f_c). \tag{2.13}$$

This is a very fundamental result. Equation (2.13) states that the Fourier transform of a bandpass signal is simply derived from the spectrum of the complex envelope. For positive values of f, $X_c(f)$ is obtained by translating $X_z(f)$ to f_c and scaling the amplitude by $1/\sqrt{2}$. For negative values of f, $X_c(f)$ is obtained by flipping $X_z(f)$ around the origin, taking the complex conjugate, translating the result to $-f_c$, and scaling the amplitude by $1/\sqrt{2}$. This also demonstrates that if $X_c(f)$ only takes values when the absolute value of f is in $[f_c - B_T/2, f_c + B_T/2]$, then $X_z(f)$ only takes values in $[-B_T/2, B_T/2]$. The energy spectrum of $x_c(t)$ can also be expressed in terms of the energy spectrum of $x_z(t)$ as

$$G_{x_c}(f) = \frac{1}{2}G_{x_z}(f-f_c) + \frac{1}{2}G_{x_z}(-f-f_c). \tag{2.14}$$

Since $E_{x_c} = \int_{-\infty}^{\infty} G_{x_c}(f)df = \int_{-\infty}^{\infty} G_{x_z}(f)df$, (2.14) guarantees that the energy of the complex envelope is identical to the energy of the bandpass signal. Additionally (2.14) guarantees that the energy spectrum of the bandpass signal is an even function of frequency as it should be for a real signal. Considering these results, the spectrum of the complex envelope of the signal shown in Fig. 2.1 will have a form shown in Fig. 2.9 when $f_c = f_C$. Other values of f_c would produce a different but equivalent complex envelope representation. This discussion of the spectral characteristics of $x_c(t)$ and $x_z(t)$ should reinforce the idea that the complex envelope contains all the information in a bandpass waveform.

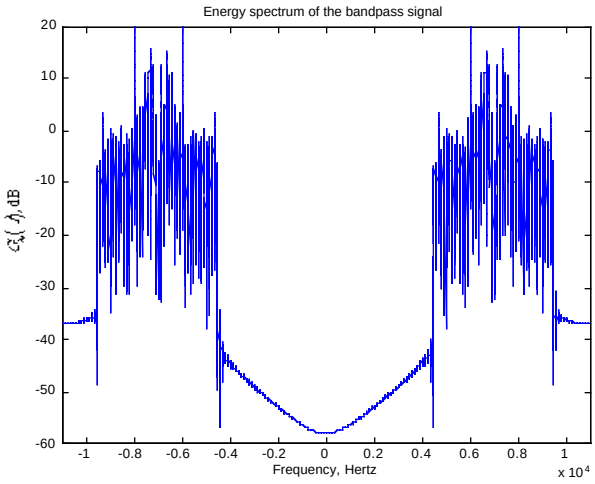


Figure 2.10: The bandpass spectrum corresponding to Fig. 2.8. $B_T = 5000Hz$

Example 2.4: (Example 2.1 continued)

$$x_I(t) = 2\cos(2\pi f_m t) \qquad x_Q(t) = \sin(2\pi f_m t)$$

$$X_I(f) = \delta(f - f_m) + \delta(f + f_m) \qquad X_Q(f) = \frac{1}{2j}\delta(f - f_m) - \frac{1}{2j}\delta(f + f_m)$$

$$X_z(f) = X_I(f) + jX_Q(f) = 1.5\delta(f - f_m) + 0.5\delta(f + f_m)$$

and using (2.13) gives the bandpass signal spectrum as

$$X_c(f) = \frac{1.5}{\sqrt{2}}\delta(f - f_c - f_m) + \frac{1}{2\sqrt{2}}\delta(f - f_c + f_m) + \frac{1.5}{\sqrt{2}}\delta(f + f_c + f_m) + \frac{1}{2\sqrt{2}}\delta(f + f_c - f_m).$$

Note in this example $B_T = 2f_m$.

Example 2.5: For the complex envelope derived in Example 2.3 the measured bandpass energy spectrum for f_c=7000Hz is shown in Fig. 2.10. Again the measured output is exactly predicted by (2.14). In this example we have $B_T = 5000Hz$.

2.4.2 Bandwidth of Bandpass Signals

The ideas of bandwidth of a signal extend in an obvious way to bandpass signals. Recall engineers define bandwidth as being the amount of positive spectrum that a signal occupies. For bandpass energy signals we have the following two definitions

Definition 2.2 *If a signal $x_c(t)$ has an energy spectrum $G_{x_c}(f)$ then B_X is determined as*

$$10\log\left(\max_f G_{x_c}(f)\right) = X + 10\log\left(G_{x_c}(f_1)\right) \tag{2.15}$$

where $G_{x_c}(f_1) > G_{x_c}(f)$ *for* $0 < f < f_1$ *and*

$$10\log\left(\max_f G_{x_c}(f)\right) = X + 10\log\left(G_{x_c}(f_2)\right) \tag{2.16}$$

where $G_{x_c}(f_2) > G_{x_c}(f)$ *for* $f > f_2$ *where* $f_2 - f_1 = B_X$.

Definition 2.3 *If a signal* $x_c(t)$ *has an energy spectrum* $G_{x_c}(f)$ *then* $B_P = \min(f_2 - f_1)$ *such that*

$$P = \frac{2\int_{f_1}^{f_2} G_{x_c}(f)df}{E_{x_c}} \tag{2.17}$$

where $f_2 > f_1$.

Note the reason for the factor of 2 in (2.20) is that half of the energy of the bandpass signal is associated with positive frequencies and half of the energy is associated with negative frequencies.

Again, for bandpass power signals similar ideas hold with $G_{x_c}(f)$ being replaced with $S_{x_c}(f,T)$.

Definition 2.4 *If a signal* $x_c(t)$ *has a sampled power spectral density* $S_{x_c}(f,T)$ *then* B_X *is determined as*

$$10\log\left(\max_f S_{x_c}(f,T)\right) = X + 10\log\left(S_{x_c}(f_1,T)\right) \tag{2.18}$$

where $S_{x_c}(f_1,T) > S_{x_c}(f,T)$ *for* $0 < f < f_1$ *and*

$$10\log\left(\max_f S_{x_c}(f,T)\right) = X + 10\log\left(S_{x_c}(f_2,T)\right) \tag{2.19}$$

where $S_{x_c}(f_2,T) > S_{x_c}(f,T)$ *for* $f > f_2$ *where* $f_2 - f_1 = B_X$.

Definition 2.5 *If a signal* $x_c(t)$ *has an power spectrum* $S_{x_c}(f,T)$ *then* $B_P = \min(f_2 - f_1)$ *such that*

$$P = \frac{2\int_{f_1}^{f_2} S_{x_c}(f,T)df}{P_{x_c}(T)} \tag{2.20}$$

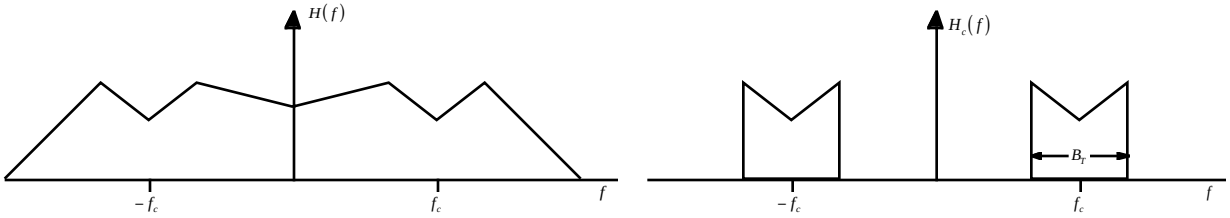
where $f_2 > f_1$.

2.5 Linear Systems and Bandpass Signals

This section discusses methods for calculating the output of a linear, time-invariant (LTI) filter with a bandpass input signal using complex envelopes. Linear system outputs are characterized by the convolution integral given as

$$y_c(t) = \int_{-\infty}^{\infty} x_c(\tau)h(t-\tau)d\tau \tag{2.21}$$

where $h(t)$ is the impulse response of the filter. Since the input signal is bandpass, the effects of an arbitrary filter, $h(t)$, in (2.21) can be modeled with an equivalent bandpass filter, $h_c(t)$, with no loss in generality. The bandpass filter, $H_c(f)$, only needs to equal the true filter, $H(f)$ over the frequency support of the the bandpass signal and the two filters need not be equal otherwise. Because of this characterization the bandpass filter is often simpler to model (and for instance simulate).



a) A lowpass filter. b) An input/output equivalent bandpass filter.

Figure 2.11: An example of a filter and its bandpass equivalent filter.

Example 2.6: For example consider the lowpass filter given in Fig. 2.11-a). Since the bandpass signal only has a non-zero spectrum in a bandwidth of B_T around the carrier frequency, f_c, the bandpass filter shown in Fig. 2.11-b) would be input–output equivalent to the filter in Fig. 2.11-a).

This bandpass LTI system also has a canonical representation given as

$$h_c(t) = 2h_I(t)\cos(2\pi f_c t) - 2h_Q(t)\sin(2\pi f_c t). \tag{2.22}$$

The complex envelope for this bandpass impulse response and transfer function associated with this complex envelope are given by

$$h_z(t) = h_I(t) + jh_Q(t) \qquad H_z(f) = H_I(f) + jH_Q(f)$$

where the bandpass system impulse response is

$$h_c(t) = 2\Re\left[h_z(t)\exp\left[j2\pi f_c t\right]\right].$$

The representation of the bandpass system in (2.22) has a constant factor of $\sqrt{2}$ difference from the bandpass signal representation of (2.1). This factor results because the system response at baseband and at bandpass should be identical. This notational convenience permits a simpler expression for the system output (as is shown shortly). Using similar techniques as in Section 2.4, the transfer function is expressed as

$$H_c(f) = H_z(f - f_c) + H_z^*(-f - f_c). \tag{2.23}$$

Example 2.7: Consider the signal

$$x_c(t) = (\cos(2\pi f_m t) + \cos(6\pi f_m t))\cos(2\pi f_c t) - (\sin(2\pi f_m t) + \sin(6\pi f_m t))\sin(2\pi f_c t) \tag{2.24}$$

that is input into a bandpass filter with a transfer function of

$$H_c(f) = \begin{cases} 2 & f_c - 2f_m \leq |f| \leq f_c + 2f_m \\ 0 & \text{elsewhere.} \end{cases} \tag{2.25}$$

Since the frequency domain representation of $x_c(t)$ is

$$X_c(f) = \frac{1}{\sqrt{2}}\left[\delta\left(f - (f_c + f_m)\right) + \delta(f - (f_c + 3f_m) + \delta(f + (f_c + f_m) + \delta(f + (f_c + 3f_m)\right] \tag{2.26}$$

the output bandpass signal wil have the frequency domain representation of

$$Y_c(f) = \frac{2}{\sqrt{2}}\left[\delta(f - (f_c + f_m) + \delta(f + (f_c + f_m)\right] \tag{2.27}$$

The complex envelopes of the input and output signals are

$$x_z(t) = \exp\left[j2\pi f_m t\right] + \exp\left[j6\pi f_m t\right] \qquad y_z(t) = 2\exp\left[j2\pi f_m t\right] \tag{2.28}$$

consequently it makes sense to have

$$H_z(f) = \begin{cases} 2 & |f| \leq 2f_m \\ 0 & \text{elsewhere.} \end{cases} \tag{2.29}$$

and $H_c(f) = H_z(f - f_c) + H_z^*(-f - f_c)$.

Equation (2.23) and (2.13) combined with the convolution theorem of the Fourier transform produces an expression for the Fourier transform of $y_c(t)$ given as

$$Y_c(f) = X_c(f)H_c(f) = \frac{1}{\sqrt{2}}\left[X_z(f - f_c) + X_z^*(-f - f_c)\right]\left[H_z(f - f_c) + H_z^*(-f - f_c)\right].$$

Since both $X_z(f)$ and $H_z(f)$ only take values in $[-B_T/2, B_T/2]$, the cross terms in this expression will be zero and $Y_c(f)$ is given by

$$Y_c(f) = \frac{1}{\sqrt{2}}\left[X_z(f - f_c)H_z(f - f_c) + X_z^*(-f - f_c)H_z^*(-f - f_c)\right]. \tag{2.30}$$

Since $y_c(t)$ will also be a bandpass signal, it will also have a complex baseband representation. A comparison of (2.30) with (2.13) demonstrates the Fourier transform of the complex envelope of $y_c(t)$, $y_z(t)$, is given as

$$Y_z(f) = X_z(f)H_z(f).$$

Linear system theory produces the desired form

$$y_z(t) = \int_{-\infty}^{\infty} x_z(\tau)h_z(t - \tau)d\tau = x_z(t) * h_z(t). \tag{2.31}$$

Example 2.8: (Example 2.1 continued) The input signal and Fourier transform are

$$x_z(t) = 2\cos(2\pi f_m t) + j\sin(2\pi f_m t) \qquad X_z(f) = 1.5\delta(f - f_m) + 0.5\delta(f + f_m).$$

Assume a bandpass filter with

$$H_I(f) = \begin{cases} 2 & -2f_m \le f \le 2f_m \\ 0 & \text{elsewhere} \end{cases} \qquad H_Q(f) = \begin{cases} j\frac{f}{f_m} & -f_m \le f \le f_m \\ j & f_m \le f \le 2f_m \\ -j & -2f_m \le f \le -f_m \\ 0 & \text{elsewhere} \end{cases} \tag{2.32}$$

This produces

$$H_z(f) = \begin{cases} 2 - \frac{f}{f_m} & -f_m \le f \le f_m \\ 1 & f_m \le f \le 2f_m \\ 3 & -2f_m \le f \le -f_m \\ 0 & \text{elsewhere} \end{cases} \tag{2.33}$$

and now the Fourier transform of the complex envelope of the filter output is

$$Y_z(f) = H_z(f)X_z(f) = 1.5\delta(f - f_m) + 1.5\delta(f + f_m). \tag{2.34}$$

The complex envelope and bandpass signal are given as

$$y_z(t) = 3\cos(2\pi f_m t) \qquad y_c(t) = 3\cos(2\pi f_m t)\sqrt{2}\cos(2\pi f_c t) \tag{2.35}$$

In other words, convolving the complex envelope of the input signal with the complex envelope of the filter response produces the complex envelope of the output signal. The different scale factor was introduced in (2.22) so that (2.31) would have a familiar form. This result is significant since $y_c(t)$ can be derived by computing a convolution of baseband (complex) signals which is generally much simpler than computing the bandpass convolution. Since $x_z(t)$ and $h_z(t)$ are complex, $y_z(t)$ is given in terms of the I/Q components as

$$y_z(t) = y_I(t) + jy_Q(t) = [x_I(t) * h_I(t) - x_Q(t) * h_Q(t)] + j\,[x_I(t) * h_Q(t) + x_Q(t) * h_I(t)]\,.$$

Fig 2.12 shows the lowpass equivalent model of a bandpass system. The two biggest advantages of using the complex baseband representation are that it simplifies the analysis of communication systems and permits accurate digital computer simulation of filters and the effects on communication systems performance.

2.6 Conclusions

The complex baseband representation of bandpass signals permits accurate characterization and analysis of communication signals independent of the carrier frequency. This greatly simplifies the job of the communication system engineer. A linear system is often an accurate model for a communication system, even with the associated transmitter filtering, channel distortion, and receiver filtering. As demonstrated in Fig 2.13, the complex baseband methodology truly simplifies the models for a communication system performance analysis.

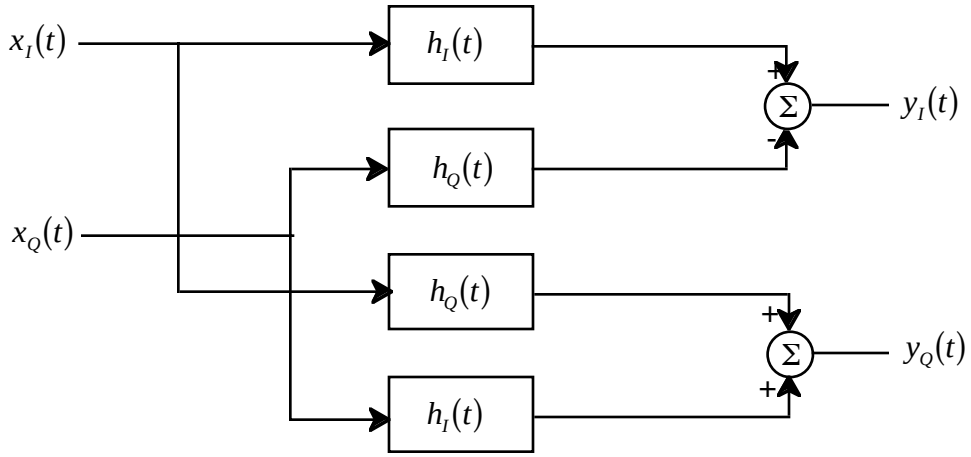


Figure 2.12: Block diagram illustrating the relation between the input and output complex envelope of bandpass signals for a linear time invariant system.

2.7 Homework Problems

Problem 2.1. Many integrated circuit implementations of the quadrature upconverters produce a bandpass signal having a form

$$x_c(t) = x_I(t)\sqrt{2}\cos(2\pi f_c t) + x_Q(t)\sqrt{2}\sin(2\pi f_c t) \tag{2.36}$$

from the lowpass signals $x_I(t)$ and $x_Q(t)$ as opposed to (2.1). How does this sign difference affect the transmitted spectrum? Specifically for the complex envelope energy spectrum given in Fig. 2.9 plot the transmitted bandpass energy spectrum.

Problem 2.2. Find the form of $x_I(t)$ and $x_Q(t)$ for the following $x_c(t)$

a) $x_c(t) = \sin(2\pi(f_c - f_m)t)$.

b) $x_c(t) = \cos(2\pi(f_c + f_m)t)$.

c) $x_c(t) = \cos(2\pi f_c t + \phi_p)$.

Problem 2.3. If the lowpass components for a bandpass signal are of the form

$$x_I(t) = 12\cos(6\pi t) + 3\cos(10\pi t)$$

and

$$x_Q(t) = 2\sin(6\pi t) + 3\sin(10\pi t).$$

a) Calculate the Fourier series of $x_I(t)$ and $x_Q(t)$.

b) Calculate the Fourier series of $x_z(t)$.

c) Assuming f_c=40Hz calculate the Fourier series of $x_c(t)$.

d) Calculate and plot $x_A(t)$. Computer might be useful.

e) Calculate and plot $x_P(t)$. Computer might be useful.

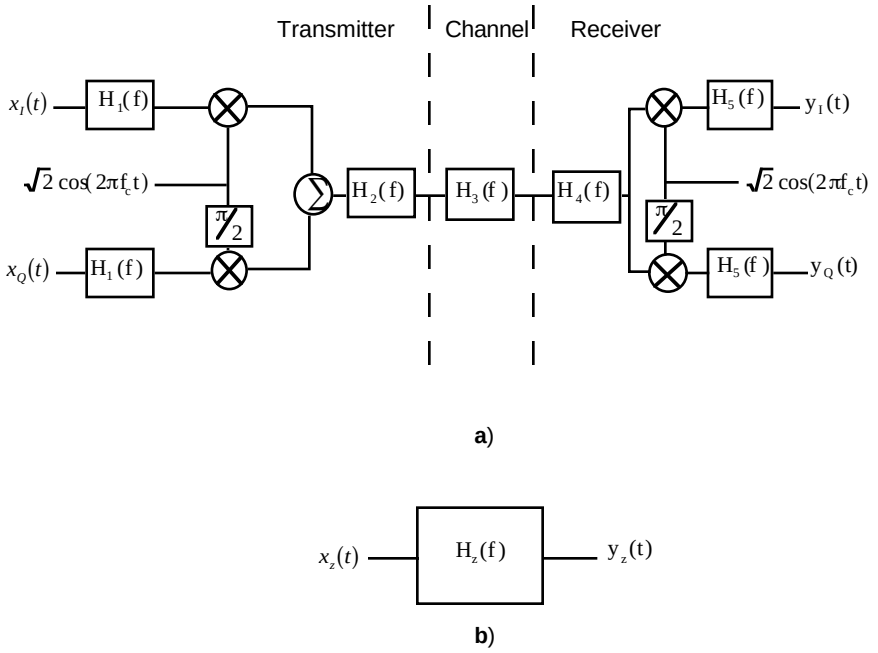


Figure 2.13: A comparison between **a**) the actual communication system model and **b**) the complex baseband equivalent model.

Problem 2.4. A bandpass filter has the following complex envelope representation for the impulse response

$$h_z(t) \;=\; \begin{cases} 2\left(\frac{1}{2}\exp\left[-\frac{t}{2}\right]\right) + j2\left(\frac{1}{4}\exp\left[-\frac{t}{4}\right]\right) & t \geq 0 \\ 0 & \text{elsewhere} \end{cases} \tag{2.37}$$

a) Calculate $H_z(f)$. *Hint: The transforms you need are in a table somewhere.*

b) With $x_z(t)$ from Problem 2.3 as the input, calculate the Fourier series for the filter output, $y_z(t)$.

c) Plot the output amplitude, $y_A(t)$, and phase, $y_P(t)$.

d) Plot the resulting bandpass signal, $y_c(t)$ using f_c=40Hz.

Problem 2.5. The picture of a color television set proposed by the National Television System Committee (NTSC) is composed by scanning in a grid pattern across the screen. The scan is made up of three independent beams (red, green, blue). These independent beams can be combined to make any color at a particular position. In order to make the original color transmission compatible with black and white televisions the three color signals ($x_r(t)$, $x_g(t)$, $x_b(t)$) are transformed into a luminance signal (black and white level), $x_L(t)$, and two independent chrominance signals, $x_I(t)$ and $x_Q(t)$. These chrominance signals are modulated onto a carrier of 3.58MHz to produce a bandpass signal for transmission. A commonly used tool for video engineers to understand these coloring patterns is the vectorscope representation shown in Figure 2.14.

a) If the video picture is making a smooth transition from a blue color (at t=0) to green color (at t=1), make a plot of the waveforms $x_I(t)$ and $x_Q(t)$.

b) Plot $x_I(t)$ and $x_Q(t)$ that would represent a scan across a red and green striped area. For consistency in the answers assume the red starts and t=0 and extends to $t = 1$, the green starts at $t = 1^+$ and extends to $t = 2, \cdots$.

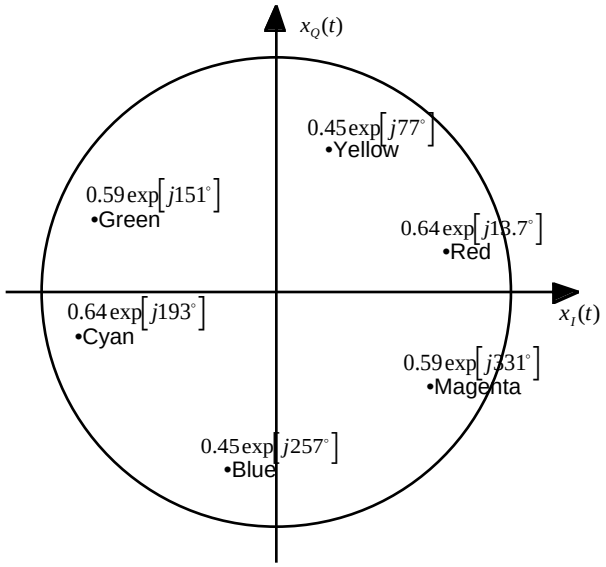


Figure 2.14: Vector scope representation of the complex envelope of the 3.58MHz chrominance carrier.

Problem 2.6. Consider two lowpass spectra, $X_I(f)$ and $X_Q(f)$ in Figure 2.15 and sketch the energy spectrum of the complex envelope, $G_{x_z}(f)$.

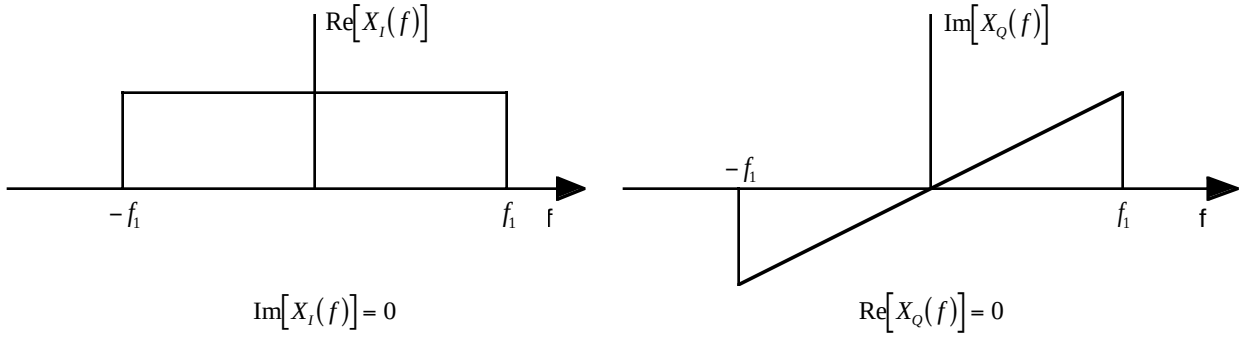


Figure 2.15: Two lowpass Fourier transforms.

Problem 2.7. (Design Problem) A key component in the quadrature up/down converter is the generator of the sine and cosine functions. This processing is represented in Figure 2.16 as a shift in the phase by 90° of a carrier signal. This function is done in digital processing in a trivial way but if the carrier is generated by an analog source the implementation is more tricky. Show that this phase shift can be generated with a time delay as in Figure 2.17. If the carrier frequency is 100MHz find the value of the delay to achieve the 90° shift.

Problem 2.8. The lowpass signals, $x_I(t)$ and $x_Q(t)$, which comprise a bandpass signal are given in Figure 2.18.

a) Give the form of $x_c(t)$, the bandpass signal with a carrier frequency f_c, using $x_I(t)$ and $x_Q(t)$.

b) Find the amplitude, $x_A(t)$, and the phase, $x_P(t)$, of the bandpass signal.

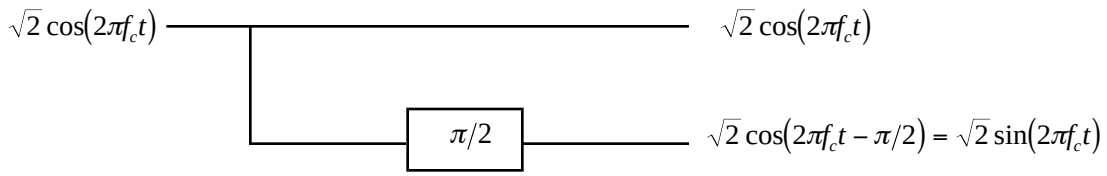


Figure 2.16: Sine and cosine generator.

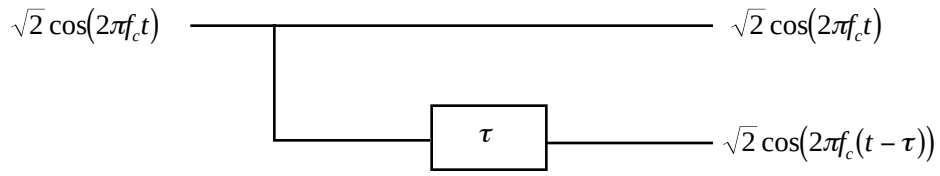


Figure 2.17: Sine and cosine generator implementation for analog signals.

c) Give the simplest form for the bandpass signal over $[2T, 3T]$.

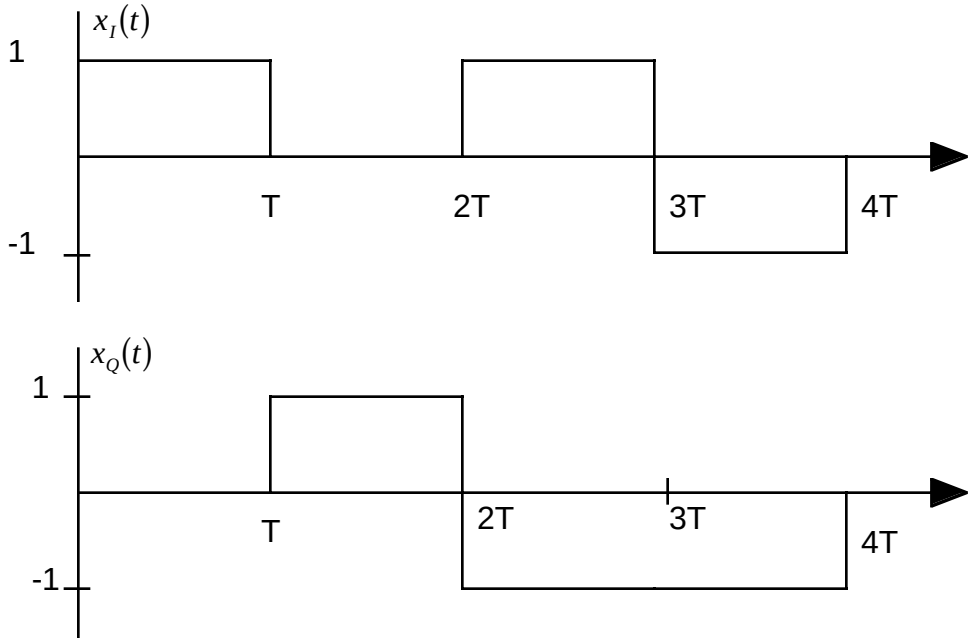


Figure 2.18: $x_I(t)$ and $x_Q(t)$.

Problem 2.9. The amplitude and phase of a bandpass signal is plotted in Figure 2.19. Plot the in-phase and quadrature signals of this baseband representation of a bandpass signal.

Problem 2.10. The block diagram in Fig. 2.20 shows a cascade of a quadrature upconverter and a quadrature downconverter where the phases of the two (transmit and receive) carriers are not the same. Show that $y_z(t) = y_I(t) + jy_Q(t) = x_z(t)\exp[-j\theta(t)]$. Specifically consider the case when the frequencies of the two carriers are not the same and compute the resulting output energy spectrum $G_{Y_z}(f)$.

Problem 2.11. A periodic real signal of bandwidth W and period T is $x_I(t)$ and $x_Q(t) = 0$ for a bandpass signal of carrier frequency $f_c > W$.

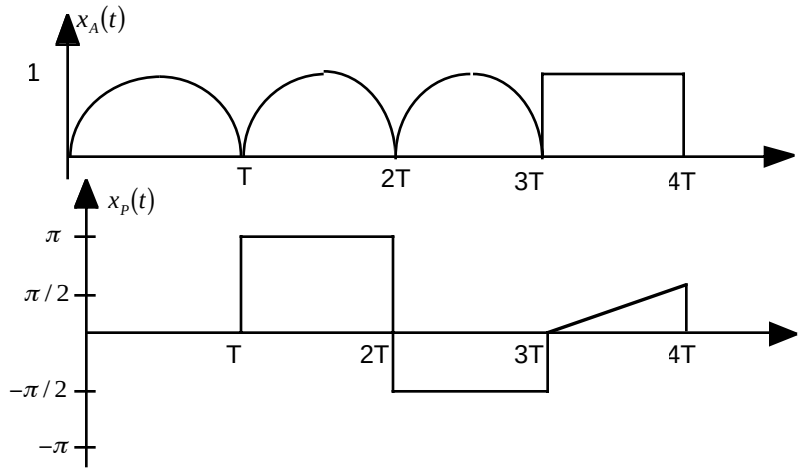


Figure 2.19: The amplitude and phase of a bandpass signal.

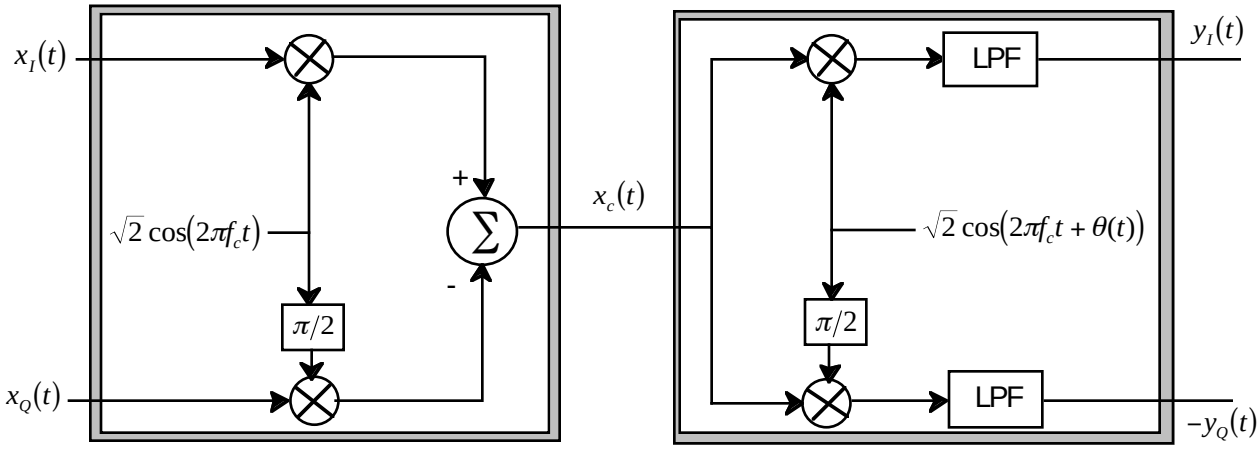


Figure 2.20: A downconverter with a phase offset.

a) Can the resulting bandpass signal, $x_c(t)$, be periodic with a period of $T_c < T$? If yes give an example.

b) Can the resulting bandpass signal, $x_c(t)$, be periodic with a period of $T_c > T$? If yes give an example.

c) Can the resulting bandpass signal, $x_c(t)$, be periodic with a period of $T_c = T$? If yes give an example.

d) Can the resulting bandpass signal, $x_c(t)$, be aperiodic? If yes give an example.

Problem 2.12. In communication systems bandpass signals are often processed in digital processors. To accomplish the processing, the bandpass signal must first be converted from an analog signal to a digital signal. For this problem assume this is done by ideal sampling. Assume the sampling frequency, f_s, is set at four times the carrier frequency.

a) Under what conditions on the complex envelope will this sampling rate be greater than the Nyquist sampling rate (see Section **??**) for the bandpass signal?

b) Give the values for the bandpass signal samples for $x_c(0)$, $x_c\left(\frac{1}{4f_c}\right)$, $x_c\left(\frac{2}{4f_c}\right)$, $x_c\left(\frac{3}{4f_c}\right)$, and $x_c\left(\frac{4}{4f_c}\right)$.

c) By examining the results in b) can you postulate a simple way to downconvert the analog signal when $f_s = 4f_c$ and produce $x_I(t)$ and $x_Q(t)$? This simple idea is frequently used in engineering practice and is known as $f_s/4$ downconversion.

Problem 2.13. A common implementation problem that occurs in an I/Q upconverter is that the sine carrier is not exactly 90° out of phase with the cosine carrier. This situation is depicted in Fig. 2.21.

a) What is the actual complex envelope, $x_z(t)$, produced by this implementation as a function of $\tilde{x}_I(t)$, $\tilde{x}_Q(t)$, and θ?

b) Often in communication systems it is possible to correct this implementation error by preprocessing the baseband signals. If the desired output complex envelope was $x_z(t) = x_I(t) + jx_Q(t)$ what should $\tilde{x}_I(t)$ and $\tilde{x}_Q(t)$ be set to as a function of $x_I(t)$, $x_Q(t)$, and θ to achieve the desired complex envelope with this implementation?

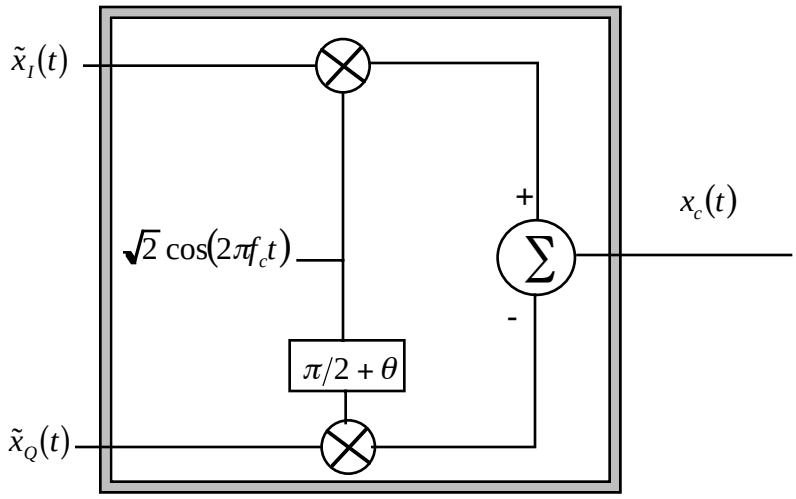


Figure 2.21: The block diagram for Problem 2.13.

Problem 2.14. A commercial airliner is flying 15,000 feet above the ground and pointing its radar down to aid traffic control. A second plane is just leaving the runway as shown in Fig. 2.22. The transmitted waveform is just a carrier tone, $x_z(t) = 1$ or $x_c(t) = \sqrt{2}\cos(2\pi f_c t)$

The received signal return at the radar receiver input has the form

$$y_c(t) = A_P\sqrt{2}\cos(2\pi(f_c + f_P)t + \theta_P) + A_G\sqrt{2}\cos(2\pi(f_c + f_G)t + \theta_G) \tag{2.38}$$

where the P subscript refers to the signal returns from the plane taking off and the G subscript refers to the signal returns from the ground. The frequency shift is due to the Doppler effect you learned about in your physics classes.

a) Why does the radar signal bouncing off the ground (obviously stationary) produce a Doppler frequency shift?

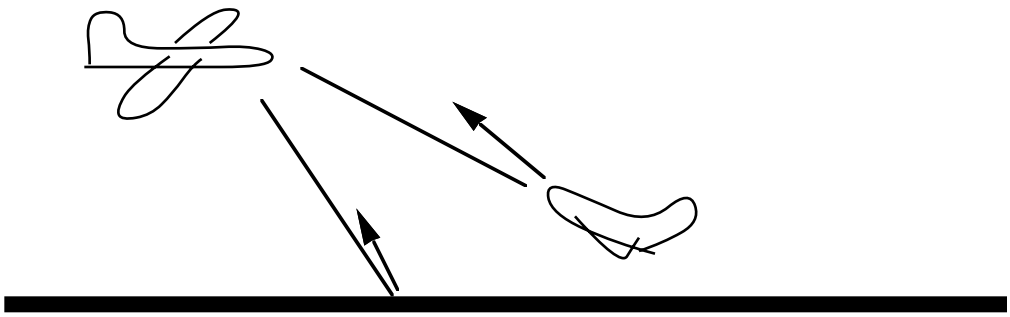


Figure 2.22: An airborne air traffic control radar example.

b) Give the complex baseband form of this received signal.

c) Assume the radar receiver has a complex baseband impulse response of

$$h_z(t) = \delta(t) + \beta\delta(t - T) \tag{2.39}$$

where β is a possibly complex constant, find the value of β which eliminates the returns from the ground at the output of the receiver. This system was a common feature in early radar systems and has the common name Moving Target Indicator (MTI) as stationary target responses will be canceled in the filter given in (2.39).

Modern air traffic control radars are more sophisticated than this problem suggests. An important point of this problem is that radar and communication systems are similar in many ways and use the same analytical techniques for design.

Problem 2.15. A baseband signal (complex exponential) and two linear systems are shown in Fig. 2.23. The top linear system in Fig. 2.23 has an impulse response of

$$h_z(t) \;=\; \begin{cases} \dfrac{1}{\sqrt{T_p}} & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{2.40}$$

The bottom linear system in Fig. 2.23 is an ideal delay element (i.e., $y_z(t) = x_z(t - \tau_d)$).

a) Give the bandpass frequency response $H_c(f)$.

b) What is the input power? Compute $y_z(t)$.

c) Select a delay, τ_d, in the bottom system in Fig. 2.23 such that $\arg[y_z(t)] = 2\pi f_0(t - \tau_d)$ for all f_0.

d) What is the output power as a function of f_0, $P_{y_z}(f_0)$?

e) How large can f_0 be before the output power, $P_{y_z}(f_0)$, is reduced by 10dB compared to the output power when $f_0 = 0$, $P_{y_z}(0)$?

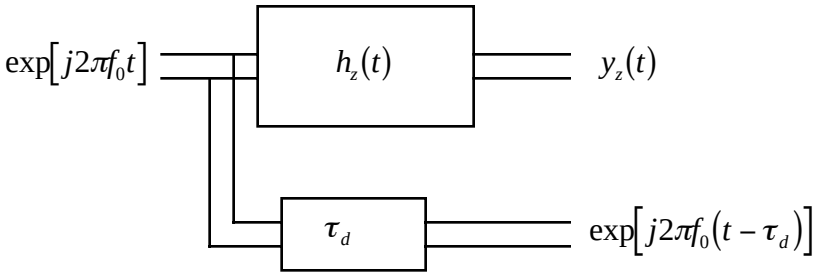


Figure 2.23: The block diagram for Problem 2.15.

Problem 2.16. The following bandpass filter has been implemented in a communication system that you have been tasked to simulate:

$$H_c(f) = \begin{cases} 1 & f_c + 7500 \leq |f| \leq f_c + 10000 \\ 2 & f_c + 2500 \leq |f| < f_c + 7500 \\ \frac{4}{3} & f_c \leq |f| < f_c + 2500 \\ \frac{3}{4} & f_c - 2500 \leq |f| < f_c \\ 0 & \text{elsewhere} \end{cases} \tag{2.41}$$

You know because of your great engineering education that it will be much easier to simulate the system using complex envelope representation.

a) Find $H_z(f)$.

b) Find $H_I(f)$ and $H_Q(f)$.

c) If $x_z(t) = \exp(j2\pi f_m t)$ find $x_c(t)$.

d) If $x_z(t) = \exp(j2\pi f_m t)$ compute $y_z(t)$ for $2000 \leq f_m < 9000$.

Problem 2.17. Consider two bandpass filters

$$h_{z1}(t) = \begin{cases} \frac{1}{\sqrt{0.2}} & 0 \leq t \leq 0.2 \\ 0 & \text{elsewhere,} \end{cases} \tag{2.42}$$

$$h_{z2}(t) = \frac{\sin(10\pi t)}{10\pi t} \tag{2.43}$$

Consider the filters and an input signal having a complex envelope of $x_z(t) = \exp(j2\pi f_m t)$.

a) Find $x_c(t)$.

b) Find $H_z(f)$.

c) Find $H_c(f)$.

d) For $f_m = 0, 7, 14$Hz find $y_z(t)$.

Problem 2.18. Find the amplitude signal, $x_A(t)$, and phase signal, $x_P(t)$ for

a) $x_z(t) = z(t)\exp(j\phi)$ where $z(t)$ is a complex valued signal.

b) $x_z(t) = m(t)\exp(j\phi)$ where $m(t)$ is a real valued signal.

Problem 2.19. A bandpass signal has a complex envelope given as

$$x_z(t) = j\exp\left[-j2\pi f_m t\right] + 3\exp\left[j2\pi f_m t\right] \tag{2.44}$$

where $f_m > 0$.

a) Find $x_I(t)$ and $x_Q(t)$.

b) Plot the frequency domain representation of this periodic signal using impulse functions.

c) Plot the frequency domain representation of the bandpass signal using impulse functions.

d) What is the bandpass bandwidth of this signal, B_T?

Problem 2.20. A bandpass signal has a complex envelope given as

$$x_z(t) = j\exp\left[-j2\pi f_m t\right] + 3\exp\left[j2\pi f_m t\right] \tag{2.45}$$

where $f_m > 0$. This signal is put into a bandpass filter which has a complex envelope characterized with

$$H_Q(f) = \begin{cases} 1 & |f| \leq 4000 \\ 0 & \text{elsewhere} \end{cases} \qquad H_I(f) = \begin{cases} -j & -4000 \leq f \leq 0 \\ j & 0 \leq f \leq 4000 \\ 0 & \text{elsewhere.} \end{cases} \tag{2.46}$$

The output of the filter at bandpass is denoted $y_c(t)$ and at baseband is denoted $y_z(t)$.

a) What is $H_z(f)$.

b) Find $y_z(t)$ as a function of f_m.

c) Plot the frequency domain representation of the bandpass signal using impulse functions for the case $f_m = 2000$Hz.

Problem 2.21. If $\mathcal{F}\{x_z(t)\} = X_z(f)$ what is $\mathcal{F}\{x_z^*(t)\}$?

Problem 2.22. An often used pulse in radar systems has a complex envelope of

$$x_z(t) = \begin{cases} A_c\exp\left[j2\pi g_0 t^2\right] & 0 \leq t \leq T_p \\ 0 & \text{elsewhere} \end{cases} \tag{2.47}$$

a) What are $x_A(t)$ and $x_P(t)$.

b) Plot the bandpass signal, $x_c(t)$ for $f_c = 10$Hz and $g_0 = 100$Hz/sec, $A_c = 1$, and $T_p = 1$.

c) What is the E_{x_z}?

d) Plot $G_{x_z}(f)$? Estimate B_{98} when $g_0 = 10$Hz/sec and T_p.

Problem 2.23. The amplitude and phase of a bandpass signal is plotted in Figure 2.24. Plot the in-phase and quadrature signals of this baseband representation of a bandpass signal.

Problem 2.24. A complex baseband signal is given as $x_z(t) = x_I(t) + jx_Q(t)$ where $x_I(t)$ and $x_Q(t)$ are real lowpass signals. Find

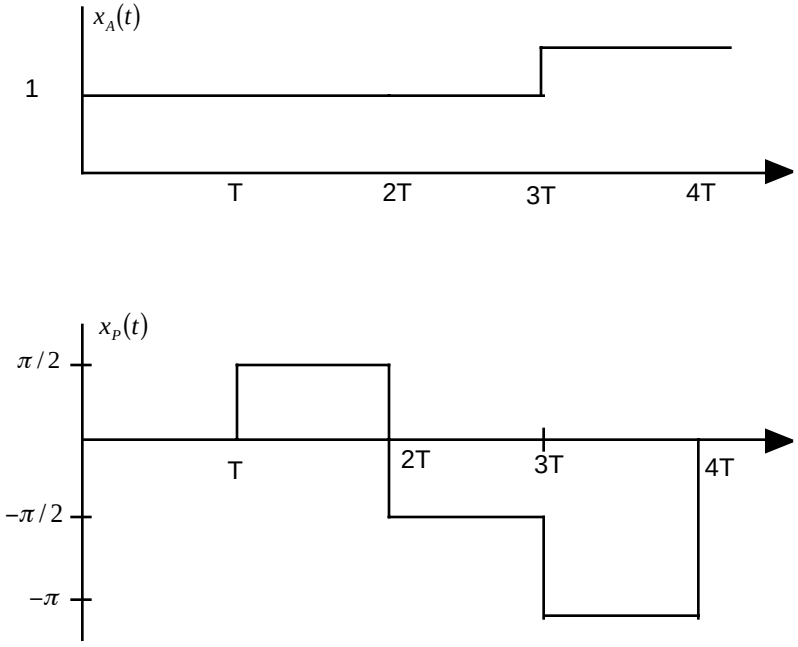


Figure 2.24: The amplitude and phase of a bandpass signal.

a) The bandpass signal, $x_c(t)$, represented by $x_z(t)$,

b) $X_z(f)$ and $X_z^*(f)$ in terms of $X_I(f)$ and $X_Q(f)$,

c) $X_Q(f)$ in terms of $X_z(f)$,

d) $X_c(f)$ in terms of $X_I(f)$ and $X_Q(f)$,

e) Show that $|X_c(f)|$ is an even function of frequency.

2.8 Example Solutions

Problem 2.2.

a) Using $\sin(a-b) = \sin(a)\cos(b) - \cos(a)\sin(b)$ gives

$$x_c(t) = \sin(2\pi f_c t)\cos(2\pi f_m t) - \cos(2\pi f_c t)\sin(2\pi f_m t). \tag{2.48}$$

By inspection we have

$$x_I(t) = \frac{-1}{\sqrt{2}}\sin(2\pi f_m t) \qquad x_Q(t) = \frac{-1}{\sqrt{2}}\cos(2\pi f_m t). \tag{2.49}$$

b) Recall $x_c(t) = x_A(t)\sqrt{2}\cos(2\pi f_c t + x_P(t))$ so by inspection we have

$$x_z(t) = \frac{1}{\sqrt{2}}\exp(j2\pi f_m t) \qquad x_I(t) = \frac{1}{\sqrt{2}}\cos(2\pi f_m t) \qquad x_Q(t) = \frac{1}{\sqrt{2}}\sin(2\pi f_m t). \tag{2.50}$$

c) Recall $x_c(t) = x_A(t)\sqrt{2}\cos(2\pi f_c t + x_P(t))$ so by inspection we have

$$x_z(t) = \frac{1}{\sqrt{2}}\exp(j\phi_p) \qquad x_I(t) = \frac{1}{\sqrt{2}}\cos(\phi_p) \qquad x_Q(t) = \frac{1}{\sqrt{2}}\sin(\phi_p). \tag{2.51}$$

Problem 2.6.
We can write

$$X_I(f) = A\text{rect}\left(\frac{f}{2f_1}\right) \qquad \text{and} \qquad X_Q(f) = j\frac{Af}{f_1}\text{rect}\left(\frac{f}{2f_1}\right)$$

Then

$$\begin{aligned}
G_{X_I}(f) &= |X_I(f)|^2 = A^2\text{rect}\left(\frac{f}{2f_1}\right) \\
G_{X_Q}(f) &= |X_Q(f)|^2 = \frac{A^2f^2}{f_1^2}\text{rect}\left(\frac{f}{2f_1}\right) \\
X_I(f)X_Q^*(f) &= -\frac{jA^2f}{f_1}\text{rect}\left(\frac{f}{2f_1}\right) \\
\Im X_I(f)X_Q^*(f) &= -\frac{A^2f}{f_1}\text{rect}\left(\frac{f}{2f_1}\right)
\end{aligned}$$

Hence

$$\begin{aligned}
G_{X_z}(f) &= G_{X_I}(f) + G_{X_Q}(f) + 2\Im X_I(f)X_Q^*(f) \\
&= A^2\left[\text{rect}\left(\frac{f}{2f_1}\right) + \left(\frac{f}{f_1}\right)^2\text{rect}\left(\frac{f}{2f_1}\right) - 2\frac{f}{f_1}\text{rect}\left(\frac{f}{2f_1}\right)\right] \\
G_{X_z}(f) &= \left(f - \frac{f}{f_1}\right)^2\text{rect}\left(\frac{f}{2f_1}\right).
\end{aligned}$$

Problem 2.11. $x_I(t)$ is periodic with period T. Because of this $x_I(t)$ can be represented in a Fourier series expansion

$$x_I(t) = \sum_{k=-\infty}^{\infty} x_k \exp\left[\frac{j2\pi kt}{T}\right] \tag{2.52}$$

If the bandwidth of the signal is less than WHz then the Fourier series will be truncated to a finite summation. Define k_m to be the largest integer such that $k_m/T \le W$ then

$$x_I(t) = \sum_{k=-k_m}^{k_m} x_k \exp\left[\frac{j2\pi kt}{T}\right]. \tag{2.53}$$

The bandpass signal will have the form

$$\begin{aligned}
x_c(t) &= x_I(t)\sqrt{2}\cos(2\pi f_c t) \\
&= x_I(t)\left[\frac{1}{\sqrt{2}}\exp(j2\pi f_c t) + \frac{1}{\sqrt{2}}\exp(-j2\pi f_c t)\right] \\
&= \sum_{k=-k_m}^{k_m}\frac{1}{\sqrt{2}}x_k\left(\exp\left[j2\pi\left(\frac{k}{T} + f_c\right)t\right] + \exp\left[j2\pi\left(\frac{k}{T} - f_c\right)t\right]\right).
\end{aligned} \tag{2.54}$$

Since the bandpass signal has a representation as a sum of weighted sinusoids there is a possibility that the bandpass signal will be periodic. This bandpass signal will only be period if all the frequencies are an integer multiple of a fundamental frequency, $1/T_c$.

a) The bandpass signal can be periodic with $T_c < T$. For example choose $x_I(t)$ to be a 2Hz sinusoid, i.e., $x_I(t) = \cos(2\pi(2)t)$ and the carrier frequency to be $f_c = 6$Hz. Clearly here $T = 1/2$ seconds. The bandpass signal in this case is

$$\begin{aligned} x_c(t) &= x_I(t)\sqrt{2}\cos(2\pi(6)t) \\ &= \frac{\sqrt{2}}{4}\left(\exp[j2\pi(-8)t] + \exp[j2\pi(-4)t] + \exp[j2\pi(4)t] + \exp[j2\pi(8)t]\right). \end{aligned} \quad (2.55)$$

The bandpass signal is a sum of four sinusoids that have frequencies of $f_1 = -8$, $f_2 = -4$, $f_3 = 4$, $f_4 = 8$. Clearly the fundamental frequency bandpass signal is 4Hz and $T_c = 0.25 < 0.5$.

b) The bandpass signal can be periodic with $T_c < T$. For example choose $x_I(t)$ to be a 3Hz sinusoid, i.e., $x_I(t) = \cos(2\pi(3)t)$ and the carrier frequency to be $f_c = 5$Hz. Here $T = 1/3$ seconds. The bandpass signal in this case is

$$\begin{aligned} x_c(t) &= x_I(t)\sqrt{2}\cos(2\pi(5)t) \\ &= \frac{\sqrt{2}}{4}\left(\exp[j2\pi(-8)t] + \exp[j2\pi(-2)t] + \exp[j2\pi(2)t] + \exp[j2\pi(8)t]\right). \end{aligned} \quad (2.56)$$

The bandpass signal is a sum of four sinusoids that have frequencies of $f_1 = -8$, $f_2 = -2$, $f_3 = 2$, $f_4 = 8$. The fundamental frequency of the bandpass signal is 2Hz and $T_c = 0.5 > 1/3$.

c) The bandpass signal can be periodic with $T_c = T$. For example choose $x_I(t)$ to be a 2Hz sinusoid, i.e., $x_I(t) = \cos(2\pi(2)t)$ and the carrier frequency to be $f_c = 4$Hz. Here $T = 0.5$ seconds. The bandpass signal in this case is

$$\begin{aligned} x_c(t) &= x_I(t)\sqrt{2}\cos(2\pi(4)t) \\ &= \frac{\sqrt{2}}{4}\left(\exp[j2\pi(-6)t] + \exp[j2\pi(-2)t] + \exp[j2\pi(2)t] + \exp[j2\pi(6)t]\right). \end{aligned} \quad (2.57)$$

The bandpass signal is a sum of four sinusoids that have frequencies of $f_1 = -6$, $f_2 = -2$, $f_3 = 2$, $f_4 = 6$. The fundamental frequency of the bandpass signal is 2Hz and $T_c = 0.5$.

d) The bandpass signal can be aperiodic. For example again choose $x_I(t)$ to be a 2Hz sinusoid, i.e., $x_I(t) = \cos(2\pi(2)t)$. Here $T = 0.5$ seconds. The bandpass signal in this case is

$$\begin{aligned} x_c(t) &= x_I(t)\sqrt{2}\cos(2\pi(f_c)t) \\ &= \frac{\sqrt{2}}{4}\left(\exp[j2\pi(-2-f_c)t] + \exp[j2\pi(-f_c+2)t] + \exp[j2\pi(f_c-2)t] + \exp[j2\pi(f_c+2)t]\right). \end{aligned} \quad (2.58)$$

The bandpass signal will be aperiodic if the two frequency are not integer multiples of a common frequency. That implies that the two bandpass frequencies must be irrational numbers. Choosing $f_c = \sqrt{2}$ will produce an aperiodic signal.

2.9 Mini-Projects

Goal: To give exposure

1. to a small scope engineering design problem in communications

2. to the dynamics of working with a team

3. to the importance of engineering communication skills (in this case oral presentations).

Presentation: The forum will be similar to a design review at a company (only much shorter) The presentation will be of 5 minutes in length with an overview of the given problem and solution. The presentation will be followed by questions from the audience (your classmates and the professor). All team members should be prepared to give the presentation.

Project 2.1. In engineering often in the course of system design or test anomalous performance characteristics often arise. Your job as an engineer is to identify the causes of these characteristics and correct them. Here is an example of such a case.

Get the Matlab file `bpex1.m` from the class web page. In this file the carrier frequency was chosen as 7kHz. If the carrier frequency is chosen as 8kHz an anomalous output is evident from the quadrature downconverter. This is most easily seen in the output energy spectrum, $G_{y_z}(f)$. Postulate a reason why this behavior occurs. *Hint:* It happens at 8kHz but not at 7kHz and Matlab is a sampled data system. What problems might one have in sampled data system? Assume that this upconverter and downconverter were products you were designing how would you specify the performance characteristics such that a customer would never see this anomalous behavior?

Project 2.2. A signal has a form

$$x_c(t) = x_I(t)\sqrt{2}\cos(2\pi f_c t + \theta) - x_Q(t)\sqrt{2}\sin(2\pi f_c t + \theta) \tag{2.59}$$

It is known that

$$x_I(t) = \cos(200\pi t) \qquad x_I(t) = \cos(400\pi t)\,. \tag{2.60}$$

The values of θ and f_c are unknown. The signal has been recorded with a sample frequency of 22050Hz (assume Nyquist criterion was satisfied) and is available in the file `CompenvPrjt2.mat` on the class web page. Using this file make an estimate of θ and f_c. Detail the logic that led to the estimates.

Chapter 3

Random Processes

An additive noise is characteristic of almost all communication systems. This additive noise typically arises from thermal noise generated by random motion of electrons in the conductors comprising the receiver. In a communication system the thermal noise having the greatest effect on system performance is generated at and before the first stage of amplification. This point in a communication system is where the desired signal takes the lowest power level and consequently the thermal noise has the greatest impact on the performance. This characteristic is discussed in more detail in Chapter 4. This chapter's goal is to introduce the mathematical techniques used by communication system engineers to characterize and predict the performance of communication systems in the presence of this additive noise. The characterization of noise in electrical systems could comprise a course in itself and often does at the graduate level. Textbooks that provide more detailed characterization of noise in electrical systems are [DR87, Hel91, LG89, Pap84]

A canonical problem formulation needed for the analysis of the performance of a communication systems design is given in Fig. 3.1. The thermal noise generated within the receiver is denoted $W(t)$. This noise is then processed in the receiver and will experience some level of filtering, represented with the transfer function $H_R(f)$. The simplest analysis problem is examining a particular point in time, t_s and characterize the resulting noise sample, $N(t_s)$, to extract a parameter of interest (e.g., average signal–to–noise ratio (SNR)). Additionally we might want to characterize two or more samples, e.g., $N(t_1)$ and $N(t_2)$, output from this filter.

To accomplish this analysis task this chapter first characterizes the thermal noise, $W(t)$. It turns out that $W(t)$ is accurately characterized as a stationary, Gaussian, and white random process. Consequently our first task is to define a random process (see Section 3.1). The exposition of the characteristics of a Gaussian random process (Section 3.2) and a stationary Gaussian random process (Section 3.3) then will follow. A brief discussion of the characteristics of thermal noise is then followed by an analysis of stationary random processes and linear systems. In conclusion, we return and solve the canonical problem posed in Fig. 3.1 in Section 3.6.

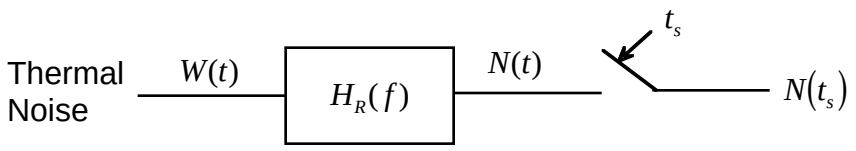


Figure 3.1: The canonical communications noise formulation.

3.1 Basic Definitions

Understanding random processes is fundamental to communications engineering. Random or stochastic processes are indexed families of random variables where the index is normally associated with different time instances.

Definition 3.1 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A real random process, $X(\omega, t)$, is a single-valued function or mapping from Ω to real valued functions of an index set variable t.*

The index set in this text will always be time but generalizations of the concept of a random process to other index sets is possible (i.e., space in image processing). The idea behind the definition of a random process is shown in Fig. 3.2. Essentially there is an underlying random experiment. The outcome of this random experiment is then mapped to a function of time. The example in Fig. 3.2 shows the mapping of three of the possible experiment outcomes into three different functions of time. The function $X(\omega_1, t)$ for ω_1 fixed is called a **sample path** of the random process. Communication engineers observe the sample paths of random processes and the goal of this chapter is to develop the tools to characterize these sample paths. From this point forward in the text the experimental outcome index will be dropped and random processes will be represented as $X(t)$.

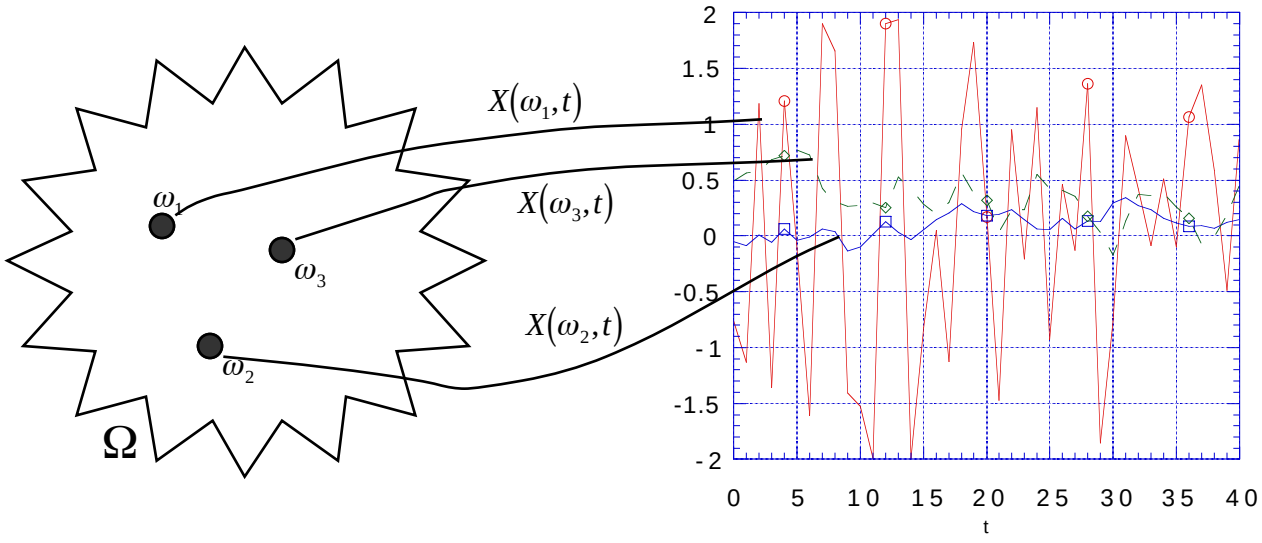


Figure 3.2: A pictorial explanation of the definition of a random process .

Property 3.1 *A sample of a random process is a random variable.*

If we focus on a random process at a particular time instance then we have a mapping from a random experiment to a real number. This is exactly the definition of a random variable (see Section **??**). Consequently, the first important question in the canonical problem stated at the beginning of this chapter has been answered: *the sample $N(t_s)$ of the random process, $N(t)$, is a random variable.* Consequently to characterize this sample of a random process we use all the tools that characterize a random variable (distribution function, density function, etc.). If random processes are indexed sets of random variables then a complete statistical characterization of the random process would be available if the random variables comprising the random process were completely characterized. In other words

for a random process, $N(t)$, if M samples of the random process are taken at various times and if the PDF given by

$$f_{N(t_1),N(t_2),\cdots,N(t_M)}(n_1, n_2, \cdots, n_M)$$

for arbitrary M is known then the random process is completely characterized. This full characterization, in general, is very difficult to do. Fortunately, the case of the Gaussian random process is the exception to the rule since Gaussian random processes are very often accurate models for noise in communication systems.

Point 1: The canonical problem can be solved when the random variable $N(t_s)$ or set of random variables, $N(t_1)$ and $N(t_2)$, can be statistically characterized with a PDF or a CDF

3.2 Gaussian Random Processes

Definition 3.2 *A Gaussian random process is a random process where any set of samples taken from the random process are jointly Gaussian random variables.*

Many important random processes in communication system analysis are well modeled as Gaussian random processes. This is important since Gaussian random variables are simply characterized. A complete characterization of Gaussian random variables (their joint PDF) is obtained by knowing their first and second order moments.

Property 3.2 *If $N(t)$ is a Gaussian random process then one sample of this process, $N(t_s)$, is completely characterized with the PDF*

$$f_{N(t_s)}(n_s) = \frac{1}{\sqrt{2\pi\sigma_N^2(t_s)}} \exp\left(-\frac{(n_s - m_N(t_s))^2}{2\sigma_N^2(t_s)}\right) \tag{3.1}$$

where

$$m_N(t_s) = E\left[N(t_s)\right] \qquad \sigma_N^2(t_s) = var\left(N(t_s)\right).$$

Consequently the complete characterization of one sample of a random process only requires the mean value and the variance of $N(t_s)$ to be evaluated. Section **??** has more discussion a PDF of a Gaussian random variable.

Property 3.3 *Most thermally generated noise corrupting a communication system typically has a zero mean.*

This is motivated by the physics of thermal noise generation. If the average voltage or current in a conductor is not zero that means that electrons are on average leaving the device. Physically this non-zero mean is not possible without an external force to induce an average voltage or current flow. **All random processes will be assumed to have a zero mean throughout the remainder of this chapter**.

Property 3.4 *If $N(t)$ is a Gaussian random process then two samples of this process, $N(t_1)$ and $N(t_2)$, are completely characterized with the PDF*

$$\begin{aligned} f_{N(t_1)N(t_2)}(n_1, n_2) &= \frac{1}{2\pi\sqrt{\sigma_N^2(t_1)\sigma_N^2(t_2)(1-\rho_N(t_1,t_2)^2)}} \times \\ &\quad \exp\left[\frac{-1}{1-\rho_N(t_1,t_2)^2}\left(\frac{n_1^2}{2\sigma_N^2(t_1)} - \frac{2\rho_N(t_1,t_2)n_1n_2}{2\sigma_N(t_1)\sigma_N(t_2)} + \frac{n_2^2}{2\sigma_N^2(t_2)}\right)\right] \end{aligned}$$

where

$$\sigma_N^2(t_i) = E\left[N^2(t_i)\right] \qquad \rho_N(t_1,t_2) = \frac{E\left[N(t_1)N(t_2)\right]}{\sigma_N(t_1)\sigma_N(t_2)}.$$

The joint PDF of two Gaussian random variables can be characterized by evaluating the variance of each of the samples and the correlation coefficient between the two samples. Recall a correlation coefficient describes how similar two random variables behave. The two random variables in this case are samples from the same random process. Section **??** has more discussion a PDF of a bivariate Gaussian random variable. Similarly three or more samples of a Gaussian random process can be characterized by evaluating the variance of each of the samples and the correlation coefficient between each of the samples [LG89, DR87].

Property 3.5 *The correlation function,*

$$R_N(t_1,t_2) = E\left[N(t_1)N(t_2)\right]. \tag{3.2}$$

contains all the information needed to characterize the joint distribution of any set of samples from a zero mean Gaussian random process.

Proof: Property 3.4 shows that only $\sigma^2(t)$ and $\rho(t_1,t_2)$ need to be identified to characterize the joint PDF of Gaussian random variables. To this end

$$\sigma_N^2(t) = R_N(t,t) = E\left[N^2(t)\right] \qquad \rho_N(t_1,t_2) = \frac{R_N(t_1,t_2)}{\sqrt{R_N(t_1,t_1)R_N(t_2,t_2)}}. \qquad \square$$

The correlation function plays a key role in the analysis of Gaussian random processes.

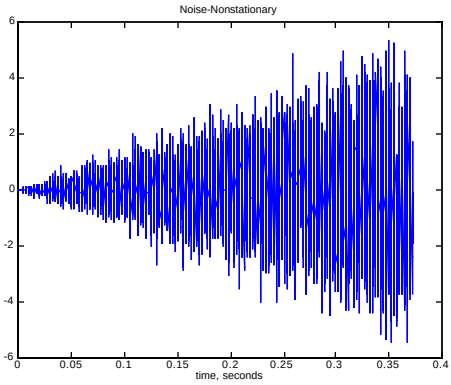
Point 2: When the noise, $N(t)$, is a Gaussian random process the canonical problem can be solved with knowledge of the correlation function, $R_N(t_1,t_2)$. The correlation function completely characterizes any joint PDF of samples of the process $N(t)$.

3.3 Stationary Random Processes

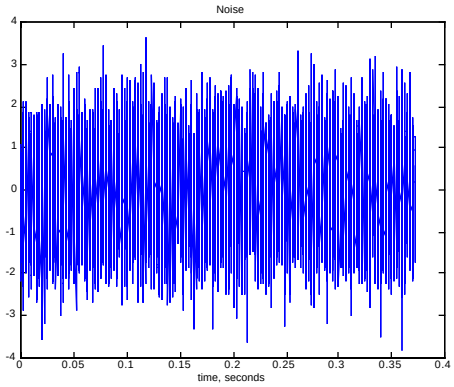
If the statistical description of a random process does not change over time, it is said to be a stationary random process. For example Fig. 3.3-a) shows a sample function of a nonstationary noise and Fig. 3.3-b) shows a sample function of a stationary noise. The random process plotted in Fig. 3.3-a) appears to have a variance, $\sigma_N^2(t)$, that is growing with time, while the random process in Fig. 3.3-b) appears to have a constant variance. Examples of random processes which are not stationary are prevalent

- Temperature in Columbus, Ohio
- The level of the Dow-Jone's Industrial Average.

In these nonstationary random processes the statistical description of the resulting random variables will change greatly depending on when the process is sampled. For example the average temperature in Columbus, OH is significantly different in July than in January. Nonstationary random processes are much more difficult to characterize than stationary processes. Fortunately the thermal noise that is often seen in communication systems is well modeled as stationary over the time spans which are of interest in understanding communication system's performance.



a) Nonstationary noise



b) Stationary noise

Figure 3.3: Sample functions of random processes.

3.3.1 Basics

Definition 3.3 *A random process, $N(t)$, is called stationary if the density function describing M samples of the random process is independent of time shifts in the sample times, i.e.,*

$$f_{N(t_1),N(t_2),\cdots,N(t_M)}(n_1,n_2,\cdots,n_M) = f_{N(t_1+t_0),N(t_2+t_0),\cdots,N(t_M+t_0)}(n_1,n_2,\cdots,n_M)$$

for any value of M and t_0.

In particular the random variable obtained by sampling a stationary random processes is statistically identical regardless of the selected sample time, i.e., $f_{N(t)}(n_1) = f_{N(t+t_0)}(n_1)$. Also two samples taken from a stationary random process will have a statistical description that is a function of the time difference between the two time samples but not the absolute location of the time samples, i.e., $f_{N(0),N(t_1)}(n_1,n_2) = f_{N(t_0),N(t_1+t_0)}(n_1,n_2)$.

3.3.2 Gaussian Processes

Stationarity for Gaussian random processes implies a fairly simple complete description for the random process. Recall a Gaussian random process is completely characterized with $R_N(t_1,t_2) = E[N(t_1)N(t_2)]$. Since $f_{N(t)}(n_1) = f_{N(t+t_0)}(n_1)$ this implies that the mean ($E[N(t)] = 0$) and the variance ($\sigma_N^2(t_i) = \sigma_N^2$) are constants. Since $f_{N(0),N(t_1)}(n_1,n_2) = f_{N(t_0),N(t_1+t_0)}(n_1,n_2)$ we know from Property 3.4 that $R_N(0,t_1) = R_N(t_0,t_0+t_1)$ for all t_0. This implies that the correlation function is essentially a function of only one variable, $R_N(t_1,t_2) = R_N(\tau) = E[N(t)N(t-\tau)]$ where $\tau = t_1 - t_2$. For a stationary Gaussian process this correlation function $R_N(\tau)$ completely characterizes any statistical event associated with the noise $N(t)$.

The correlation function gives a description of how rapidly the random process is changing with time. For a stationary random process the correlation function can be expressed as

$$R_N(\tau) = \sigma_N^2 \rho_N(\tau) \tag{3.3}$$

where $\rho_N(\tau)$ is the correlation coefficient between two samples taken τ seconds apart. Recall that if $\rho_N(\tau) \approx 0$ then the two samples will behave statistically in an independent fashion. If $\rho_N(\tau) \approx 1$ then the two samples will be very close in value (statistically dependent).

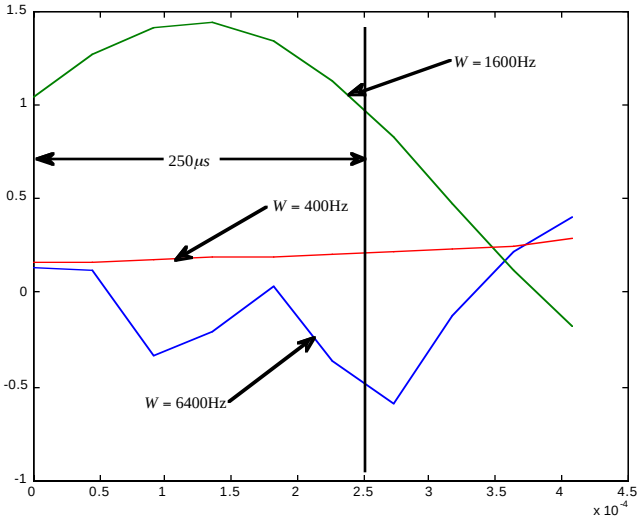


Figure 3.4: Sample functions for several Gaussian random processes parameterized by (3.5). Note this plot show samples of random processes where the sample rate is $f_s = 22050$Hz. The jaggedness and piecewise linear nature of the high bandwidth noise is due to this sampling. A higher sampling rate would result in a smoother appearance.

The idea of a correlation function has been introduced before in the context of deterministic signal analysis. For deterministic signals the correlation function of a signal $x(t)$ was defined as

$$V_x(\tau) = \int_{-\infty}^{\infty} x(t)x^*(t-\tau)dt \tag{3.4}$$

This correlation function also was a measure of how rapidly a signal varied in time just like the correlation function for random processes measures the time variations of a noise. Table 3.1 summarizes the comparison between the correlation functions for random processes and deterministic energy signals.

Point 3: When the noise, $N(t)$, is a stationary Gaussian random process the canonical problem can be solved solely by knowing the form of $R_N(\tau) = E\left[N(t)N(t-\tau)\right]$

Example 3.1: Consider a correlation coefficient parameterized as

$$\rho_N(\tau) = \text{sinc}\left(2W\tau\right). \tag{3.5}$$

For two samples taken τ=250μs apart, a process characterized with $W = 400$Hz would have these samples behaving very similar ($\rho_N(.00025) = 0.935$) while a process characterized with $W = 6400$Hz would have $\rho_N(.00025) = -0.058$ and essentially the two samples will be independent. Fig. 3.4 shows a plot of several sample functions of Gaussian random processes whose correlation function is given (3.5) with different values of W. The process parameterized with $W = 6400$Hz varies rapidly over 250μs while the process parameterized with $W = 400$Hz is nearly constant.

Signal	Definition	Units	Correlation value at $\tau = 0$
Deterministic	$V_x(\tau) = \int_{-\infty}^{\infty} x(t)x^*(t-\tau)dt$	Joules	$V_x(0) = E_x$ (Energy of $x(t)$)
Random	$R_N(\tau) = E[N(t)N(t-\tau)]$	Watts	$R_N(0) = \sigma_N^2$ (Average power of $N(t)$)

Table 3.1: Summary of the important characteristics of the correlation function.

3.3.3 Frequency Domain Representation

Stationary noise can also be described in the frequency domain. Frequency domain analysis has been very useful for providing an alternate view (compared to time domain) in deterministic signal analysis. This characteristic holds true for stationary random processes. Using the definition of a finite time Fourier transform given in (**??**), i.e.,

$$N_{T_m}(f) = \int_{-T_m}^{T_m} N(t) \exp[-j2\pi ft]\, dt, \tag{3.6}$$

the finite time average power spectral density (units Watts per Hertz) is given as

$$S_N(f, T_m) = \frac{1}{2T_m} E\left[|N_{T_m}(f)|^2\right]. \tag{3.7}$$

Definition 3.4 *The power spectral density of a random process $N(t)$ is*

$$S_N(f) = \lim_{T_m \to \infty} S_N(f, T_m).$$

Intuitively, random processes that vary rapidly must have power at higher frequencies in a very similar way as deterministic signals. This definition of the power spectral density puts that intuition on a solid mathematical basis. The power spectral density is a function which defines how the average power of a random process is distributed as a function of frequency. An additional relation to the correlation function can be made.

Property 3.6 *The power spectrum of a real valued*[1] *random process is always a non-negative and even function of frequency, i.e., $S_N(f) \geq 0$ and $S_N(f) = S_N(-f)$.*

Proof: The positivity of $S_N(f)$ is a direct consequence of Definition 3.4 and (3.7). The evenness is due to Property **??**. □

Property 3.7 (Wiener-Khinchin) *The power spectral density for a stationary random process is given by*

$$S_N(f) = \mathcal{F}\{R_N(\tau)\}.$$

Proof: The power spectral density (see Definition 3.4) can be rewritten as

$$S_N(f) = \lim_{T_m \to \infty} \frac{1}{2T_m} E\left[|N_{T_m}(f)|^2\right].$$

[1]This chapter only considers real valued random processes but the next chapter will extend these ideas to the complex envelope of a bandpass noise, hence the need to distinguish between a real and a complex random process.

Using the definition in (3.6) gives

$$\begin{aligned} E\left[|N_{T_m}(f)|^2\right] &= E\left[\int_{-T_m}^{T_m} N(t_1)\exp\left[-j2\pi f t_1\right]dt_1 \int_{-T_m}^{T_m} N(t_2)\exp\left[j2\pi f t_2\right]dt_2\right] \\ &= \int_{-T_m}^{T_m}\int_{-T_m}^{T_m} E\left[N(t_1)N(t_2)\right]\exp\left[-j2\pi f(t_1-t_2)\right]dt_1 dt_2. \end{aligned} \tag{3.8}$$

Making a change of variables $\tau = t_1 - t_2$ and using the stationarity of $N(t)$ reduces (3.8) to

$$E\left[|N_{T_m}(f)|^2\right] = \int_{-T_m}^{T_m}\int_{-T_m-t_2}^{T_m-t_2} R_N(\tau)\exp\left[-j2\pi f\tau\right]d\tau dt_2. \tag{3.9}$$

Taking the limit gives the desired result. □

Actually the true Wiener-Khinchin theorem is a bit more general but this form is all that is needed for this text.

The power spectral density can be used to find the average power of a random process with a direct analogy to Rayleigh's energy theorem. Recall that

$$R_N(\tau) = \mathcal{F}^{-1}\left\{S_N(f)\right\} = \int_{-\infty}^{\infty} S_N(f)\exp\left(j2\pi f\tau\right)df \tag{3.10}$$

so that the total average power of a random process is given as

$$R_N(0) = \sigma_N^2 = \int_{-\infty}^{\infty} S_N(f)df. \tag{3.11}$$

Equation (3.11) demonstrates that the average power of a random process is the area under the power spectral density curve.

Example 3.2: If

$$R_N(\tau) = \sigma_N^2 \mathrm{sinc}\left(2W\tau\right). \tag{3.12}$$

then

$$S_N(f) = \begin{cases} \dfrac{\sigma_N^2}{2W} & |f| \le W \\ 0 & \text{elsewhere} \end{cases} \tag{3.13}$$

A random process having this correlation–spectral density pair has power uniformly distributed over the frequencies from $-W$ to W.

The duality of the power spectral density of random signals, $S_N(f)$ and the energy spectral density of deterministic energy signals, $G_x(f)$, is also strong. Table 3.2 summarizes the comparison between the correlation functions for random processes and deterministic energy signals.

Point 4: When the noise, $N(t)$, is a stationary Gaussian random process the canonical problem can be solved with knowledge of the power spectral density of the random process, $S_N(f)$. The power spectral density determines the correlation function, $R_N(\tau)$ through an inverse Fourier transform. The correlation function completely characterizes any joint PDF of samples of the process $N(t)$.

Signal	Definition	Units	Correlation Value at $\tau = 0$	
Deterministic	$G_x(f) = \mathcal{F}\{V_x(\tau)\}$	Joules/Hz	$V_x(0) = E_x = \int_{-\infty}^{\infty} G_x(f)df$	$G_x(f) \geq 0$
Random	$S_N(f) = \mathcal{F}\{R_N(\tau)\}$	Watts/Hz	$R_N(0) = \sigma_N^2 = \int_{-\infty}^{\infty} S_N(f)df$	$S_N(f) \geq 0$

Table 3.2: Summary of the duality of spectral densities.

3.4 Thermal Noise

An additive noise is characteristic of almost all communication systems. This additive noise typically arises from thermal noise generated by random motion of electrons in conductors. From the kinetic theory of particles, given a resistor R at a physical temperature of T degrees Kelvin, there exists a random additive noise voltage, $W(t)$, across the resistor. This noise is zero mean, stationary[2], and Gaussian distributed (due to the large number of randomly moving electrons and the Central Limit Theorem), with a spectral density given as [Zie86]

$$S_W(f) = \frac{2Rh|f|}{\exp\left[\frac{h|f|}{KT}\right] - 1} \tag{3.14}$$

where $h = 6.62 \times 10^{-34}$ is Planck's constant and $K = 1.3807 \times 10^{-23}$ is Boltzmann's constant. The major characteristic to note is that this spectral density is constant up to about 10^{12} Hz at a value of $2KTR$. Since a majority of communications occurs at frequencies much lower than 10^{12} Hz[3] an accurate model of the thermal noise is to assume a flat spectral density (white noise). This spectral density is a function of the receiver temperature. The lower the temperature of the receiver the lower the noise power.

Additive white Gaussian noise (AWGN) is a basic assumption made about the corrupting noise in most communication systems performance analyses. White noise has a constant spectral density given as

$$S_W(f) = \frac{N_0}{2} \tag{3.15}$$

where $N_0/2$ is the noise power spectral density. Note that since

$$E[W^2(t)] = \int_{-\infty}^{\infty} S_W(f)df \tag{3.16}$$

this thermal noise model has an infinite average power. This characteristic is only a result of modeling (3.14) with the simpler form in (3.15). Traditionally some engineers have been loath to admit to the idea of negative frequencies hence they are not comfortable with the model in (3.15). These engineers like to think about noise power distributed over positive only frequencies with a spectral density of double the height. In the literature one will often see people denote $N_0/2$ as the two-sided noise spectral density and N_0 as the one-sided noise spectral density.

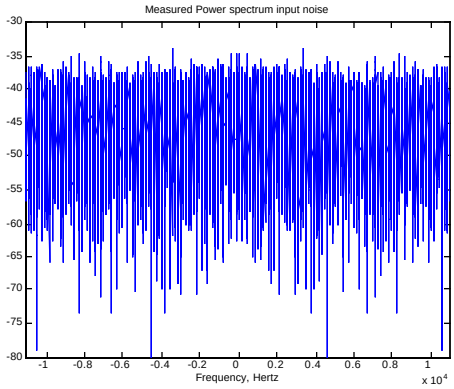
The autocorrelation function of AWGN is given by

$$R_W(\tau) = E\left[W(t)W(t-\tau)\right] = \frac{N_0}{2}\delta(\tau). \tag{3.17}$$

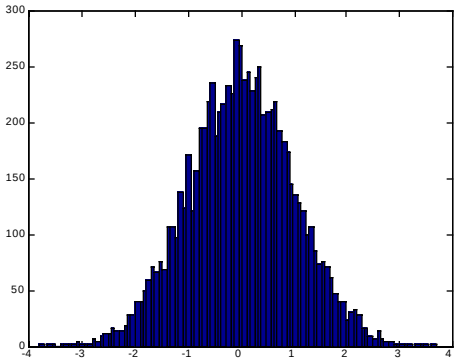
AWGN is a stationary random process so a great majority of the analysis in communication systems design can utilize the extensive theory of stationary random processes. White noise has the interesting

[2] As long as the temperature is constant.

[3] Fiber optic communication is a notable exception to this statement.



a) The measured PSD.



b) The measured histogram.

Figure 3.5: Characteristics of sampled wideband Gaussian noise.

characteristic that two samples of the process $W(t)$ will always be independent no matter how closely the samples are taken. Consequently white noise can be thought of as a very rapidly varying noise with no predictability.

A resistor in isolation will produce an approximate white noise with a two-sided spectral density of $2KTR$. If this system is placed in an electronic system the greatest value of the average power that would be able to be transferred out of the resistor is one fourth of value that would be exist if the resistor was shorted (using the maximum power transfer theorem of linear systems, see Problem 3.13). Hence an important number for communication system engineers to remember is $N_0 = KT$. For room temperature ($T = 290°K$) $N_0 = 4 \times 10^{-21}$ Watts/Hz (-174dBm/Hz). When signals get amplified by active components in a receiver the value of this noise spectral density will change. We explore some of these ideas in the homework and refer the interested student to the detailed references [Zie86, Cou93, Skl88] on the ideas of noise figure, link budgets, and computing output signal to noise ratio.

This noise spectral density at the input to a communication system is very small. If this thermal noise is to corrupt the communication signal then the signal must also take a small value. Consequently it is apparent from this discussion that thermal noise is only going to be significant when the signal at the receiver has been attenuated to a relatively small power level.

Example 3.3: The noise plotted in Fig. 3.3-b) could result from sampling an AWGN at a sample rate of $f_s = 22050$Hz. The rapidly varying characteristic of the noise samples predicted by (3.17) is evident from Fig. 3.3-b). The measured power spectrum (see (3.7)) for this process is shown in Fig. 3.5-a). This PSD, while extremely jagged due to the consideration of a finite time interval, clearly shows the whiteness of the noise. Fig. 3.5-a) show the histogram of the samples of the AWGN. The Gaussian nature of this histogram is also clearly evident.

Point 5: The dominant noise in communication systems is often generated by the random motion of electrons in conductors. This noise process is accurately modeled as stationary, Gaussian, and white.

3.5 Linear Systems and Random Processes

The final step in solving the canonical problem is to characterize the output of a linear time-invariant filter when the input is a white Gaussian stationary random process.

Property 3.8 *Random processes that result from linear filtering of Gaussian random processes are also Gaussian processes.*

Proof: The technical machinery needed to prove this property is beyond the scope of this class but most advanced texts concerning random processes will prove this. For example [DR87]. It should be noted that a majority of the important ideas can be understood by examining Problem **??**.24. □

Property 3.9 *A random process that results from linear filtering of a zero mean process will also be zero mean.*

Proof: Assume the filter input is $W(t)$, the filter impulse response is $h_R(t)$, and the filter output is $N(t)$.

$$E\left[N(t)\right] = E\left[\int_{-\infty}^{\infty} h_R(\lambda)W(t-\lambda)d\lambda\right] = \int_{-\infty}^{\infty} h_R(\lambda)E\left[W(t-\lambda)\right]d\lambda = 0. \qquad \square \tag{3.18}$$

Property 3.10 *A Gaussian random process that results from linear time-invariant filtering of another stationary Gaussian random process is also a stationary Gaussian process.*

Proof: Assume the filter input is $W(t)$, the filter impulse response is $h_R(t)$, and the filter output is $N(t)$. Since we know the output is Gaussian, the stationarity of $N(t)$ only requires that the correlation function of $N(t)$ be a function of τ only, i.e., $R_N(t_1,t_2) = R_N(\tau)$. The two argument correlation function of $N(t)$ is given as

$$\begin{aligned} R_N(t_1,t_2) &= E\left[N(t_1)N(t_2)\right] \\ &= E\left[\int_{-\infty}^{\infty} h_R(\lambda_1)W(\lambda_1-t_1)d\lambda_1 \int_{-\infty}^{\infty} h_R(\lambda_2)W(\lambda_2-t_2)d\lambda_2\right]. \end{aligned} \tag{3.19}$$

Rearranging gives

$$\begin{aligned} R_N(t_1,t_2) &= \int_{-\infty}^{\infty} h_R(\lambda_1)\int_{-\infty}^{\infty} h_R(\lambda_2)E\left[W(\lambda_1-t_1)W(\lambda_2-t_2)\right]d\lambda_1 d\lambda_2 \\ &= \int_{-\infty}^{\infty} h_R(\lambda_1)\int_{-\infty}^{\infty} h_R(\lambda_2)R_W(\lambda_1-t_1-\lambda_2+t_2)d\lambda_1 d\lambda_2 \\ &= \int_{-\infty}^{\infty} h_R(\lambda_1)\int_{-\infty}^{\infty} h_R(\lambda_2)R_W(\lambda_1-\lambda_2-\tau)d\lambda_1 d\lambda_2 \\ &= g(\tau). \qquad \square \end{aligned} \tag{3.20}$$

Point 6: The output noise in the canonical problem, $N(t)$, is accurately modeled as a stationary Gaussian process.

The canonical problem given in Fig. 3.1 can now be solved by finding the correlation function or spectral density of the stationary Gaussian random process $N(t)$. For the canonical problem $W(t)$ is a white noise so that (3.20) can be rewritten as

$$R_N(\tau) = \int_{-\infty}^{\infty} h_R(\lambda_1)\int_{-\infty}^{\infty} h_R(\lambda_2)\frac{N_0}{2}\delta(\lambda_1-\lambda_2-\tau)d\lambda_1 d\lambda_2. \tag{3.21}$$

Using the sifting property of the delta function gives

$$R_N(\tau) = \frac{N_0}{2}\int_{-\infty}^{\infty} h_R(\lambda_1)h_R(\lambda_1-\tau)d\lambda_1 = \frac{N_0}{2}V_{h_R}(\tau) \tag{3.22}$$

where $V_{h_R}(\tau)$ is the correlation function of the deterministic impulse response $h_R(t)$ as defined in Chapter **??**. This demonstrates that the output correlation function from a linear filter with a white noise input is only a function of the filter impulse response and the white noise spectral density.

The frequency domain description of the random process $N(t)$ is equally simple. The power spectral density is given as

$$S_N(f) = \mathcal{F}\{R_N(\tau)\} = \frac{N_0}{2}\mathcal{F}\{V_{h_R}(\tau)\} = \frac{N_0}{2}G_{h_R}(f) = \frac{N_0}{2}|H_R(f)|^2. \tag{3.23}$$

The average output noise power is given using either (3.22) or (3.23) as

$$\sigma_N^2 = E\left[N^2(t)\right] = R_N(0) = \frac{N_0}{2}\int_{-\infty}^{\infty}|H_R(f)|^2\,df = \frac{N_0}{2}\int_{-\infty}^{\infty}|h_R(t)|^2\,dt. \tag{3.24}$$

It should be noted that (3.24) is another example of a Parseval theorem result.

Example 3.4: Consider an ideal low pass filter of bandwidth W, i.e.,

$$H_R(f) = \begin{cases} 1 & |f| \le W \\ 0 & \text{elsewhere} \end{cases} \tag{3.25}$$

The output spectral density is given as

$$S_N(f) = \begin{cases} \dfrac{N_0}{2} & |f| \le W \\ 0 & \text{elsewhere} \end{cases} \tag{3.26}$$

and the average output power is given as $\sigma_N^2 = N_0 W$.

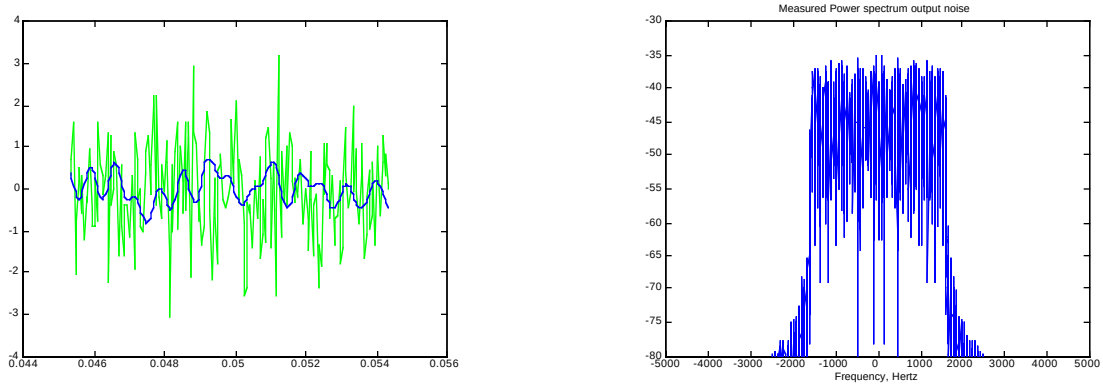
Engineers noted the simple form for the average output noise power expression in Example 3.4 and decided to parameterize the output noise power of all filters by a single number, the noise equivalent bandwidth.

Definition 3.5 *The noise equivalent bandwidth of a filter is*

$$B_N = \frac{1}{2|H_R(0)|^2}\int_{-\infty}^{\infty}|H_R(f)|^2\,df.$$

For an arbitrary filter the average output noise power will be $\sigma_N^2 = N_0 B_N |H_R(0)|^2$. Note for a constant $|H_R(0)|$ the smaller the noise equivalent bandwidth the smaller the noise power. This implies that noise power can be minimized by making B_N as small as possible. In the homework problems we will show that the output signal to noise ratio is not a strong function of $|H_R(0)|$ so that performance in the presence of noise is well parameterized by B_N. In general, $H_R(f)$ is chosen as a compromise between complexity of implementation, signal distortion, and noise mitigation.

Example 3.5: The sampled wideband Gaussian noise shown in Fig. 3.3-b) is put into a lowpass filter of bandwith 1600Hz. This lowpass filter will significantly smooth out the sampled noise and reduce the variance. Two output sample functions; one for the wideband noise and one for the filtered noise, are shown in Fig. 3.6-a). The lowering of the output power of the noise and the resulting smoother time variations are readily apparent in this figure. The output measured PSD of the filtered noise is shown in Fig. 3.6-b). This measured PSD validates (3.23), as this PSD is clearly the result of multiplying a white noise spectrum with a lowpass filter transfer function. The histogram of the output samples, shown in Fig. 3.7, again demonstrates that this noise is well modeled as a zero mean Gaussian random process.



a) Sample paths of wideband and filtered noise. b) Measured PSD.

Figure 3.6: Measured characteristics of filtered noise.

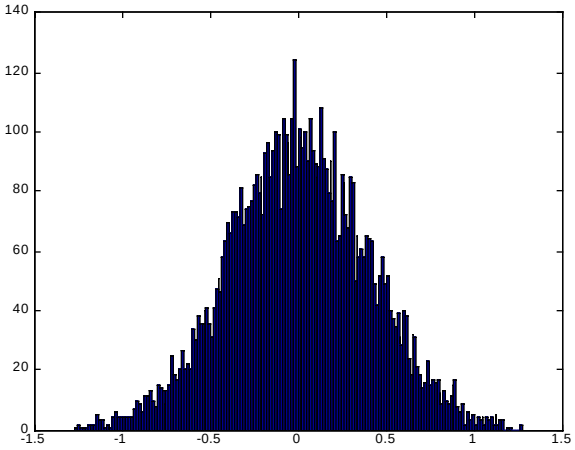


Figure 3.7: A histogram of the samples of the filtered noise.

Point 7. A stationary, white, and Gaussian noise of two sided spectral density $N_0/2$ when passed through a filter will produce a stationary Gaussian noise whose power spectral density is given as

$$S_N(f) = \frac{N_0}{2} \left| H_R(f) \right|^2 .$$

3.6 The Solution of the Canonical Problem

Putting together the results of the previous five sections leads to a solution of the canonical problem posed in this chapter.

First, the characterization of one sample, $N(t_s)$, of the random process $N(t)$ is considered. This case requires a four step process summarized as

1. Identify N_0 and $H_R(f)$.

2. Compute $S_N(f) = \frac{N_0}{2}\left|H_R(f)\right|^2$.

3. $\sigma_N^2 = E\left[N^2(t)\right] = R_N(0) = \frac{N_0}{2}\int_{-\infty}^{\infty}\left|H_R(f)\right|^2 df = N_0 B_N |H_R(0)|^2$

4. $f_{N(t_s)}(n_1) = \frac{1}{\sqrt{2\pi\sigma_N^2}}\exp\left(-\frac{n_1^2}{2\sigma_N^2}\right)$

Example 3.6: Consider two ideal lowpass filters with

$$H_{R(i)}(f) \;=\; \begin{cases} 1 & |f| \le W_i \\ 0 & \text{elsewhere} \end{cases} \tag{3.27}$$

with W_1=400Hz and W_2=6400Hz and a noise spectral density of

$$N_0 = \frac{1}{1600}.$$

Step 2 gives

$$S_{N_i}(f) \;=\; \begin{cases} \dfrac{1}{3200} & |f| \le W_i \\ 0 & \text{elsewhere.} \end{cases} \tag{3.28}$$

Step 3 gives $\sigma_{N_1}^2 = 0.25$ and $\sigma_{N_2}^2 = 4$. The average noise power between the two filter outputs is different by a factor of 16. The PDFs in step 4 are plotted in Fig. 3.8.

Second, the characterization of two samples, $N(t_1)$ and $N(t_2)$, of the random process $N(t)$ is considered. This case requires a five step process summarized as

1. Identify N_0 and $H_R(f)$.

2. Compute $S_N(f) = \frac{N_0}{2}\left|H_R(f)\right|^2$.

3. $R_N(\tau) = \mathcal{F}^{-1}\left\{S_N(f)\right\}$.

4. $\sigma_N^2 = R_N(0)$ and $\rho_N(\tau) = R_N(\tau)/\sigma_N^2$.

5. $f_{N(t_1)N(t_2)}(n_1, n_2) = \frac{1}{2\pi\sigma_N^2\sqrt{(1-\rho_N(\tau)^2)}}\exp\left[\frac{-1}{2\sigma_N^2(1-\rho_N(\tau)^2)}\left(n_1^2 - 2\rho_N(\tau)n_1n_2 + n_2^2\right)\right]$.

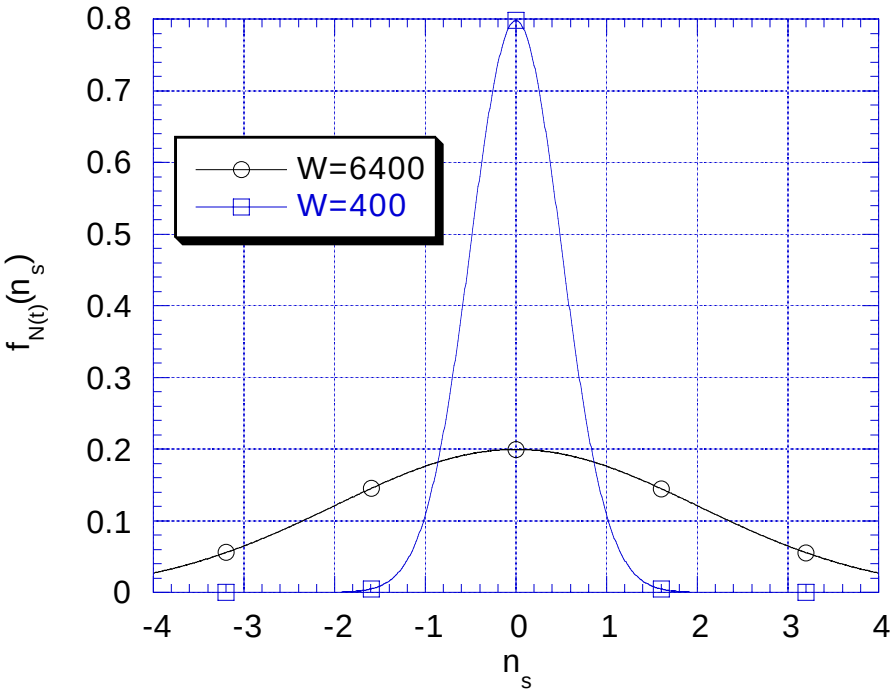


Figure 3.8: The PDF of a filter output sample for W=400 and W=6400Hz.

Example 3.7: Consider two ideal lowpass filters with

$$H_{R(i)}(f) = \begin{cases} 1 & |f| \leq W_i \\ 0 & \text{elsewhere} \end{cases} \tag{3.29}$$

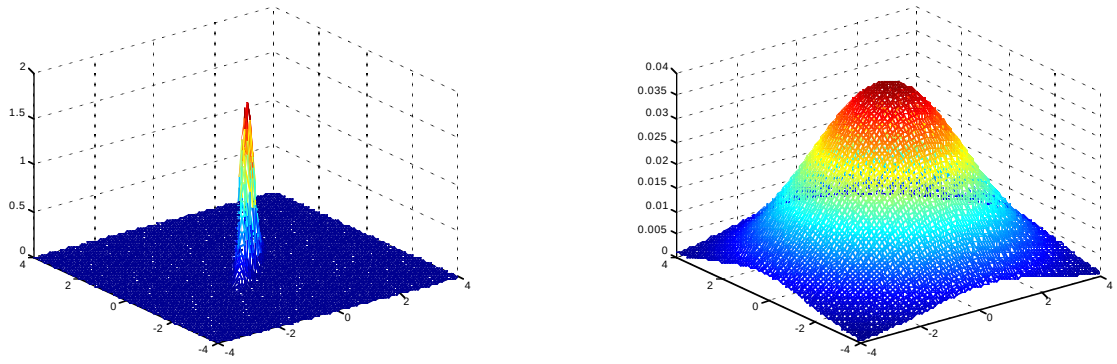
with W_1=400Hz and W_2=6400Hz and a noise spectral density of

$$N_0 = \frac{1}{1600}.$$

Step 2 gives

$$S_{N_i}(f) = \begin{cases} \frac{1}{3200} & |f| \leq W_i \\ 0 & \text{elsewhere.} \end{cases} \tag{3.30}$$

Step 3 gives $R_{N_1}(\tau) = 0.25\text{sinc}\,(800\tau)$ and $R_{N_2}(\tau) = 4\text{sinc}\,(12800\tau)$. Step 4 has the average noise power between the two filter outputs being different by a factor of 16 since $\sigma^2_{N_1} = 0.25$ and $\sigma^2_{N_2} = 4$. For an offset of τ seconds the correlation coefficients are $\rho_{N_1}(\tau) = \text{sinc}\,(800\tau)$ and $\rho_{N_2}(\tau) = \text{sinc}\,(12800\tau)$. The PDFs in step 5 are plotted in Fig. 3.9-a) and Fig. 3.9-b) for $\tau = 250\mu$ s. With the filter having a bandwidth of 400Hz the variance of the output samples is less and the correlation between samples is greater. This higher correlation implies that the probability that the two random variables take significantly different values is very small. This correlation is evident due to the knife edge like joint PDF. These two joint PDFs show the characteristics in an analytical form that were evident in the noise sample paths shown in Fig. 3.4.



a) W=400Hz. b) W=6400Hz.

Figure 3.9: The PDF of two samples separated by $\tau = 250\mu$s.

Similarly three or more samples of the filter output could be characterized in a very similar fashion. The tools developed in this chapter give a student the ability to analyze many of the important properties of noise that are of interest in communication system design.

3.7 Homework Problems

Problem 3.1. For this problem consider the canonical block diagram shown in Fig. 3.10 with $s_i(t) = \cos(2\pi(200)t)$ and with $W(t)$ being an additive white Gaussian noise with two-sided spectral density of $N_0/2$. Assume $H_R(f)$ is an ideal lowpass filter with bandwidth of 400Hz. Two samples are taken from the filter output at time $t = 0$ and $t = \tau$, i.e., $Y_o(0)$ and $Y_o(\tau)$.

a) Choose N_0 such that the output noise power, $E\left[N^2(t)\right] = 1$.

b) $Y_o(0)$ is a random variable. Compute $E\left[Y_o(0)\right]$ and $E\left[Y_o^2(0)\right]$, and $\text{var}\left(Y_o(0)\right)$.

c) $Y_o(\tau)$ is a random variable. Compute $E\left[Y_o(\tau)\right]$ and $E\left[Y_o^2(\tau)\right]$, and $\text{var}\left(Y_o(\tau)\right)$.

d) Find the correlation coefficient between $Y_o(0)$ and $Y_o(\tau)$i.e.,

$$\rho_{Y_o}(\tau) = \frac{E\left[(Y_o(0) - E\left[Y_o(0)\right])(Y_o(\tau) - E\left[Y_o(\tau)\right])\right]}{\sqrt{\text{var}\left(Y_o(0)\right)\text{var}\left(Y_o(\tau)\right)}}. \tag{3.31}$$

e) Plot the joint density function, $f_{Y_o(0)Y_o(\tau)}(y_1, y_2)$, of these two samples for $\tau = 2.5$ms.

f) Plot the joint density function, $f_{Y_o(0)Y_o(\tau)}(y_1, y_2)$, of these two samples for $\tau = 25\mu$s.

g) Compute the time average signal to noise ratio where the instantaneous signal power is defined as $E\left[Y_o(t)\right]^2$.

Problem 3.2. Real valued additive white Gaussian noise (two sided spectral density $N_0/2$) is input into a linear time invariant filter with a real valued impulse response $h_R(t)$ where $h_R(t)$ is constrained to have a unit energy (see Fig. 3.10 where $s_i(t) = 0$).

a) Calculate the output correlation function, $R_N(\tau)$.

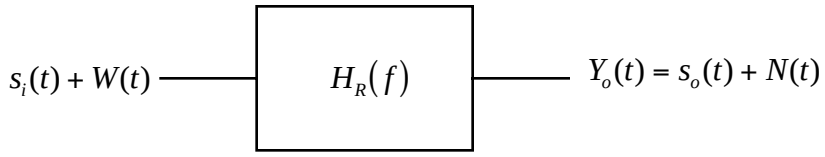


Figure 3.10: Block diagram for signals, noise, and linear systems.

b) The output signal is sampled at a rate of f_s. Give conditions on the impulse response, $h_R(t)$, that will make these samples independent. Give one example of a filter where the output samples will be independent.

c) Give the PDF of one sample, e.g., $f_{N(t_1)}(n_1)$.

d) Give the PDF of any two sample, e.g., $f_{N(t_1)N(t_1+k/f_s)}(n_1, n_2)$ where k is a nonzero integer and the filter satisfies the conditions in b).

e) Give an expression for $P\left(N(t_1) > 2\right)$.

Problem 3.3. Real valued additive white Gaussian noise, $W(t)$ with a two sided spectral density $N_0/2 = 0.1$ is input into a linear time invariant filter with a transfer function of

$$H(f) \;=\; \begin{cases} 2 & |f| \leq 10 \\ 0 & \text{elsewhere} \end{cases} \tag{3.32}$$

The output is denoted $N(t)$.

a) What is the correlation function, $R_W(\tau)$, that is used to model $W(t)$?

b) What is $E\left[N(t)\right]$?

c) Calculate the output power spectral density, $S_N(f)$.

d) What is $E\left[N^2(t)\right]$?

e) Give the PDF of one sample, e.g., $f_{N(t_1)}(n_1)$.

f) Give an expression for $P\left(N(t_1) > 3\right)$.

Problem 3.4. For this problem consider the canonical block diagram shown in Fig. 3.10 with $s_i(t) = A$ and with $W(t)$ being an additive white Gaussian noise with two-sided spectral density of $N_0/2$. Assume a 1Ω system and break the filter response into the DC gain term, $H_R(0)$ and the normalized filter response, $H_N(f)$, i.e., $H_R(f) = H_R(0)H_N(f)$.

a) What is the input signal power?

b) What is the output signal power, P_s?

c) What is the average output noise power, σ_N^2?

d) Show that the output SNR=P_s/σ_N^2 is not a function of $H_R(0)$.

e) Give the output SNR as a function of A, N_0, and B_N.

Problem 3.5. For the canonical problem given in Fig. 3.1, the output noise spectral density is given as

$$S_N(f) \;=\; \begin{cases} N_0 f^2 & |f| \le W \\ 0 & \text{elsewhere} \end{cases} \tag{3.33}$$

a) What is the average ouput noise power?

b) Give a possible $H_R(f)$ that would result in this output noise?

c) Think of a way to implement this $H_R(f)$ using a lowpass filter cascaded with another linear system that is used in the demodulation of an analog modulation.

Problem 3.6. Show that $S_N(f) = S_W(f)\,|H_R(f)|^2$ by taking the Fourier transform of (3.20).

Problem 3.7. Consider the following linear system with a signal, $s_i(t)$, plus a white Gaussian noise, $W(t)$, as inputs.

$$Y(t) = s_o(t) + N(t) = \int_{t-T_i}^{t} x(\lambda)d\lambda = \int_{t-T_i}^{t} \left(s_i(\lambda) + W(\lambda)\right) d\lambda \tag{3.34}$$

where T_i is defined as the integration time.

a) What is the impulse response of this linear system, $h_r(t)$.

b) Show that

$$\sigma_N^2 = \frac{N_0}{2} \int_{-\infty}^{\infty} |h_r(t)|^2 \, dt \tag{3.35}$$

c) If the input to the filter is a DC term plus the AWGN, i.e.,

$$Y(t) = \int_{t-T_i}^{t} (A + W(\lambda))d\lambda \tag{3.36}$$

find the output SNR$= P_s/\sigma_N^2$ as a function of T_i.

d) For the case considered in c) show that SNR grows monotonically with T_i.

e) If the input to the filter is a sinusoid signal plus the AWGN, i.e.,

$$Y(t) = \int_{t-T_i}^{t} (\sqrt{2}A\cos(2\pi f_m t) + W(\lambda))d\lambda \tag{3.37}$$

find the output SNR as a function of T_i.

f) For the case considered in e) show that SNR does not grows monotonically with T_i and find the optimum T_i for each f_m.

Problem 3.8. Real valued additive white Gaussian noise, $W(t)$ with a two sided spectral density $N_0/2 = 0.1$ is input into a linear time invariant filter with a transfer function of

$$H_R(f) \;=\; \begin{cases} 2 & |f| \le W \\ 0 & \text{elsewhere} \end{cases} \tag{3.38}$$

The output is denoted $N(t)$.

a) What is the correlation function, $R_W(\tau)$, that is used to model $W(t)$?

b) What is $E[N(t)]$?

c) Calculate the output power spectral density, $S_N(f)$.

d) Select W to give $E\left[N^2(t)\right]$=10.

e) Give the PDF of one sample, e.g., $f_{N(t_1)}(n_1)$ when $E\left[N^2(t)\right]$=10.

f) Given an expression for $P(N(t_1) > 3)$ when $E\left[N^2(t)\right]$=10.

Problem 3.9. As simple model of a radar receiver consider the following problem. A radar sends out a pulse and listens for the return at the receiver whose filtering is represented with an impulse response $h_R(t)$. The situation when no target is present is represented with the input being modeled as a real valued additive white Gaussian noise, $Y(t) = W(t)$, ($W(t)$ has a two sided spectral density $N_0/2$) being input into a linear time invariant filter with a real valued impulse response $h_R(t)$. Denote the output as $V(t)$. The situation where a target is present can be represented as the same filter with a input $Y(t) = A + W(t)$. A target detection decision is given by a simple threshold test at time t

$$Y(t) \underset{\text{Target Absent}}{\overset{\text{Target Present}}{\gtrless}} \gamma \tag{3.39}$$

a) For the case when no target is present, completely characterize a sample of the process $V(t)$.

b) For a particular time t, find a value of the threshold, γ, where the probability that a target will be declared present when one is not present is 10^{-5}.

c) For the case when a target is present completely characterize a sample of the process $V(t)$.

d) For the value of the threshold selected in b) find how large A should be to correctly detect a target present with a probability of detection greater than 0.9 for any value of t.

Problem 3.10. A real valued additive white Gaussian noise, $W(t)$, (two sided spectral density $N_0/2$) is input into a linear time invariant filter with a real valued impulse response $h_R(t)$. Assume $h_R(t)$ represents an ideal lowpass filter with bandwidth B_T and $N(t)$ is the output of the filter.

a) Find $R_N(\tau)$.

b) Choose B_T such that $E\left[N^2(t)\right] = 1$.

c) For the random process $Z(t) = N(t) - N(t - t_1)$ find $R_Z(\tau)$.

d) Select a value of t_1 such that $E\left[Z^2(t)\right] = 0.5$ for the B_T computed in b).

Problem 3.11. (Adapted from PD) Professor Fitz likes to gamble but he is so busy thinking up problems to torment students that he can never find time to get to Las Vegas. So instead he decides to gamble with the students in his communication classes. The game has real valued additive white Gaussian noise (two sided spectral density $N_0/2 = 1$), $W(t)$, input into a linear time invariant filter with a real valued impulse response

$$h_R(t) = \begin{cases} 2 & 0 \le t \le 0.5 \\ -2 & 0.5 < t \le 1 \\ 0 & \text{elsewhere.} \end{cases} \tag{3.40}$$

Professor Fitz samples the noise, $N(t)$, at a time of his choosing, t_0, and notes the sign of the noise sample. The student playing the game gets to choose a later time, $t_0+\tau, \tau > 0$ to sample the noise and if the sign of the noise sample is the opposite of Professor Fitz's sample then the student gets out of the final exam.

a) Calculate the output correlation function, $R_N(\tau)$.

b) Prof. Fitz samples at $t_0 = 1$ find $E\left[N^2(1)\right]$.

c) Give the probability density function of Prof. Fitz's sample, e.g., $f_{N(t_0)}(n_1)$.

d) If you are playing the game and Professor Fitz samples the noise at $t_0 = 1$, select a sample time to optimize your odds of getting out of the test. Justify your answer. Could you do better if you were allowed to sample the process at a time before Professor Fitz samples the process?

e) If Professor Fitz gets a realization $N(1) = 2$ in his sample, compute the probability of your getting out of the test for the optimized sample time you gave in d).

Problem 3.12. Consider the model in Fig. 3.1 where the noise is zero mean white Gaussian with a one sided noise spectral density of N_0 Watts/Hz and filter has an impulse response of

$$h_R(t) \;=\; \begin{cases} \exp\left[-at\right] & t \geq 0 \\ 0 & \text{elsewhere} \end{cases} \tag{3.41}$$

where $a > 0$.

a) Find $E\left[N^2(t_s)\right]$.

b) Find $P\left(N(t_s) \leq 1\right)$.

c) Find the noise equivalent bandwidth of this filter, B_N.

d) Find $P\left(N(t_s+\tau) \leq 0 \,|N(t_s) = 1\right)$. *Hint: (**??**) might be useful.*

Problem 3.13. Consider the system in Fig. 3.11 where voltage noise source produce a voltage of $N(t)$ where $N(t)$ is a stationary noise with $E\left[N^2(t)\right] = P_N$.

a) Find the average power delivered to the source as a function of R_s and R.

b) What is the value of R that would maximize this power and what would be the corresponding maximum average power delivered.

Problem 3.14. Given Fig. 3.12 and $W(t)$ being an additive white Gaussian noise ($R_W(\tau) = N_0/2\delta(\tau)$).

a) Express the probability density function (PDF) of the random variable $N_1(t_1)$ for a fixed t_1 in terms of $h_1(t)$ and N_0.

b) Find the $E\left[N_1(t_1)N_2(t_1)\right]$ for a fixed t_1 in terms of $h_1(t)$, $h_2(t)$, and N_0. Specify the joint PDF of $N_1(t_1)$ and $N_2(t_1)$.

c) If $h_1(t)$ is given in Fig. 3.13 then find an impulse response for $h_2(t)$ such that $N_1(t_1)$ and $N_2(t_1)$ for a fixed t_1 are independent random variables with equal variances (Hint: the answer is not unique).

Problem 3.15. For this problem consider the canonical block diagram shown in Fig. 3.10 with $s_i(t) = A\cos\left(200\pi t\right)$ and with $W(t)$ being an additive white Gaussian noise with two-sided spectral density of $N_0/2 = 0.01$. $H_R(f)$ is a filter with a transfer function given in Fig. 3.14.

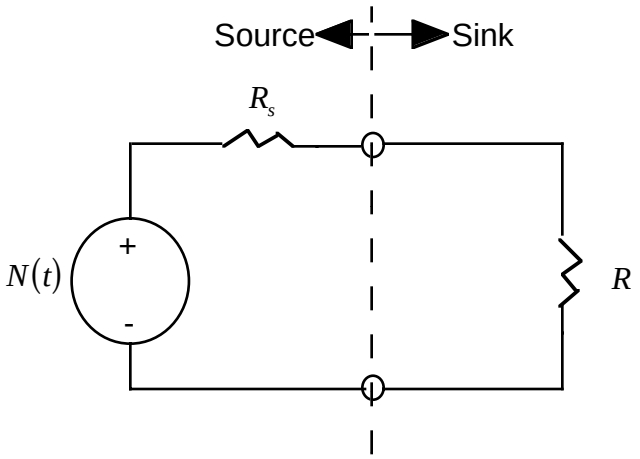


Figure 3.11: A noise source and sink.

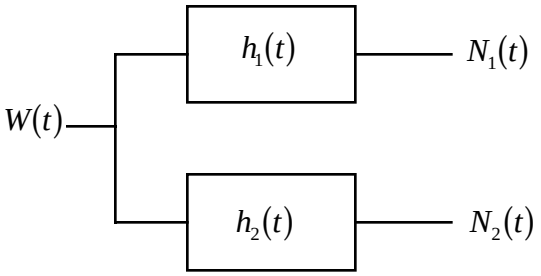


Figure 3.12: White noise into a linear filter.

a) What is $s_o(t)$ and what is P_{s_o}?

b) If two samples are taken from $W(t)$, e.g., $W(t_1)$ and $W(t_2)$, what spacing in the time samples, $t_2 - t_1$, must be maintained to have two samples that are statistically independent random variables?

c) What is the spectral density of $N(t), S_N(f)$?

d) What is the noise power, $E[N^2(t)]$?

e) Choose A such that the output SNR=10.

Problem 3.16. Consider the model in Fig. 3.1 where the noise is zero mean white Gaussian with a one sided noise spectral density of N_0 Watts/Hz and filter is an ideal bandpass filter of bandwidth B_T that has a transfer function of

$$H_R(f) = \begin{cases} 3 & f_c - \frac{B_T}{2} \leq |f| \leq f_c + \frac{B_T}{2} \\ 0 & \text{elsewhere.} \end{cases} \tag{3.42}$$

Assume $f_c \geq B_T$.

a) Find $S_N(f)$

b) Find $E\left[N^2(t_s)\right]$.

Figure 3.13: An example impulse response.

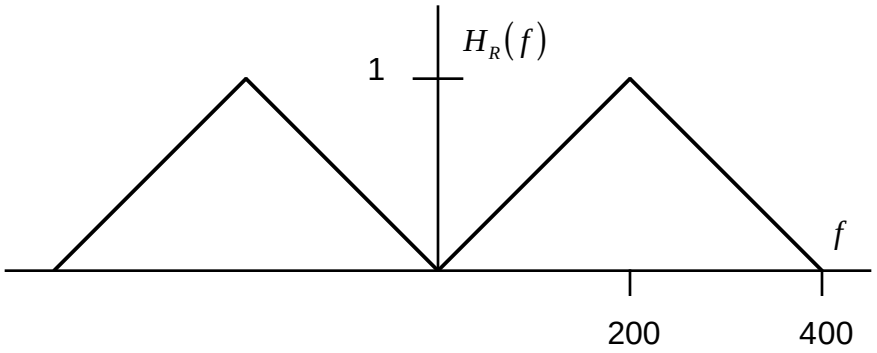


Figure 3.14: Transfer function of $H_R(f)$.

c) Find $R_N(\tau)$

d) Find $P\left(N(t_s) \leq 1\right)$.

e) Find $P\left(N(t_s+\tau) \leq 0 \,|N(t_s)=1\right)$.

f) Find the value of τ that maximizes $P\left(N(t_s+\tau) \leq 0 \,|N(t_s)=1\right)$. *Hint: (**??**) might be useful.*

Problem 3.17. A zero mean, stationary, Gaussian noise, $N(t)$, is characterized by the correlation function given in Fig. 3.15.

a) Give the average power of this noise.

b) Give the joint probability density function of two samples of this noise, $N(1)$ and $N(1.02)$.

c) Find $P\left(N(1) \geq 4\right)$.

d) If this noise is the output of an ideal lowpass filter with $H_R(0)=1$ driven by white noise, estimate the filter bandwidth and two-sided noise spectral density of the noise, $N_0/2$.

Problem 3.18. Two noise processes, $N_1(t)$ and $N_2(t)$, are zero mean, stationary, Gaussian proceses, and have the following correlation functions:

$$R_{N_1}(\tau) = 100\text{sinc}\left(\frac{\tau}{4}\right) \qquad R_{N_2}(\tau) = 50\text{sinc}\left(25\tau\right) \tag{3.43}$$

a) Which random process has a greater power? Why?

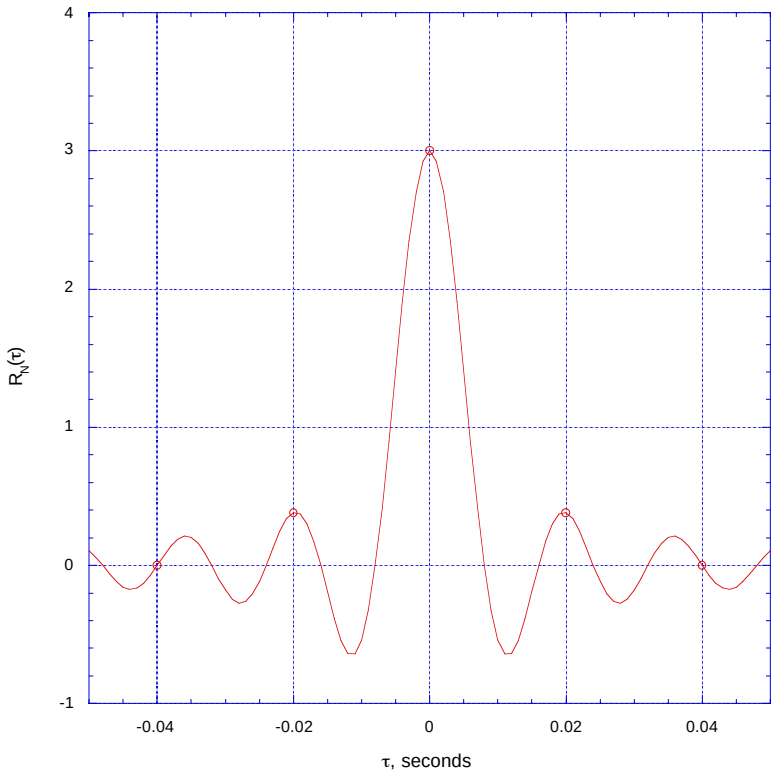


Figure 3.15: A correlation function for a stationary noise.

b) Two samples are taken from $N_2(t)$, $N_2(t_1)$ and $N_2(t_1+\tau)$. What is the smallest value of τ such that the two samples are orthogonal random variables?

c) For $N_1(t)$, what is the correlation coefficient corresponding to $\tau = 1$? Give the joint PDF of two samples of $N_1(t)$ taken 1 second apart.

d) Find $P(N_1(t) < -10)$,

e) Using only linear devices (amplifiers and filters), generate a random process statistically identical to $N_1(t)$ from $N_2(t)$.

Problem 3.19. Two random processes are defined as

$$N_1(t) = X\cos(2\pi f_o t) + Y \tag{3.44}$$

and

$$N_2(t) = X\cos(2\pi f_o t) + Y\sin(2\pi f_o t) \tag{3.45}$$

where X and Y are independent zero mean jointly Gaussian random variables with unity variance and f_o is a deterministic constant.

a) Are these processes Gaussian processes?

b) For each of the processes that are Gaussian given a joint density function of the samples of the random process taken at $t = 0$ and $t = 1/(4f_o)$.

c) Are these processes stationary? *Hint: Trigonometry formulas will be useful here.*

Problem 3.20. Consider the block diagram in Fig. 3.16 where $W_i(t)$ are independent white Gaussian noises with $R_{W_i}(\tau) = N_0/2\delta(\tau), \quad i = 1, 2.$

a) Find $f_{N(t_0)}(n)$.

b) Find $P(N(t_0) > -1)$.

c) Find $f_{N(t_0)N(t_0-\tau)}(n_1, n_2)$.

d) Find $f_{N(t_0)}(n\,|N(t_0-\tau) = n_1)$.

e) Find the simplest expression for $S_E(f)$.

Problem 3.21. Consider an additive white Gaussian noise input into a linear time invariant system as given in Fig. 3.17. Suppose that

$$H_R(f) = \begin{cases} A & |f| \le W \\ 0 & \text{elsewhere} \end{cases} \tag{3.46}$$

and that $S_W(f) = \frac{N_0)}{2}$.

a) Choose A such that $\sigma^2_{N(t)} = 1$.

b) If $W = 100$Hz and $N(t_0) = 2$ find the conditional mean $E[N(t_0-\tau)\,|N(t_0) = 2]$.

c) If $W = 100$Hz, $N(t_0) = 2$, and $\tau > 0.01$ find the value of τ that will minimize

$$\text{var}(N(t_0-\tau)\,|N(t_0) = 2). \tag{3.47}$$

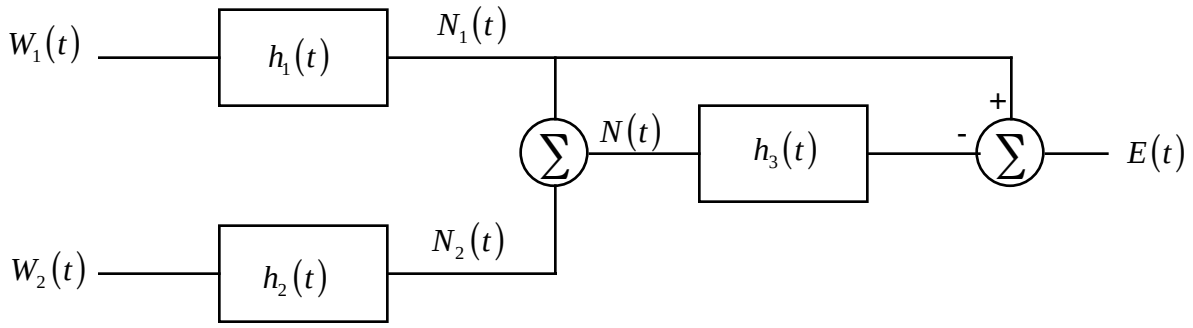


Figure 3.16: A system block diagram.

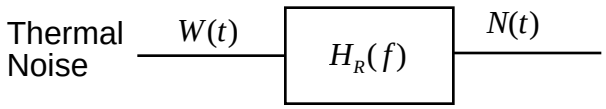


Figure 3.17: The noise model.

3.8 Example Solutions

Problem 3.4.

a) $P_{s_i} = A^2$

b) Since the input signal is a DC signal ($f = 0$) and has a Fourier transform $S_i(f) = A\delta(f)$ the output is $s_o(t) = AH_R(0)$. Consequently $P_s = A^2|H_R(0)|^2$

c)

$$\sigma_N^2 = \frac{N_0}{2}\int_{-\infty}^{\infty}|H_R(f)|^2 df = \frac{N_0}{2}|H_R(0)|^2\int_{-\infty}^{\infty}|H_N(f)|^2 df \tag{3.48}$$

d)

$$SNR = \frac{P_s}{\sigma_N^2} = \frac{2A^2}{N_0}\left[\int_{-\infty}^{\infty}|H_N(f)|^2\, df\right]^{-1} \tag{3.49}$$

e) Using the definition of B_N gives

$$B_N = \frac{1}{2|H_R(0)|^2}\int_{-\infty}^{\infty}|H_R(f)|^2 df = \frac{1}{2|H_R(0)|^2}\int_{-\infty}^{\infty}|H_R(0)|^2|H_N(f)|^2 df. \tag{3.50}$$

Therefore, $B_N = \frac{1}{2}\int_{-\infty}^{\infty}|H_N(f)|^2\, df$ and SNR $= \frac{A^2}{N_0 B_N}$

3.9 Mini-Projects

Goal: To give exposure

1. to a small scope engineering design problem in communications

2. to the dynamics of working with a team

3. to the importance of engineering communication skills (in this case oral presentations).

Presentation: The forum will be similar to a design review at a company (only much shorter) The presentation will be of 5 minutes in length with an overview of the given problem and solution. The presentation will be followed by questions from the audience (your classmates and the professor). Each team member should be prepared to give the presentation.

Project 3.1. You have a voice signal with average power of -10dBm and bandwidth of 4kHz that you want to transmit over a cable from one building on campus to another building on campus. The system block diagram is shown in Fig. 3.18. The noise in the receiver electronics is accurately modeled as an AWGN with a one sided noise spectral density of N_0=-174dBm/Hz. The cable is accurately modeled as a filter with the following impulse response

$$h_c(t) = L_p \delta(t - \tau_p). \tag{3.51}$$

where L_p is the cable loss. You are using cable with a loss of 2dB/1000ft. How long of a cable can be laid and still achieve at least 10dB SNR? If the signal was changed to a video signal with -10dBm average power and bandwidth of 6MHz, how long of a cable can be laid and still achieve at least 10dB SNR?

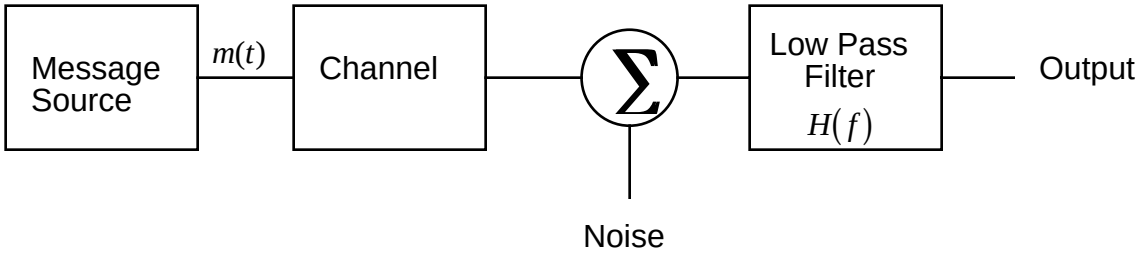


Figure 3.18: Block diagram of a baseband (cable) communication system.

Chapter 4

Noise in Bandpass Communication Systems

Noise in communication systems is produced from a filtering operation on the wideband noise that appears at the receiver input. Most communication occurs in a fairly limited bandwidth around some carrier frequency. Since Chapter 3 showed that the input noise in a communication system is well modeled as an additive white Gaussian noise, it make sense to limit the effects of this noise by filtering around this carrier frequency. Consequently, the canonical model for noise in a bandpass communication system is given in Fig. 4.1. The white noise, $W(t)$, models a wide band noise which is present at the receiver input. This noise can come from a combination of the receiver electronics, man-made noise sources, or galactic noise sources. This noise is often well modeled as a zero mean, stationary, and Gaussian random process with a power spectral density of $N_0/2$ Watts/Hz. The filter, $H_R(f)$, represents the frequency response of the receiver system. This frequency response is often designed to have a bandpass characteristic to match the transmission band and is bandpass around the carrier frequency, f_c. Chapter 3 showed that $N_c(t)$ is a stationary Gaussian process with a power spectral density (PSD) of

$$S_{N_c}(f) = |H_R(f)|^2 \frac{N_0}{2}. \tag{4.1}$$

The noise, $N_c(t)$, will consequently have all it's power concentrated around the carrier frequency, f_c. Noise with this characteristic is denoted bandpass noise. The average power of the noise is given as

$$E\left[N_c^2(t)\right] = \sigma_N^2 = \int_{-\infty}^{\infty} S_{N_c}(f)df = \frac{N_0}{2}\int_{-\infty}^{\infty} |H_R(f)|^2 \, df. \tag{4.2}$$

This noise in a bandpass communication system will then be passed through an I/Q downconverter to produce an in-phase noise, $N_I(t)$, and a quadrature noise, $N_Q(t)$. This chapter provides the necessary tools to characterize the low pass noise processes, $N_I(t)$ and $N_Q(t)$, that result in this situation.

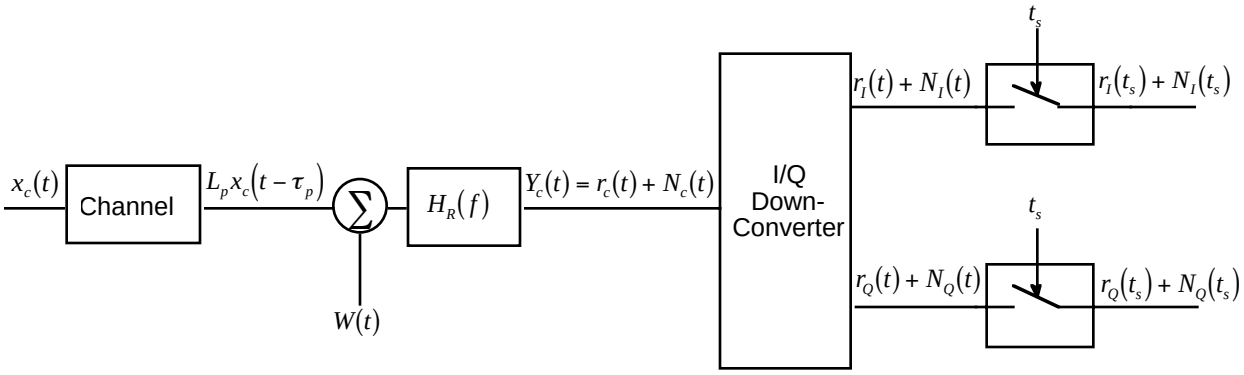


Figure 4.1: The canonical model for communication system analysis.

Example 4.1: Consider a receiver system with an ideal bandpass filter response of bandwidth B_T centered at f_c, i.e.,

$$H_R(f) = \begin{cases} A & ||f| - f_c| \leq \frac{B_T}{2} \\ 0 & \text{elsewhere} \end{cases} \tag{4.3}$$

where A is a real, positive constant. The bandpass noise will have a spectral density of

$$S_{N_c}(f) = \begin{cases} \dfrac{A^2 N_0}{2} & ||f| - f_c| \leq \frac{B_T}{2} \\ 0 & \text{elsewhere} \end{cases} \tag{4.4}$$

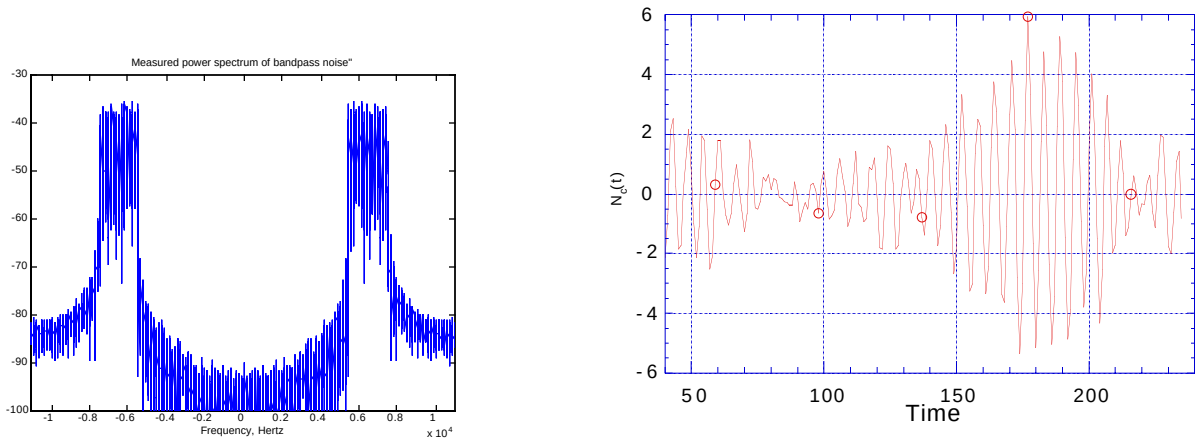
Consequently

$$E\left[N_c^2(t)\right] = \sigma_N^2 = \int_{-\infty}^{\infty} S_{N_c}(f) df = A^2 N_0 B_T \tag{4.5}$$

These tools will enable an analysis of the performance of bandpass communication systems in the presence of noise. This ability to analyze the effects of noise is what distinguishes the competent communication system engineer from the hardware designers and the technicians they work with. The simplest analysis problem is examining a particular point in time, t_s, and characterizing the resulting noise samples, $N_I(t_s)$ and $N_Q(t_s)$, to extract a parameter of interest (e.g., average signal–to–noise ratio (SNR)). Bandpass communication system performance can be characterized completely if probability density functions (PDF) of the noise samples of interest can be identified, e.g.,

$$f_{N_I}(n_1), \qquad f_{N_I(t)N_I(t+\tau)}(n_1, n_2), \quad \text{and/or} \quad f_{N_Q(t)N_I(t+\tau)}(n_1, n_2). \tag{4.6}$$

To accomplish this analysis task this chapter first introduces notation for discussing the bandpass noise and the lowpass noise. The resulting lowpass noise processes are shown to be zero mean, jointly stationary, and jointly Gaussian. Consequently the PDFs like those detailed in (4.6) will entirely be a function of the second order statistics of the lowpass noise processes. These second order statistics can be produced from the PSD given in (4.1).



a) The measured PSD b) The sample function

Figure 4.2: The measured characteristics of bandpass filtered white noise.

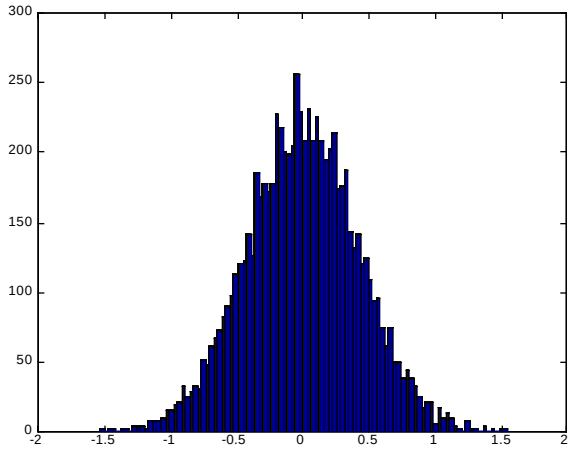


Figure 4.3: A histogram of the samples taken from the bandpass noise resulting from bandpass filtering a white noise.

Example 4.2: If a sampled white Gaussian noise like that considered in Fig. 3.3-b) with a measured spectrum like that given in Fig.3.5-a) is put through a bandpass filter a bandpass noise will result. For the case of a bandpass filter with a center frequency of 6500Hz and a bandwidth of 2000Hz a measured output PSD is shown in Fig. 4.2-a). The validity of (4.1) is clearly evident in this output PSD. A resulting sample function of this output bandpass noise is given in Fig.4.2-b). A histogram of output samples of this noise process is shown in Fig.4.3. This histogram again demonstrates that a bandpass noise is well modeled as a zero mean Gaussian random process.

Point 1: In bandpass communications the input noise, $N_c(t)$, is a stationary Gaussian random process with power spectral density given in (4.1). The effect of noise on the performance of a bandpass communication system can be analyzed if the PDFs of the noise samples of interest can be characterized as in (4.6).

4.1 Bandpass Random Processes

A bandpass random process will have the same form as a bandpass signal. Consequently it can be written as

$$\begin{aligned} N_c(t) &= N_I(t)\sqrt{2}\cos(2\pi f_c t) - N_Q(t)\sqrt{2}\sin(2\pi f_c t) && (4.7)\\ &= N_A(t)\sqrt{2}\cos(2\pi f_c t + N_P(t)). && (4.8)\end{aligned}$$

$N_I(t)$ in (4.7) is normally referred to as the **in-phase (I)** component of the noise and $N_Q(t)$ is normally referred to as the **quadrature (Q)** component of the bandpass noise. The **amplitude** of the bandpass noise is $N_A(t)$ and the **phase** of the bandpass noise is $N_P(t)$. As in the deterministic case the transformation between the two representations are given by

$$N_A(t) = \sqrt{N_I(t)^2 + N_Q(t)^2} \qquad N_P(t) = \tan^{-1}\left[\frac{N_Q(t)}{N_I(t)}\right]$$

and

$$N_I(t) = N_A(t)\cos(N_P(t)) \qquad N_Q(t) = N_A(t)\sin(N_P(t)).$$

A bandpass random process has sample functions which appear to be a sinewave of frequency f_c with a slowly varying amplitude and phase. An example of a bandpass random process is shown in Fig. 4.2-b). As in the deterministic signal case, a method of characterizing a bandpass random process which is independent of the carrier frequency is desired. The complex envelope representation provides such a vehicle.

The **complex envelope** of a bandpass random process is defined as

$$N_z(t) = N_I(t) + jN_Q(t) = N_A(t)\exp\left[jN_P(t)\right]. \tag{4.9}$$

The original bandpass random process can be obtained from the complex envelope by

$$N_c(t) = \sqrt{2}\Re\left[N_z(t)\exp\left[j2\pi f_c t\right]\right].$$

Since the complex exponential is a deterministic function, the complex random process $N_z(t)$ contains all the randomness in $N_c(t)$. In a similar fashion as a bandpass signal, a bandpass random process can be generated from its I and Q components and a complex baseband random process can be generated from the bandpass random process. Fig. 4.4 shows these transformations. Fig. 4.5 shows a bandpass noise and the resulting in-phase and quadrature components that are output from a downconverter structure shown in Fig. 4.4.

Correlation functions are important in characterizing random processes. Autocorrelation functions were a focus of Chapter 3. For bandpass processes the crosscorrelation function will be important as well.

Definition 4.1 *Given two real random processes, $N_I(t)$ and $N_Q(t)$, the crosscorrelation function between these two random processes is given as*

$$R_{N_I N_Q}(t_1, t_2) = E\left[N_I(t_1)N_Q(t_2)\right]. \tag{4.10}$$

The crosscorrelation function is a measure of how similar the two random processes behave. In an analogous manner to the discussion in Chapter 3 a crosscorrelation coefficient can be defined.

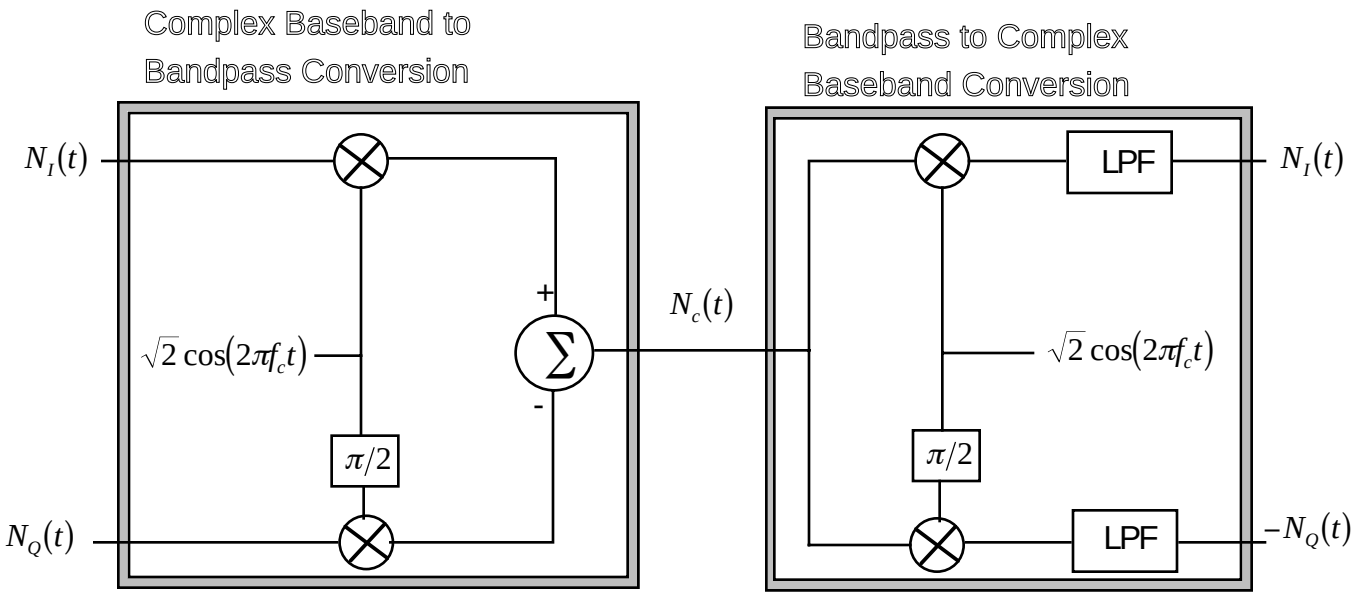


Figure 4.4: The transformations between bandpass noise and the baseband components.

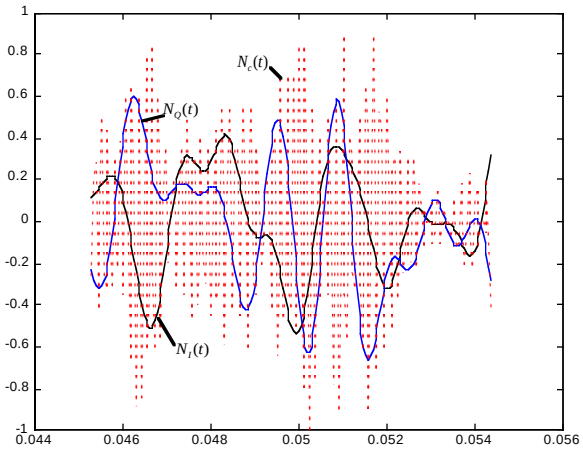


Figure 4.5: A bandpass noise and the resulting in-phase and quadrature noises.

The correlation function of a bandpass random process, which is derived using (4.7), is given by

$$\begin{aligned} R_{N_c}(t_1,t_2) &= E\left[N_c(t_1)N_c(t_2)\right] \\ &= 2R_{N_I}(t_1,t_2)\cos(2\pi f_c t_1)\cos(2\pi f_c t_2) - 2R_{N_I N_Q}(t_1,t_2)\cos(2\pi f_c t_1)\sin(2\pi f_c t_2) \\ &\quad -2R_{N_Q N_I}(t_1,t_2)\sin(2\pi f_c t_1)\cos(2\pi f_c t_2) \\ &\quad +2R_{N_Q}(t_1,t_2)\sin(2\pi f_c t_1)\sin(2\pi f_c t_2). \end{aligned} \tag{4.11}$$

Consequently the correlation function of the bandpass noise is a function of both the correlation function of the two lowpass noise processes and the crosscorrelation between the two lowpass noise processes.

Definition 4.2 *The correlation function of the complex envelope of a bandpass random process is*

$$R_{N_z}(t_1,t_2) = E\left[N_z(t_1)N_z^*(t_2)\right]. \tag{4.12}$$

Using the definition of the complex envelope given in (4.9) produces

$$R_{N_z}(t_1,t_2) = R_{N_I}(t_1,t_2) + R_{N_Q}(t_1,t_2) + j\left[-R_{N_I N_Q}(t_1,t_2) + R_{N_Q N_I}(t_1,t_2)\right]. \tag{4.13}$$

The correlation function of the bandpass signal, $R_{N_c}(t_1,t_2)$, is derived from the complex envelope, via

$$R_{N_c}(t_1,t_2) = 2E(\Re\left[N_z(t_1)\exp\left[j2\pi f_c t_1\right]\right]\Re\left[N_z^*(t_2)\exp\left[-j2\pi f_c t_2\right]\right]). \tag{4.14}$$

This complicated function can be simplified in some practical cases. The case when the bandpass random process, $N_c(t)$, is a stationary random process is one of them.

4.2 Characteristics of the Complex Envelope

4.2.1 Three Important Results

This section shows that the lowpass noise, $N_I(t)$ and $N_Q(t)$ derived from a bandpass noise, $N_c(t)$, are zero mean, jointly Gaussian, and jointly stationary when $N_c(t)$ is zero mean, Gaussian, and stationary. This will be shown to simplify the description of $N_I(t)$ and $N_Q(t)$ considerably. Important quantities in this analysis will be

$$\begin{aligned} N_1(t) &= N_c(t)\sqrt{2}\cos(2\pi f_c t) \\ N_2(t) &= -N_c(t)\sqrt{2}\sin(2\pi f_c t). \end{aligned} \tag{4.15}$$

Property 4.1 *If the bandpass noise, $N_c(t)$, has a zero mean, then $N_I(t)$ and $N_Q(t)$ both have a zero mean.*

Proof: This property's validity can be proved by considering how $N_I(t)$ (or $N_Q(t)$) is generated from $N_c(t)$ as shown in Fig. 4.4. The I component of the noise is expressed as

$$N_I(t) = N_1(t) * h_L(t) = \sqrt{2}\int_{-\infty}^{\infty} h_L(t-\tau)N_c(\tau)\cos(2\pi f_c\tau)d\tau$$

where $h_L(t)$ is the impulse response of the lowpass filter. Neither $h_L(t)$ nor the cosine term are random, so the linearity property of the expectation operator can be used to obtain

$$E(N_I(t)) = \sqrt{2}\int_{-\infty}^{\infty} h_L(t-\tau)E(N_c(\tau))\cos(2\pi f_c\tau)d\tau = 0.$$

The same ideas holds for $N_Q(t)$. □

This property is important since the input thermal noise to a communication system is typically zero mean consequently the I and Q components of the resulting bandpass noise will also be zero mean.

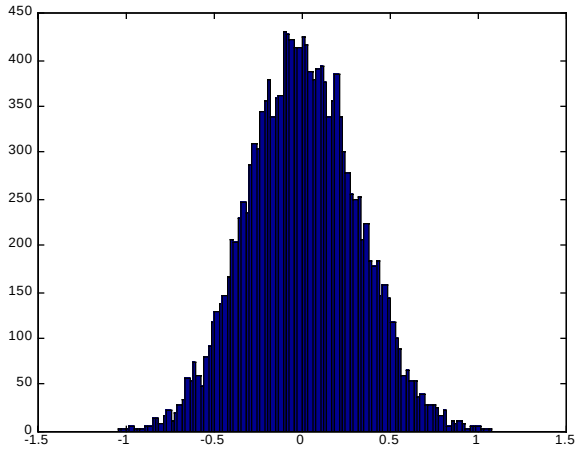


Figure 4.6: A histogram of the samples taken from $N_I(t)$ after downconverting a bandpass noise process.

Definition 4.3 *Two random processes $N_I(t)$ and $N_Q(t)$ are jointly Gaussian random processes if any set of samples taken from the two processes are a set of joint Gaussian random variables.*

Property 4.2 *If the bandpass noise, $N_c(t)$, is a Gaussian random process then $N_I(t)$ and $N_Q(t)$ are jointly Gaussian random processes.*

Proof: The detailed proof techniques are beyond the scope of this course but are contained in most advanced texts concerning random processes (e.g., [DR87]). A sketch of the ideas needed in the proof is given here. A random process which is a product of a deterministic waveform and a Gaussian random process is still a Gaussian random process. Hence $N_1(t)$ and $N_2(t)$ are jointly Gaussian random processes. $N_I(t)$ and $N_Q(t)$ are also jointly Gaussian processes since they are linear functionals of $N_1(t)$ and $N_2(t)$ (i.e., $N_I(t) = N_1(t) * h_L(t)$). □

Again this property implies that the I and Q components of the bandpass noise in most communication systems will be well modeled as Gaussian random processes. $N_c(t)$ can be sampled at $t = k/(2f_c)$ where k is an integer. Since $N_c(t)$ is a Gaussian random process examining these samples, which take the form $N_c(k/(2f_c)) = (-1)^k N_I(k/(2f_c))$, show that many samples of the lowpass processes are Gaussian random variables. Property 4.2 simply implies that all jointly considered samples of $N_I(t)$ and $N_Q(t)$ are jointly Gaussian random variables.

Example 4.3: Consider the previous bandpass filtered noise example where f_c=6500Hz and the bandwidth is 2000Hz. Fig. 4.6 shows a histogram of the samples taken from $N_I(t)$ after downconverting a bandpass noise process. Again is it apparent from this figure that the lowpass noise is well modeled as a zero mean Gaussian random process.

Property 4.3 *If a bandpass signal, $N_c(t)$, is a stationary Gaussian random process, then $N_I(t)$ and $N_Q(t)$ are also jointly stationary, jointly Gaussian random processes.*

Proof: Define the random process $N_{z1}(t) = N_1(t) + jN_2(t) = N_c(t)\sqrt{2}\exp[-j2\pi f_c t]$. Since $N_c(t)$ is a stationary Gaussian random process then $R_{N_{z1}}(t_1, t_2) = 2R_{N_c}(\tau)\exp[-j2\pi f_c \tau] = R_{N_{z1}}(\tau)$ where $\tau = t_1 - t_2$. Since $N_z(t) = N_{z1}(t) * h_L(t)$, if $N_{z1}(t)$ is stationary then the stationarity of the output

complex envelope is due to Property 3.10. Using (4.13) gives

$$R_{N_z}(\tau) = R_{N_I}(t_1,t_2) + R_{N_Q}(t_1,t_2) + j\left[-R_{N_IN_Q}(t_1,t_2) + R_{N_QN_I}(t_1,t_2)\right] \tag{4.16}$$

which implies that both the real and imaginary part of $R_{N_z}(\tau)$ must only be functions of τ so that

$$R_{N_I}(t_1,t_2) + R_{N_Q}(t_1,t_2) = g_1(\tau) \qquad -R_{N_IN_Q}(t_1,t_2) + R_{N_QN_I}(t_1,t_2) = g_2(\tau). \tag{4.17}$$

Alternately, since $N_c(t)$ has the form given in (4.11) a rearrangement (by using trigonometric identities) with $t_2 = t_1 - \tau$ gives

$$\begin{aligned} R_{N_c}(\tau) &= \left[R_{N_I}(t_1,t_2) + R_{N_Q}(t_1,t_2)\right]\cos(2\pi f_c\tau) + \left[R_{N_I}(t_1,t_2) - R_{N_Q}(t_1,t_2)\right]\cos(2\pi f_c(2t_2+\tau)) \\ &\quad + \left[R_{N_IN_Q}(t_1,t_2) - R_{N_QN_I}(t_1,t_2)\right]\sin(2\pi f_c\tau) \\ &\quad - \left[R_{N_IN_Q}(t_1,t_2) + R_{N_QN_I}(t_1,t_2)\right]\sin(2\pi f_c(2t_2+\tau)) \end{aligned} \tag{4.18}$$

Since $N_c(t)$ is Gaussian and stationary this implies that the right hand side of (4.18) is a function only of the time difference, τ and not the absolute time t_2. Consequently the factors multiplying the sinusoidal terms having arguments containing t_2 must be zero. Consequently a second set of constraints is

$$R_{N_I}(t_1,t_2) = R_{N_Q}(t_1,t_2) \qquad R_{N_IN_Q}(t_1,t_2) = -R_{N_QN_I}(t_1,t_2). \tag{4.19}$$

The only way for (4.17) and (4.19) to be satisfied is if

$$R_{N_I}(t_1,t_2) = R_{N_Q}(t_1,t_2) = R_{N_I}(\tau) = R_{N_Q}(\tau) \tag{4.20}$$

$$R_{N_IN_Q}(t_1,t_2) = -R_{N_QN_I}(t_1,t_2) = R_{N_IN_Q}(\tau) = -R_{N_QN_I}(\tau) \tag{4.21}$$

Since all correlation functions and crosscorrelation functions are functions of τ then $N_I(t)$ and $N_Q(t)$ are jointly stationary. □

Point 2: If the input noise, $N_c(t)$, is a zero mean, stationary, Gaussian random process then $N_I(t)$ and $N_Q(t)$ are zero mean, jointly Gaussian, and jointly stationary.

4.2.2 Important Corollaries

This section discusses several important corollaries to the important results derived in the last section. The important result from the last section is summarized as follows: if $N_c(t)$ is zero mean, Gaussian and stationary then $N_I(t)$ and $N_Q(t)$ are zero mean, jointly Gaussian, and jointly stationary.

Property 4.4

$$R_{N_I}(\tau) = R_{N_Q}(\tau) \tag{4.22}$$

Surprisingly, this property, given in (4.20), implies that both $N_I(t)$ and $N_Q(t)$ behave in a statistically identical manner. Consequently the power of the noise in each component is identical, i.e., $E\left[N_I^2(t)\right] = E\left[N_Q^2(t)\right] = \sigma_{N_I}^2$

Property 4.5

$$R_{N_IN_Q}(\tau) = -R_{N_IN_Q}(-\tau) \tag{4.23}$$

This property, due to (4.21), implies $R_{N_IN_Q}(\tau)$ is an odd function with respect to τ.

Property 4.6 $R_{N_z}(\tau) = 2R_{N_I}(\tau) - j2R_{N_IN_Q}(\tau)$

Proof: This is shown by using (4.22) and (4.23) in (4.13). □

This implies that the real part of $R_{N_z}(\tau)$ is an even function of τ and the imaginary part of $R_{N_z}(\tau)$ is an odd function of τ.

Property 4.7

$$R_{N_c}(\tau) = 2R_{N_I}(\tau)\cos(2\pi f_c\tau) + 2R_{N_IN_Q}(\tau)\sin(2\pi f_c\tau) = \Re\left[R_{N_z}(\tau)\exp(j2\pi f_c\tau)\right] \tag{4.24}$$

Proof: This is shown by using (4.22) and (4.23) in (4.18). □

This property implies there is a simple relationship between the correlation function of the stationary Gaussian bandpass noise and the correlation function of the complex envelope of the bandpass noise. This relationship has significant parallels to the relationship between bandpass and the baseband signals given in Chapter 2.

Property 4.8 $var(N_c(t)) = \sigma_N^2 = var(N_I(t)) + var(N_Q(t)) = 2var(N_I(t)) = var(N_z(t)) = 2\sigma_{N_I}^2$

Proof: This is shown by using (4.22) in (4.24) for $\tau = 0$. □

This property states that the power in the bandpass noise is the same as the power in the complex envelope. This power in the bandpass noise is also the sum of the powers in the two lowpass noises which comprise the complex envelope.

Property 4.9 *For the canonical problem considered in this chapter, $N_I(t)$ and $N_Q(t)$ are completely characterized by the functions $R_{N_I}(\tau)$ and $R_{N_IN_Q}(\tau)$.*

Proof: Any joint PDF of samples taken from jointly stationary and jointly Gaussian processes are completely characterized by the first and second order moments. Note first that $N_I(t)$ and $N_Q(t)$ are zero mean processes. Consequently the characterization of the process only requires the identification of the variance, $\sigma_N^2 = R_{N_I}(0)$, and the correlation coefficient between samples. The correlation coefficient between samples from the same process is given as $\rho_{N_I}(\tau) = R_{N_I}(\tau)/R_{N_I}(0)$ while the correlation coefficient between samples taken from $N_I(t)$ and $N_Q(t)$ is $\rho_{N_IN_Q}(\tau) = R_{N_IN_Q}(\tau)/R_{N_I}(0)$. □

Example 4.4: If one sample from the in-phase noise is considered then

$$f_{N_I}(n_i) = \frac{1}{\sqrt{2\pi\sigma_{N_I}^2}}\exp\left(-\frac{n_i^2}{2\sigma_{N_I}^2}\right) \tag{4.25}$$

where $\sigma_{N_I}^2 = R_{N_I}(0)$.

Example 4.5: If two samples from the in-phase noise are considered then

$$f_{N_I(t)N_I(t-\tau)}(n_1,n_2) = \frac{1}{2\pi\sigma_{N_I}^2\sqrt{(1-\rho_{N_I}^2(\tau))}}\exp\left[\frac{-\left(n_1^2 - 2\rho_{N_I}(\tau)n_1n_2 + n_2^2\right)}{2\sigma_{N_I}^2(1-\rho_{N_I}^2(\tau))}\right]$$

where $\rho_{N_I}(\tau) = R_{N_I}(\tau)/R_{N_I}(0)$.

Example 4.6: If one sample each from the in-phase noise and the quadrature noise are considered then

$$f_{N_I(t)N_Q(t-\tau)}(n_1,n_2)=\frac{1}{2\pi\sigma_{N_I}^2\sqrt{(1-\rho_{N_IN_Q}^2(\tau))}}\exp\left[\frac{-\left(n_1^2-2\rho_{N_IN_Q}(\tau)n_1n_2+n_2^2\right)}{2\sigma_{N_I}^2(1-\rho_{N_IN_Q}^2(\tau))}\right]$$

where $\rho_{N_IN_Q}(\tau)=R_{N_IN_Q}(\tau)/\sqrt{R_{N_I}(0)R_{N_Q}(0)}=R_{N_IN_Q}(\tau)/R_{N_I}(0)$.

Property 4.10 *The random variables $N_I(t_s)$ and $N_Q(t_s)$ are independent random variables for any value of t_s.*

Proof: From (4.23) we know $R_{N_IN_Q}(\tau)$ is an odd function. Consequently $R_{N_IN_Q}(0)=0$. Since $N_I(t)$ and $N_Q(t)$ are jointly Gaussian and orthogonal random variables then they are also independent. □

Any joint PDF of samples of $N_I(t_s)$ and $N_Q(t_s)$ taken at the same time have the simple PDF

$$f_{N_I(t)N_Q(t)}(n_1,n_2)=\frac{1}{2\pi\sigma_{N_I}^2}\exp\left[\frac{-\left(n_1^2+n_2^2\right)}{2\sigma_{N_I}^2}\right]. \tag{4.26}$$

This simple PDF will prove useful in our performance analysis of bandpass communication systems.

Point 3: Since $N_I(t)$ and $N_Q(t)$ are zero mean, jointly Gaussian, and jointly stationary then a complete statistical description of $N_I(t)$ and $N_Q(t)$ is available from $R_{N_I}(\tau)$ and $R_{N_IN_Q}(\tau)$.

4.3 Spectral Characteristics

At this point we need a methodology to translate the known PSD of the bandpass noise (4.1) into the required correlation function, $R_{N_I}(\tau)$, and the required crosscorrelation function, $R_{N_IN_Q}(\tau)$. To accomplish this goal we need a definition.

Definition 4.4 *For two random processes $N_I(t)$ and $N_Q(t)$ whose cross correlation function is given as $R_{N_IN_Q}(\tau)$ the cross spectral density is*

$$S_{N_IN_Q}(f)=\mathcal{F}\left\{R_{N_IN_Q}(\tau)\right\}. \tag{4.27}$$

Property 4.11 *The PSD of $N_z(t)$ is given by*

$$S_{N_z}(f)=\mathcal{F}\left\{R_{N_z}(\tau)\right\}=2S_{N_I}(f)-j2S_{N_IN_Q}(f) \tag{4.28}$$

where $S_{N_I}(f)$ and $S_{N_IN_Q}(f)$ are the power spectrum of $N_I(t)$ and the cross power spectrum of $N_I(t)$ and $N_Q(t)$, respectively.

Proof: This is seen taking the Fourier transform of $R_{N_z}(\tau)$ as given in Property 4.6. □

Property 4.12 *$S_{N_IN_Q}(f)$ is purely an imaginary function and an odd function of f.*

Proof: This is true since $R_{N_IN_Q}(\tau)$ is an odd real valued function (see Property **??**). A power spectral density must be real and positive so by examining (4.28) it is clear $S_{N_IN_Q}(f)$ must be imaginary for $S_{N_z}(f)$ to be a valid power spectral density. □

Property 4.13 *The even part of $S_{N_z}(f)$ is due to $S_{N_I}(f)$ and the odd part is due to $S_{N_IN_Q}(f)$.*

Proof: A spectral density of a real random process is always even. $S_{N_I}(f)$ is a spectral density of a real random process. By Property 4.12, $S_{N_I N_Q}(f)$ is a purely imaginary and odd function of frequency. □

Property 4.14

$$S_{N_I}(f) = \frac{S_{N_z}(f) + S_{N_z}(-f)}{4} \qquad S_{N_I N_Q}(f) = \frac{S_{N_z}(-f) - S_{N_z}(f)}{j4}.$$

Proof: This is a trivial result of Property 4.13 □

Consequently Property 4.14 provides a simple method to compute $S_{N_I}(f)$ and $S_{N_I N_Q}(f)$ once $S_{N_z}(f)$ is known. $S_{N_z}(f)$ can be computed from $S_{N_c}(f)$ given in (4.1) in a simple way as well.

Property 4.15

$$S_{N_c}(f) = \frac{1}{2} S_{N_z}(f - f_c) + \frac{1}{2} S_{N_z}(-f - f_c).$$

Proof: Examining (4.24) and the Frequency Translation theorem of the Fourier transform, the spectral density of the bandpass noise, $N_z(t)$, is expressed as

$$S_{N_c}(f) = 2S_{N_I}(f) * \left[\frac{1}{2}\delta(f - f_c) + \frac{1}{2}\delta(f + f_c)\right] + 2S_{N_I N_Q}(f) * \left[\frac{1}{2j}\delta(f - f_c) - \frac{1}{2j}\delta(f + f_c)\right]$$

where $*$ again denotes convolution. This equation can be rearranged to give

$$S_{N_c}(f) = \left[S_{N_I}(f - f_c) - jS_{N_I N_Q}(f - f_c)\right] + \left[S_{N_I}(f + f_c) + jS_{N_I N_Q}(f + f_c)\right]. \tag{4.29}$$

Noting that due to Property 4.12

$$S_{N_z}(-f) = 2S_{N_I}(f) + j2S_{N_I N_Q}(f), \tag{4.30}$$

(4.29) reduces to the result in Property 4.15. □

This is a very fundamental result. Property 4.15 states that the power spectrum of a bandpass random process is simply derived from the power spectrum of the complex envelope and vice versa. For positive values of f, $S_{N_c}(f)$ is obtained by translating $S_{N_z}(f)$ to f_c and scaling the amplitude by 0.5 and for negative values of f, $S_{N_c}(f)$ is obtained by flipping $S_{N_z}(f)$ around the origin, translating the result to $-f_c$, and scaling the amplitude by 0.5. Likewise $S_{N_z}(f)$ is obtained from $S_{N_c}(f)$ by taking the positive frequency PSD which is centered at f_c and translating it to baseband ($f = 0$) and multiplying it by 2. Property 4.15 also demonstrates in another manner that the average power of the bandpass and baseband noises are identical since the area under the PSD is the same (this was previously shown in Property 4.8).

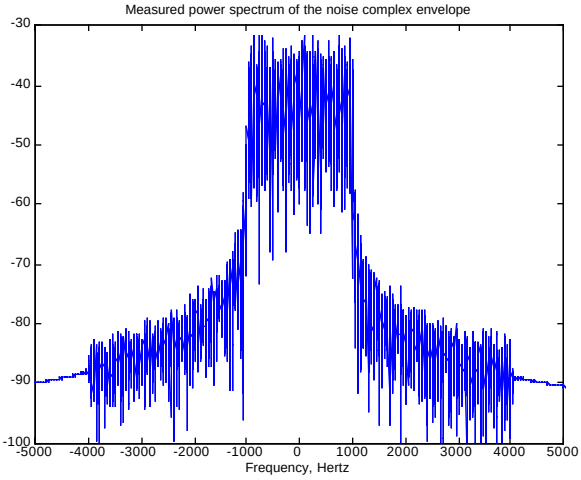


Figure 4.7: The measured PSD of the complex envelope of a bandpass process resulting from bandpass filtering a white noise.

Example 4.7: Example 4.1 showed a receiver system with an ideal bandpass filter that had a bandpass noise PSD of

$$S_{N_c}(f) = \begin{cases} \frac{A^2 N_0}{2} & ||f| - f_c| \leq \frac{B_T}{2} \\ 0 & \text{elsewhere} \end{cases} \tag{4.31}$$

For this case

$$S_{N_z}(f) = \begin{cases} A^2 N_0 & |f| \leq \frac{B_T}{2} \\ 0 & \text{elsewhere} \end{cases} \tag{4.32}$$

and

$$S_{N_I}(f) = \begin{cases} \frac{A^2 N_0}{2} & |f| \leq \frac{B_T}{2} \\ 0 & \text{elsewhere} \end{cases} \qquad S_{N_I N_Q}(f) = 0 \tag{4.33}$$

Again considering the previous example of bandpass filtered white noise with f_c=6500Hz and a bandwidth of 2000Hz a resulting measured power spectral density of the complex envelope is given in Fig. 4.7. This measured PSD demonstrates the validity of the analytical results given in (4.33).

Point 4: $S_{N_I}(f)$ and $S_{N_I N_Q}(f)$ can be computed in a straightforward fashion from $S_{N_c}(f)$ given in (4.1). An inverse Fourier transform will produce the functions $R_{N_I}(\tau)$ and $R_{N_I N_Q}(\tau)$. From $R_{N_I}(\tau)$ and $R_{N_I N_Q}(\tau)$ a complete statistical description of the complex envelope noise process can be obtained.

4.4 The Solution of the Canonical Bandpass Problem

The tools and results are now in place to completely characterize the complex envelope of the bandpass noise typically encountered in a bandpass communication system. First, the characterization of one sample, $N_I(t_s)$, of the random process $N_I(t)$ is considered (or equivalently $N_Q(t_s)$). This case requires a six step process summarized as

1. Identify N_0 and $H_R(f)$.

2. Compute $S_{N_c}(f) = \frac{N_0}{2}\left|H_R(f)\right|^2$.
3. Compute $S_{N_z}(f)$.
4. Compute
$$S_{N_I}(f) = \frac{S_{N_z}(f) + S_{N_z}(-f)}{4}.$$
5. $\sigma^2_{N_I} = R_{N_I}(0) = \int_{-\infty}^{\infty} S_{N_I}(f)df$.
6. $f_{N_I(t_s)}(n_1) = \frac{1}{\sqrt{2\pi\sigma^2_{N_I}}}\exp\left(-\frac{n_1^2}{2\sigma^2_{N_I}}\right) = f_{N_Q(t_s)}(n_1)$.

The only difference between this process and the process used for lowpass processes as highlighted in Section 3.6 is step 3–4. These two steps simply are the transformation of the bandpass PSD into the PSD for one channel of the lowpass complex envelope noise.

Example 4.8: The previous examples showed a receiver system with an ideal bandpass filter had after the completion of steps 1–4
$$S_{N_I}(f) = \begin{cases} \frac{A^2N_0}{2} & |f| \le \frac{B_T}{2} \\ 0 & \text{elsewhere.} \end{cases} \tag{4.34}$$
Consequently $\sigma^2_{N_I} = R_{N_I}(0) = \frac{A^2N_0B_T}{2}$.

Second, the characterization of two samples from one of the channels of the complex envelope, $N_I(t_1)$ and $N_I(t_2)$, is considered. This case requires a seven step process summarized as

1. Identify N_0 and $H_R(f)$.
2. Compute $S_{N_c}(f) = \frac{N_0}{2}\left|H_R(f)\right|^2$.
3. Compute $S_{N_z}(f)$.
4. Compute
$$S_{N_I}(f) = \frac{S_{N_z}(f) + S_{N_z}(-f)}{4}.$$
5. $R_{N_I}(\tau) = \mathcal{F}^{-1}\left\{S_{N_I}(f)\right\}$.
6. $\sigma^2_{N_I} = R_{N_I}(0)$ and $\rho_{N_I}(\tau) = R_{N_I}(\tau)/\sigma^2_{N_I}$.
7. $f_{N_I(t_1)N_I(t_2)}(n_1,n_2) = \frac{1}{2\pi\sigma^2_{N_I}\sqrt{(1-\rho^2_{N_I}(\tau))}}\exp\left[\frac{-1}{2\sigma^2_N(1-\rho^2_{N_I}(\tau))}\left(n_1^2 - 2\rho_{N_I}(\tau)n_1n_2 + n_2^2\right)\right]$.

The only difference between this process and the process used for lowpass processes as highlighted in Section 3.6 is step 3–4. These two steps again are the transformation of the bandpass PSD into the PSD for one channel of the lowpass complex envelope noise.

Example 4.9: The previous examples showed a receiver system with an ideal bandpass filter had after the completion of steps 1–4
$$S_{N_I}(f) = \begin{cases} \frac{A^2N_0}{2} & |f| \le \frac{B_T}{2} \\ 0 & \text{elsewhere.} \end{cases} \tag{4.35}$$
Consequently Step 5 gives $\sigma^2_{N_I} = R_{N_I}(0) = \frac{A^2N_0B_T}{2}$ and $R_{N_I}(\tau) = \frac{A^2N_0B_T}{2}\text{sinc}\left(B_T\tau\right)$. This implies that $\rho_{N_I}(\tau) = \text{sinc}\left(B_T\tau\right)$.

Finally, the characterization of two samples one from each of the channels of the complex envelope, $N_I(t_1)$ and $N_Q(t_2)$, is considered. This case requires a seven step process summarized as

1. Identify N_0 and $H_R(f)$.
2. Compute $S_{N_c}(f) = \frac{N_0}{2}\left|H_R(f)\right|^2$.
3. Compute $S_{N_z}(f)$.
4. Compute
$$S_{N_I}(f) = \frac{S_{N_z}(f) + S_{N_z}(-f)}{4} \qquad S_{N_I N_Q}(f) = \frac{S_{N_z}(-f) - S_{N_z}(f)}{j4}.$$
5. $R_{N_I N_Q}(\tau) = \mathcal{F}^{-1}\left\{S_{N_I N_Q}(f)\right\}$.
6. $\sigma^2_{N_I} = R_{N_I}(0)$ and $\rho_{N_I N_Q}(\tau) = R_{N_I N_Q}(\tau)/\sigma^2_{N_I}$.
7. $f_{N_I(t_1)N_Q(t_2)}(n_1, n_2) = \frac{1}{2\pi\sigma^2_{N_I}\sqrt{(1-\rho^2_{N_I N_Q}(\tau))}} \exp\left[\frac{-1}{2\sigma^2_{N_I}(1-\rho^2_{N_I N_Q}(\tau))}\left(n_1^2 - 2\rho_{N_I N_Q}(\tau)n_1 n_2 + n_2^2\right)\right]$.

The only difference between this process and the process for two samples from the same channel is the computation of the cross correlation function, $R_{N_I N_Q}(\tau)$.

Example 4.10: The previous examples showed a receiver system with an ideal bandpass filter had after the completion of steps 1–4

$$S_{N_I}(f) = \begin{cases} \frac{A^2 N_0}{2} & |f| \le \frac{B_T}{2} \\ 0 & \text{elsewhere} \end{cases} \qquad S_{N_I N_Q}(f) = 0. \tag{4.36}$$

Consequently Step 5 gives $\sigma^2_{N_I} = R_{N_I}(0) = \frac{A^2 N_0 B_T}{2}$ and $R_{N_I N_Q}(\tau) = 0$. This implies that $N_I(t_1)$ and $N_Q(t_2)$ are independent random variables irregardless of t_1 and t_2.

Similarly, three or more samples from either channel of the complex envelope could be characterized in a very similar fashion. The tools developed in this chapter give a student the ability to analyze many of the important properties of noise that are of interest in a bandpass communication system design.

4.5 Homework Problems

Problem 4.1. What conditions on the bandpass filter characteristic of the receiver, $H_R(f)$, must be satisfied such that $N_I(t_1)$ and $N_Q(t_2)$ are independent random variables for all t_1 and t_2 when the input is AWGN?

Problem 4.2. This problem considers noise that might be seen in a vestigial sideband demodulator. Consider a bandpass Gaussian noise at the input to a demodulator with a spectrum given as

$$S_{N_c}(f) = \begin{cases} 2 & f_c - 1000 \le |f| \le f_c + 3000 \\ 0 & \text{elsewhere.} \end{cases} \tag{4.37}$$

Assume operation is in a 1Ω system.

a) What bandpass filter, $H_R(f)$, would produce this spectrum from a white noise input with a two sided noise spectral density of $N_0/2$=0.5?

b) What is the spectral density of $N_I(t)$?

c) What is $E\left[N_I^2(t)\right]$?

d) Give the joint PDF of $N_I(t_0)$ and $N_Q(t_0)$ in a 1Ω system for a fixed t_0.

e) Compute $S_{N_I N_Q}(f)$.

Problem 4.3. This problem considers noise which might be seen in a single sideband demodulator. Consider a bandpass Gaussian noise at the input to a demodulator with a spectrum given as

$$S_{N_c}(f) = \begin{cases} 3.92 \times 10^{-3} & f_c \leq |f| \leq f_c + 3000 \\ 0 & \text{elsewhere.} \end{cases} \tag{4.38}$$

Assume operation is in a 1Ω system.

a) What is the spectral density of $N_I(t)$?

b) What is $E\left[N_I^2(t)\right]$?

c) Give the joint PDF of $N_I(t_0)$ and $N_Q(t_0)$.

d) Give the joint PDF of $N_I(t_1)$ and $N_Q(t_1 - \tau)$.

e) Plot the PDF in d) for $\tau = 0.0001$ and $\tau = 0.001$.

Problem 4.4. Consider a bandpass Gaussian noise at the input to a demodulator with a spectrum given as

$$S_{N_c}(f) = \begin{cases} 0.001 & f_c - 2000 \leq |f| \leq f_c + 2000 \\ 0 & \text{elsewhere.} \end{cases} \tag{4.39}$$

Assume operation is in a 1Ω system.

a) Find the joint density function of $N_I(t_0)$ and $N_Q(t_0)$, $f_{N_I(t_0)N_Q(t_0)}(n_i, n_q)$.

b) Plot $f_{N_I(t_0)N_Q(t_0)}(n_i, n_q)$.

c) Consider the complex random variable $\tilde{N}_z(t) = \tilde{N}_I(t) + j\tilde{N}_Q(t) = N_z(t)\exp(-j\phi_p)$ and find an inverse mapping such that

$$\tilde{N}_I(t) = g_1\left(N_I(t_0), N_Q(t_0)\right) \qquad \tilde{N}_Q(t) = g_2\left(N_I(t_0), N_Q(t_0)\right). \tag{4.40}$$

d) Using the results of Section **??** show that $f_{\tilde{N}_I(t)\tilde{N}_Q(t)}(n_i, n_q) = f_{N_I(t)N_Q(t)}(n_i, n_q)$. In other words the noise distribution is unchanged by a phase rotation. This result is very important for the study of coherent receiver performance in the presence of noise detailed in the sequel.

Problem 4.5. Show that $|S_{N_I N_Q}(f)| \leq S_{N_I}(f)$. If $|S_{N_I N_Q}(f_0)| = S_{N_I}(f_0)$ for some f_0 then this places a constraint on either $H_R(f_c + f_0)$ or $H_R(f_c - f_0)$. What is this constraint?

Problem 4.6. For a bandpass noise with complex envelope $N_z(t) = N_I(t) + jN_Q(t)$ that has a baseband autocorrelation function of $R_{N_I}(\tau)$ and a baseband crosscorrelation function of $R_{N_I N_Q}(\tau)$

a) Find $E\left[|N_z(t)|^2\right]$.

b) Find $E\left[N_z^2(t)\right]$.

c) Find $E\left[N_z(t)N_z^*(t-\tau)\right]$.

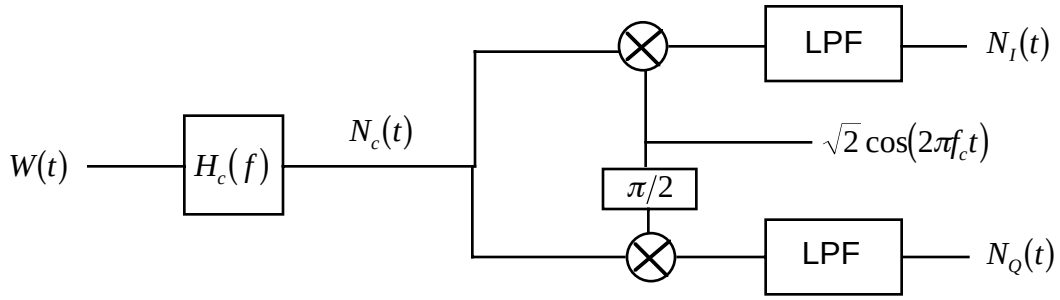


Figure 4.8: Noise and the receiver frontend processor.

d) Find $E\left[N_z(t)N_z(t-\tau)\right]$.

Problem 4.7. Given the block diagram in Fig. 4.8 where $W(t)$ is an AWGN with two-sided spectral density $N_0/2$ and $H_R(f)$ is a bandpass filter with a passband including f_c, find the conditions on the filter $H_R(f)$ such that

a) $N_I(t)$ and $N_Q(t)$ have the same correlation function.

b) $R_{N_I}(\tau) = R_{N_Q}(\tau) = \frac{\sin(\pi B\tau)}{\pi B\tau}$ and $R_{N_I N_Q}(\tau) = 0$.

Problem 4.8. Real valued additive white Gaussian noise, $W(t)$ with a one sided spectral density $N_0 = 0.01$ is input into a bandpass receiver which has a block diagram shown in Fig. 4.8. The noise power spectrum output from the down converter is measured as

$$S_{N_z}(f) \;=\; \begin{cases} 0.25 & |f| \le 100 \\ 0 & \text{elsewhere.} \end{cases} \tag{4.41}$$

a) Find $E\left[|N_z(t)|^2\right]$.

b) Find $S_{N_I}(f)$ and $S_{N_I N_Q}(f)$.

c) What is $|H_R(f)|$?

d) Give the probability density function of one sample of $N_I(t)$, i.e., $f_{N_I(t)}(n)$.

e) Compute $P\left(N_I(t) > 4\right)$.

Problem 4.9. Real valued additive white Gaussian noise, $W(t)$ with a two sided spectral density $N_0/2$ is input into a demodulator for signal sideband amplitude modulation which has a block diagram shown in Fig. 4.10. The time invariant filter has a transfer function of

$$H_R(f) \;=\; \begin{cases} 2 & f_c \le |f| \le f_c + W \\ 0 & \text{elsewhere} \end{cases} \tag{4.42}$$

a) Find the power spectral density of $N_c(t), S_{N_c}(f)$?

b) Find $S_{N_I}(f)$?

c) Calculate $R_{N_I}(\tau)$.

d) Detail out the form for the probability density function (PDF) of one sample from $N_I(t), f_{N_I(t_s)}(n)$.

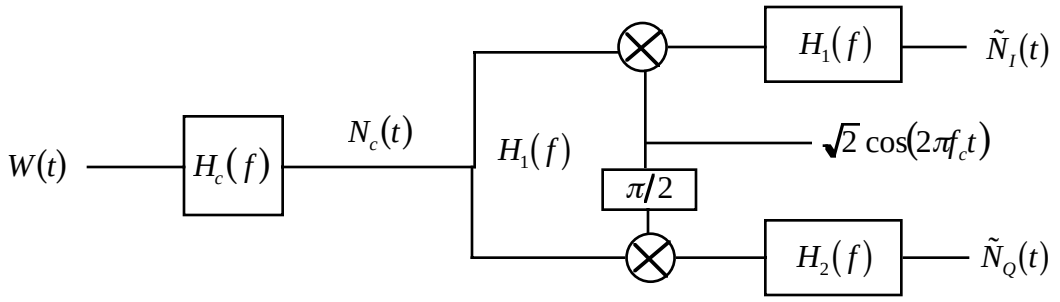


Figure 4.9: A bandpass noise model.

e) Find $S_{N_I N_Q}(f)$.

f) Find the joint PDF of samples taken from $N_I(t)$ and $N_Q(t)$ taken at the same time, $f_{N_I(t_s)N_Q(t_s)}(n_i, n_q)$.

Problem 4.10. Consider the model given in Fig. 4.9. $H_1(f)$ and $H_2(f)$ are two potentially different filters. This problem investigates how the characteristics of the bandpass noise change if the filter in the I channel is different from the filter in the Q channel.

a) Find the power spectral density of $\tilde{N}_I(t), S_{\tilde{N}_I}(f)$, in terms of $S_{N_I}(f)$.

b) Find the cross spectrum $S_{\tilde{N}_I \tilde{N}_Q}(f)$ in terms of $S_{N_I N_Q}(f)$.

c) Will $\tilde{N}_I(t)$ always be independent of $\tilde{N}_Q(t)$ as when the filters are the same? If not then give an example?

d) Give conditions on $H_1(f)$ and $H_2(f)$ such that $S_{\tilde{N}_I \tilde{N}_Q}(f)$ would be imaginary and odd as is the case for $S_{N_I N_Q}(f)$. If these conditions were satisfied would $\tilde{N}_I(t)$ always be independent of $\tilde{N}_Q(t)$?

Problem 4.11. Show that if $\tilde{N}_z(t) = N_z(t) \exp[j\phi_p]$ then

a) $R_{\tilde{N}_I}(\tau) = R_{\tilde{N}_Q}(\tau) = R_{N_I}(\tau)$.

b) $R_{\tilde{N}_I \tilde{N}_Q}(\tau) = R_{N_I N_Q}(\tau)$

Problem 4.12. Real valued additive white Gaussian noise, $W(t)$ with a two sided spectral density $N_0/2$ is input into a demodulator which has a block diagram shown in Fig. 4.10. The time invariant filter has a transfer function of

$$H_R(f) = \begin{cases} 2 & f_c - 100 \leq |f| \leq f_c + 100 \\ 0 & \text{elsewhere} \end{cases} \tag{4.43}$$

a) Find the power spectral density of $N_c(t), S_{N_c}(f)$.

b) $E\left[N_c^2(t)\right] = 2$, what is N_0?

c) Find the power spectral density of $N_I(t), S_{N_I}(f)$.

d) Find $E\left[N_I^2(t)\right]$ for the value of N_0 obtained in b).

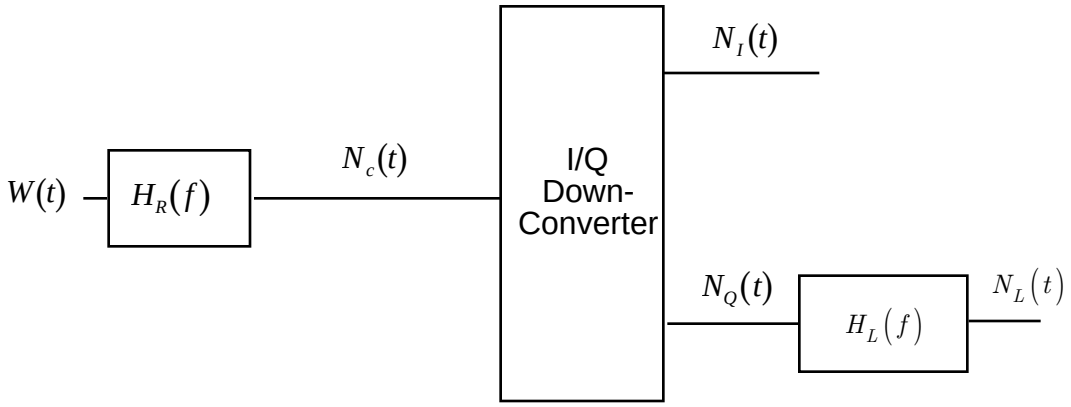


Figure 4.10: Demodulator block diagram.

e) Assume that $H_L(f)$ is an ideal lowpass filter with $H_L(0) = 1$. Choose the bandwidth of the filter such that $E\left[N_L^2(t)\right] = \frac{1}{2}E\left[N_I^2(t)\right]$.

Problem 4.13. An interesting relationship exists between the power spectral density of a phase noise and the distribution of the instanteous frequency deviation. This problem explores this relationship, which has proven useful in practice, using the techniques introduced in this chapter and in the development of angle modulation. Assume a random process is characterized as

$$N_z(t) = A_c \exp\left[jN_p(t)\right] \tag{4.44}$$

where $N_p(t)$ is a phase noise generated in, for example, a frequency synthesizer.

a) Show that

$$R_{N_z}(t,\tau) = E\left[N_z(t)N_z^*(t-\tau)\right] \approx A_c^2 E\left[\exp\left(j2\pi f_d(t)\tau\right)\right] \tag{4.45}$$

where $f_d(t) = \frac{1}{2\pi}\frac{d}{dt}N_p(t)$.

b) If $f_d(t)$ is a stationary random process, show that

$$R_{N_z}(t,\tau) \approx R_{N_z}(\tau) \approx A_c^2 \Phi_{f_d}(\tau) \tag{4.46}$$

where $\Phi_{f_d}(t)$ is the characteristic function[1] of one sample of the random process $f_d(t)$.

c) Finally show that

$$S_{N_z}(f) \approx A_c^2 f_{f_d}(f) \tag{4.47}$$

where $f_{f_d}(f)$ is the PDF of the random variable $f_d(t)$.

d) Make an engineering assessment of the validity and the shortcomings of the approximations in this problem.

[1]The characteristic function of a random variable was introduced in Problem **??**.24 in this text but is a common tool in the analysis of noise [LG89, Hel91].

4.6 Example Solutions

Problem 4.2.

a) We know $N_0 = 1$ and

$$S_{N_c}(f) = \frac{N_0}{2}|H_c(f)|^2, \tag{4.48}$$

therefore,

$$|H_R(f)|^2 = \begin{cases} 4 & f_c - 1000 \le |f| \le f_c + 3000 \\ 0 & \text{elsewhere} \end{cases} \tag{4.49}$$

Consequently

$$H_R(f) = \begin{cases} 2e^{j\theta(f)} & f_c - 1000 \le |f| \le f_c + 3000 \\ 0 & \text{elsewhere} \end{cases} \tag{4.50}$$

where $\theta(f)$ is an arbitrary phase for the filter transfer function.

b) We know

$$S_{N_c}(f) = \frac{1}{2}S_{N_z}(f - f_c) + \frac{1}{2}S_{N_z}(-f - f_c) \tag{4.51}$$

Therefore

$$S_{N_z}(f) = \begin{cases} 4 & -1000 \le f \le 3000 \\ 0 & \text{elsewhere} \end{cases} \tag{4.52}$$

Solving for $S_{N_I}(f)$ gives

$$S_{N_I}(f) = \frac{S_{N_z}(f) + S_{N_z}(-f)}{4} \tag{4.53}$$

Therefore

$$S_{N_I}(f) = \begin{cases} 1 & -3000 \le f \le -1000 \\ 2 & -1000 \le f \le 1000 \\ 1 & 1000 \le f \le 3000 \\ 0 & \text{elsewhere} \end{cases} \tag{4.54}$$

c) $E[N_I^2(t)] = \int_{-\infty}^{\infty} S_{N_I}(f)df = 8000$

d) Using Prop 4.10 and 4.24

$$f_{N_I(t)N_Q(t)}(n_1, n_2) = \frac{1}{16000\,\pi} \cdot \exp\left[-\frac{n_1^2 + n_2^2}{16000}\right] \tag{4.55}$$

e) Solving for $S_{N_I N_Q}(f)$ gives

$$S_{N_I N_Q}(f) = \frac{S_{N_z}(-f) - S_{N_z}(f)}{j4} \tag{4.56}$$

Therefore,

$$S_{N_I N_Q}(f) = \begin{cases} -j & -3000 \le f \le -1000 \\ j & 1000 \le f \le 3000 \\ 0 & \text{elsewhere} \end{cases} \tag{4.57}$$

4.7 Mini-Projects

Goal: To give exposure

1. to a small scope engineering design problem in communications
2. to the dynamics of working with a team
3. to the importance of engineering communication skills (in this case oral presentations).

Presentation: The forum will be similar to a design review at a company (only much shorter) The presentation will be of 5 minutes in length with an overview of the given problem and solution. The presentation will be followed by questions from the audience (your classmates and the professor). All team members should be prepared to give the presentation.

Project 4.1. A colleague at your company, *Wireless.com*, is working on characterizing the noise in the frontend of an intermediate frequency (IF) receiver that the company is trying to make it's first billion on. The design carrier frequency, f_c, was not documented well by the designer (he left for another startup!). The frequency is known to lie somewhere between f_c=2.5kHz and f_c=8kHz. Your colleague is getting some very anomalous results in his testing and has come to you since he knows you were taught communications theory by a world renowned professor (;-)). The bandpass noise output from the receiver is processed during testing in a programmable I/Q down converter with a carrier frequency $\tilde{f}_c$ as shown in Fig. 4.11. The lowpass filters in the I/Q downconverter is programmed to have a cutoff frequency of $\tilde{f}_c$. The anomalous results your colleague sees are

1. if he chooses $\tilde{f}_c$=5000Hz then the output noise, $N_I(t)$ and $N_Q(t)$, has a bandwidth of 5000Hz,
2. if $\tilde{f}_c$=4000Hz then the output noise, $N_I(t)$ and $N_Q(t)$, has a bandwidth of 4000Hz.

Your colleague captured and stored a sample function of the noise and it is available for downloading at
`http://fitzmac.ee.ucla.edu/~fitz/EE132A.htm`
as `noizin.mat`. Examining this file will be useful to complete this project.

a) Explain why the output noise bandwidth changes as a function $\tilde{f}_c$.

b) Assuming that DSB-AM is the design modulation for the receiver, try and identify the probable design carrier frequency.

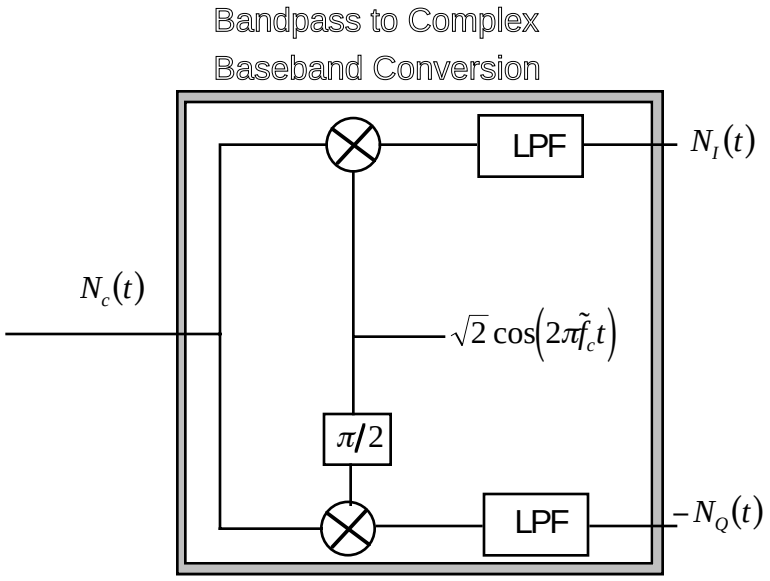


Figure 4.11: Block diagram of an I/Q downconverter with noise.

Chapter 5

Bandpass Gaussian Random Processes

In studying digital communications several additional results for Gaussian random process will be useful. This short chapter quickly reviews this material.

5.1 Preliminaries

The most common situation encountered in communication system design and analysis is for the corrupting noise to be a Gaussian random process. Analytically this is very fortunate since a simple closed form expression for the PDF of L samples from a Gaussian random process is

$$f_{\vec{N}}(\vec{n}) = \frac{1}{\sqrt{(2\pi)^L \det(\mathbf{C}_N)}} \exp\left[-\frac{1}{2}(\vec{n}-\vec{m}_N)^T \mathbf{C}_N^{-1}(\vec{n}-\vec{m}_N)\right] \tag{5.1}$$

where $\vec{N}$ is an $L \times 1$ vector defined as

$$\vec{N} = [N(t_1), ..., N(t_n)]^T,$$

$\vec{m}_N$ is the mean vector of $\vec{N}$, and $\mathbf{C}_N$ is the $L \times L$ covariance matrix of the vector $\vec{N}$. The Gaussian assumption is rather powerful. It accurately models most communication systems and the joint PDF, which is a complete probabilistic description of the random process, is given in terms of only the first and second moments.

Definition 5.1 *A cross-covariance matrix between two real random vectors is*

$$\mathbf{C}_{XY} = E[(\vec{X}-\vec{m}_x)(\vec{Y}-\vec{m}_y)^T] \tag{5.2}$$

Property 5.1 *If $\vec{X}$ and $\vec{Y}$ are both Gaussian random vectors then $\vec{X}$ conditioned on $\vec{Y} = \vec{y}$ is a Gaussian random vector with*

$$E(\vec{X}|\vec{Y}=\vec{y}) = \vec{m}_x + \mathbf{C}_{XY}\mathbf{C}_Y^{-1}(\vec{y}-\vec{m}_y)$$

and

$$\mathbf{C}_{X|Y} = E[(\vec{X}-E(\vec{X}|\vec{Y}=\vec{y}))(\vec{X}-E(\vec{X}|\vec{Y}=\vec{y}))^T] = \mathbf{C}_X - \mathbf{C}_{XY}\mathbf{C}_Y^{-1}\mathbf{C}_{XY}^T.$$

Proof: Setting $\vec{X}$ to be a vector of size L and

$$\vec{Z} = \begin{bmatrix} \vec{X} \\ \vec{Y} \end{bmatrix} \qquad \vec{m}_Z = \begin{bmatrix} \vec{m}_X \\ \vec{m}_Y \end{bmatrix} \qquad \mathbf{C}_Z = \begin{bmatrix} \mathbf{C}_X & \mathbf{C}_{XY} \\ \mathbf{C}_{XY}^T & \mathbf{C}_Y \end{bmatrix}$$

and using the definition of a conditional density function gives

$$\begin{aligned} f_{\vec{X}|\vec{Y}}(\vec{x}|\vec{y}) &= \frac{f_{\vec{X}\vec{Y}}(\vec{x},\vec{y})}{f_{\vec{Y}}(\vec{y})} = \frac{f_{\vec{Z}}(\vec{z})}{f_{\vec{Y}}(\vec{y})} \\ &= \frac{1}{(2\pi)^{L/2}} \frac{\det[\mathbf{C}_Y]}{\det[\mathbf{C}_Z]} \frac{\exp\left[-\frac{1}{2}\left(\vec{z}-\vec{m}_Z\right)^T \mathbf{C}_Z^{-1}\left(\vec{z}-\vec{m}_Z\right)\right]}{\exp\left[-\frac{1}{2}\left(\vec{y}-\vec{m}_Y\right)^T \mathbf{C}_Y^{-1}\left(\vec{y}-\vec{m}_Y\right)\right]}. \end{aligned} \tag{5.3}$$

Eqn. (5.3) can be simplified using the following two well known linear algebra identities

1.

$$\det\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \det[\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}]\det[\mathbf{D}] \tag{5.4}$$

2.

$$\begin{aligned} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} &= \begin{bmatrix} \mathbf{A}^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{I} \end{bmatrix} [\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}]^{-1} \begin{bmatrix} -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{D}^{-1} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -\mathbf{D}^{-1}\mathbf{C} \end{bmatrix} [\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}]^{-1} \begin{bmatrix} \mathbf{I} & -\mathbf{B}\mathbf{D}^{-1} \end{bmatrix} \end{aligned} \tag{5.5}$$

where $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$ are arbitrary matrices. Using (5.4) the ratio of the determinants in (5.3) is given as

$$\frac{\det[\mathbf{C}_Z]}{\det[\mathbf{C}_Y]} = \det[\mathbf{C}_X - \mathbf{C}_{XY}\mathbf{C}_Y^{-1}\mathbf{C}_{YX}] = \det[\mathbf{C}_{X|Y}]. \tag{5.6}$$

Using (5.5) to reformulate $\mathbf{C}_Z$ in (5.3) gives

$$\exp\left[-\frac{1}{2}[\vec{z}-\vec{m}_Z]^T\mathbf{C}_Z^{-1}[\vec{z}-\vec{m}_Z]\right] = \exp\left[-\frac{1}{2}[\vec{y}-\vec{m}_Y]^T\mathbf{C}_Y^{-1}[\vec{y}-\vec{m}_Y]\right] \exp\left[-\frac{1}{2}\mathbf{F}^T\mathbf{G}^{-1}\mathbf{F}\right] \tag{5.7}$$

where

$$\mathbf{F} = \vec{x} - \vec{m}_X - \mathbf{C}_{XY}\mathbf{C}_Y^{-1}[\vec{y}-\vec{m}_Y] = \vec{x} - E[\vec{X}|\vec{Y}=\vec{y}]$$

and

$$\mathbf{G} = \mathbf{C}_X - \mathbf{C}_{XY}\mathbf{C}_Y\mathbf{C}_{XY}^T = \mathbf{C}_{X|Y}$$

Using (5.6) and (5.7) in (5.3) and cancelling the common terms gives the desired result. □

Property 5.2 *The characteristic function (CHF) of a Gaussian vector has the form of*

$$\Phi_{\vec{N}}(\vec{t}) = E\left[\exp\left(j\left(\vec{t}\right)^T\vec{N}\right)\right] = \exp\left(j\left(\vec{t}\right)^T\vec{m}_N - \frac{1}{2}\left(\vec{t}\right)^T\mathbf{C}_N\vec{t}\right). \tag{5.8}$$

The Gaussian assumption is rather powerful. It accurately models the input to most communication systems and the joint PDF (which is a complete probabilistic description of the random process) is given in terms of only the first and second moments. The CHF of a Gaussian random vector is also a simple function of only the the first two moments. Consequently the random variables resulting from many types of signal processing of Gaussian processes can be characterized.

5.2 Amplitude and Phase of Noise

The most common form of noise in a communication system is zero mean, stationary, bandpass Gaussian noise. Recall from Chapter 4 that the joint density of $N_I(t)$ and $N_Q(t)$ at any time instant is given by

$$f_{N_I(t)N_Q(t)}(n_I, n_Q) = \frac{1}{2\pi\sigma_N^2} \exp\left(-\frac{n_I^2 + n_Q^2}{2\sigma_N^2}\right). \tag{5.9}$$

This is a very simple closed form for the PDF of the real and imaginary components of the complex envelope ($N_I(t)$ and $N_Q(t)$) and it will allow us to generate many interesting results.

Often in the course of analyzing communication systems, representing a narrowband noise in terms of the envelope and phase of the complex envelope is more convenient. The PDF of the amplitude and phase can be derived by a simple transformation of random variables

$$N_I(t) = N_A(t)\cos(N_P(t)) \qquad N_Q(t) = N_A(t)\sin(N_P(t)) \tag{5.10}$$

$$N_A(t) = \sqrt{N_I(t)^2 + N_Q(t)^2} \qquad N_P(t) = \tan^{-1}(N_Q(t), N_I(t)) \tag{5.11}$$

with $|J| = N_A(t)$ to produce

$$f_{N_A(t),N_P(t)}(n_a, n_p) = \begin{cases} \frac{n_a}{2\pi\sigma_N^2} \exp\left(-\frac{n_a^2}{2\sigma_N^2}\right) & n_a \geq 0, -\pi \leq n_p \leq \pi \\ 0 & \text{elsewhere.} \end{cases} \tag{5.12}$$

The marginal densities of the amplitude and phase can be obtained by integrating over this joint density. The amplitude PDF is given by

$$f_{N_A(t)}(n_a) = \begin{cases} \frac{n_a}{\sigma_N^2} \exp\left(-\frac{n_a^2}{2\sigma_N^2}\right) & n_a \geq 0 \\ 0 & \text{elsewhere.} \end{cases} \tag{5.13}$$

This is known as the Rayleigh PDF. The PDF of the phase is also easily obtained and, as one would expect, it is uniform, i.e.,

$$f_{N_P(t)}(n_p) = \frac{1}{2\pi} \qquad -\pi \leq \theta(t) \leq \pi.$$

Since $f_{N_A(t),N_P(t)}(n_a, n_p) = f_{N_P(t)}(n_p) f_{N_A(t)}(n_a)$ the envelope and phase of bandpass Gaussian noise are independent random variables at any time instant.

5.3 The Amplitude of Signal Plus Noise

If a complex baseband signal has the form $Y_z(t) = x_z(t) + N_z(t)$ with $N_z(t)$ as above, then the PDF of $Y_z(t)$ is given by

$$f_{Y_I(t),Y_Q(t)}(y_i, y_q) = \frac{1}{2\pi\sigma_N^2} \exp\left(-\frac{(y_i - x_I(t))^2 + (y_q - x_Q(t))^2}{2\sigma_N^2}\right)$$

using the same transformation as above, the joint PDF of the envelope and phase can be derived

$$f_{Y_A(t),Y_P(t)}(y_a, y_p) = \begin{cases} \frac{y_a}{2\pi\sigma_N^2} \exp\left(-\frac{y_a^2 + x_A(t)^2 - 2y_a x_A(t)\cos(y_p - x_P(t))}{2\sigma_N^2}\right) & y_a \geq 0,\ -\pi \leq y_p \leq \pi \\ 0 & \text{elsewhere.} \end{cases} \tag{5.14}$$

Note that the PDF of a baseband signal plus baseband noise cannot be expressed in the form

$$f_{Y_A(t),Y_P(t)}(y_a, y_p) = f_{Y_A(t)}(y_a) f_{Y_P(t)}(y_p). \tag{5.15}$$

This fact implies that the envelope and phase of a signal plus Gaussian noise at any time instant are dependent random variables. Again the individual marginal densities can be obtained by integrating over this joint density.

In communications an important PDF is that of the envelope or the amplitude of the complex envelope. The marginal density of the envelope or the amplitude of the complex envelope of a signal in bandpass Gaussian noise is given by

$$f_{Y_A(t)}(y_a) = \int_{-\pi}^{\pi} f_{Y_A(t),Y_P(t)}(y_a, y_p) dy_p \tag{5.16}$$

which is simplified to

$$\begin{aligned} f_{Y_A(t)}(y_a) &= \begin{cases} \dfrac{y_a}{\sigma_N^2} \exp\left(-\dfrac{y_a^2 + x_A(t)^2}{2\sigma_N^2}\right) \left[\dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} \exp\left[\dfrac{y_a x_A(t) \cos(y_p - x_P(t))}{\sigma_N^2}\right] dy_p\right] & y_a \geq 0 \\ 0 & \text{elsewhere} \end{cases} \\ &= \begin{cases} \dfrac{y_a}{\sigma_N^2} \exp\left(-\dfrac{y_a^2 + x_A(t)^2}{2\sigma_N^2}\right) I_0\left(\dfrac{y_a x_A(t)}{\sigma_N^2}\right) & y_a \geq 0 \\ 0 & \text{elsewhere.} \end{cases} \end{aligned} \tag{5.17}$$

where $I_0(x)$ is the modified Bessel function of order zero [Ae72]. The modified Bessel function of order zero is a common transcendental function appearing in the analysis of communication systems. It is defined as

$$I_0(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(x \cos\theta) d\theta.$$

The modified Bessel function of order zero is evaluated by a table look-up [Ae72], or, more commonly, numerical techniques on a computer ([PFTV86] or Matlab). The PDF given in (5.17) is known as the Ricean distribution (or Nakagami-Rice distribution) [Ric45, Ric48, Nak60]. Fig. 5.1 is a plot of the Rayleigh and Ricean PDF for various values of SNR[1] with the variance of the noise normalized to unity.

The complimentary distribution function of a Ricean PDF arises in many communications and radar problems. Marcum's Q-function is a transcendental function used to characterize this tail probability and is defined as

$$Q(a, b) = \int_b^{\infty} x \exp\left[-\frac{a^2 + x^2}{2}\right] I_0(ax) dx \tag{5.18}$$

Marcum's Q-function has been included in some toolboxes of Matlab and well studied numerical techniques are available in the literature [Shn89].

Examining Fig. 5.1 one can see that the Ricean PDF has the following characteristics:

1. When $x_A(t) = 0$ the Ricean PDF becomes the Rayleigh PDF. Hence the Rayleigh PDF is a special case of the Ricean PDF.

2. When the SNR becomes very large the Ricean PDF becomes approximately Gaussian. The intuitive reason is seen by examining the transformation

$$Y_A(t) = \sqrt{Y_I^2(t) + Y_Q^2(t)}$$

[1] $SNR = \frac{x_A(t)^2}{2\sigma_N^2}$

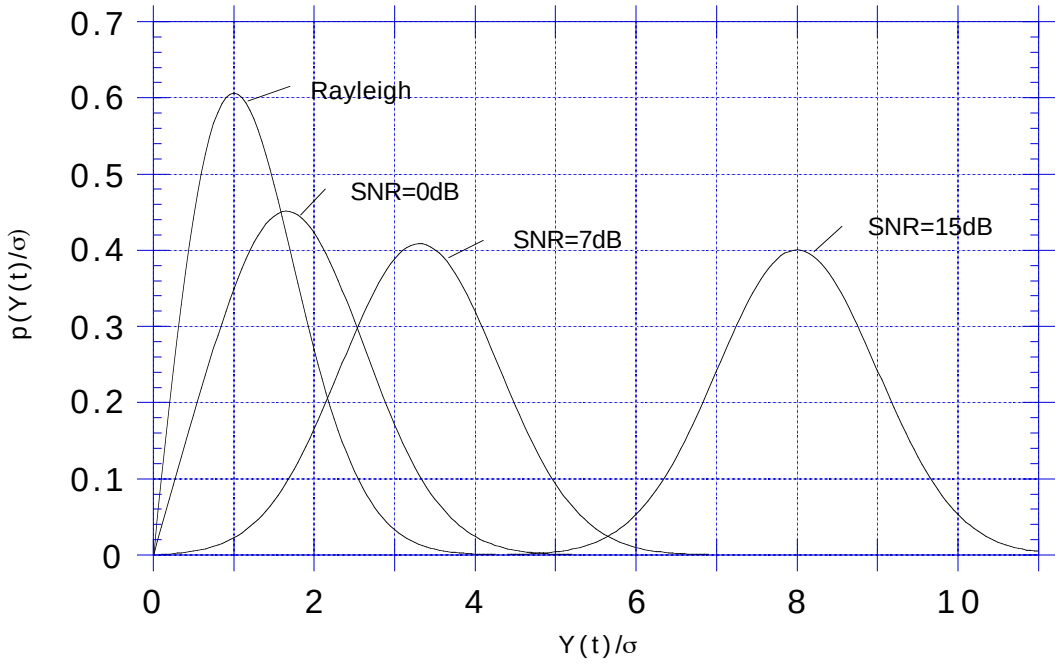


Figure 5.1: The normalized Ricean PDF for various SNRs.

and noting that

$$\begin{aligned} Y_z(t) &= \exp\left[jx_P(t)\right]\left(x_A(t) + N_z(t)\exp\left[-jx_P(t)\right]\right) \\ &= \exp\left[jx_P(t)\right]\left(x_A(t) + \tilde{N}_z(t)\right). \end{aligned} \tag{5.19}$$

The form for $Y_A(t)$ can be simplified to

$$Y_A(t) = \sqrt{(x_A(t) + \tilde{N}_I(t))^2 + \tilde{N}_Q^2(t)}.$$

At high SNR the quadrature noise, $\tilde{N}_Q(t)$, will be negligible compared to $x_A(t)$ and hence

$$Y_A(t) \approx x_A(t) + \tilde{N}_I(t). \tag{5.20}$$

Equation (5.20) implies at high SNR that only one component of the bandpass noise affects the envelope of a signal plus bandpass noise.

5.4 The Phase of Signal Plus Noise

Again if a complex baseband signal has the form $Y_z(t) = x_z(t) + N_z(t)$ with $N_z(t)$ as above, then the PDF of $Y_z(t)$ is given by

$$f_{Y_I(t),Y_Q(t)}(y_i, y_q) = \frac{1}{2\pi\sigma_N^2}\exp\left(-\frac{(y_i - x_I(t))^2 + (y_q - x_Q(t))^2}{2\sigma_N^2}\right)$$

using the same transformation as above, the joint PDF of the envelope and phase can be derived

$$f_{Y_A(t),Y_P(t)}(y_a,y_p) = \begin{cases} \dfrac{y_a}{2\pi\sigma_N^2}\exp\left(-\dfrac{y_a^2+x_A(t)^2-2y_a x_A(t)\cos(y_p-x_P(t))}{2\sigma_N^2}\right) & y_a\geq 0,\ -\pi\leq y_p\leq\pi \\ 0 & \text{elsewhere.} \end{cases} \tag{5.21}$$

The marginal density of the phase of a signal in bandpass noise is given by

$$f_{Y_P(t)}(y_p) = \int_0^\infty f_{Y_A(t),Y_P(t)}(y_a,y_p)dy_a. \tag{5.22}$$

This can be evaluated by completion of the square, i.e.,

$$f_{Y_P(t)}(y_p) = \begin{cases} \dfrac{1}{2\pi\sigma_N^2}\exp\left(-\dfrac{x_A(t)^2\sin^2(y_p-x_P(t))}{2\sigma_N^2}\right)\displaystyle\int_0^\infty y_a\exp\left[-\frac{(y_a-x_A(t)\cos(y_p-x_P(t)))^2}{2\sigma_N^2}\right]dy_a \\ \qquad -\pi < y_p\leq\pi \\ 0 \qquad \text{elsewhere} \end{cases} \tag{5.23}$$

Making the change of variables

$$u = \frac{y_a - x_A(t)\cos(y_p-x_P(t))}{\sigma_N} \qquad a = -\frac{x_A(t)\cos(y_p-x_P(t))}{\sigma_N}$$

in the integral in (5.23) produces

$$\begin{aligned} f_{Y_P(t)}(y_p) &= \frac{1}{2\pi}\exp\left(-\frac{x_A^2(t)\sin^2(y_p-x_P(t))}{2\sigma_N^2}\right)\int_a^\infty u\exp\left(-\frac{u^2}{2}\right)du \\ &+ \frac{x_A(t)\cos(y_p-x_P(t))}{2\pi\sigma_N}\exp\left(-\frac{x_A^2(t)\sin^2(y_p-x_P(t))}{2\sigma_N^2}\right)\int_a^\infty\exp\left(-\frac{u^2}{2}\right)du. \end{aligned} \tag{5.24}$$

Both integrals are well known so gain making a final change of variables

$$P = \frac{x_A^2(t)}{2\sigma_N^2} = SNR$$

gives the final form of

$$f_{Y_P(t)}(y_p) = \begin{cases} \dfrac{\exp[-P]}{2\pi} + \sqrt{\dfrac{P}{4\pi}}\cos(y_p-x_P(t))\exp(-P\sin^2(y_p-x_P(t)))\left[1+\operatorname{erf}\left(\sqrt{P}\cos(y_p-x_P(t))\right)\right] \\ \qquad -\pi\leq y_p\leq\pi \\ 0 \qquad \text{elsewhere} \end{cases} \tag{5.25}$$

Fig. 5.2 is a plot of the PDF of a sample of the phase error, $E_P(t), (E_P(t) = Y_P(t) - x_P(t))$ for various values of SNR. Examining Fig. 5.2 one can see that this PDF has the following characteristics:

1. When $x_A(t) = 0$ the PDF becomes the uniform PDF.
2. When the SNR becomes very large the PDF becomes approximately Gaussian. Recall again that

$$\begin{aligned} Y_z(t) &= \exp\left[jx_P(t)\right]\left(x_A(t)+N_z(t)\exp\left[-jx_P(t)\right]\right) \\ &= \exp\left[jx_P(t)\right]\left(x_A(t)+\tilde{N}_z(t)\right). \end{aligned} \tag{5.26}$$

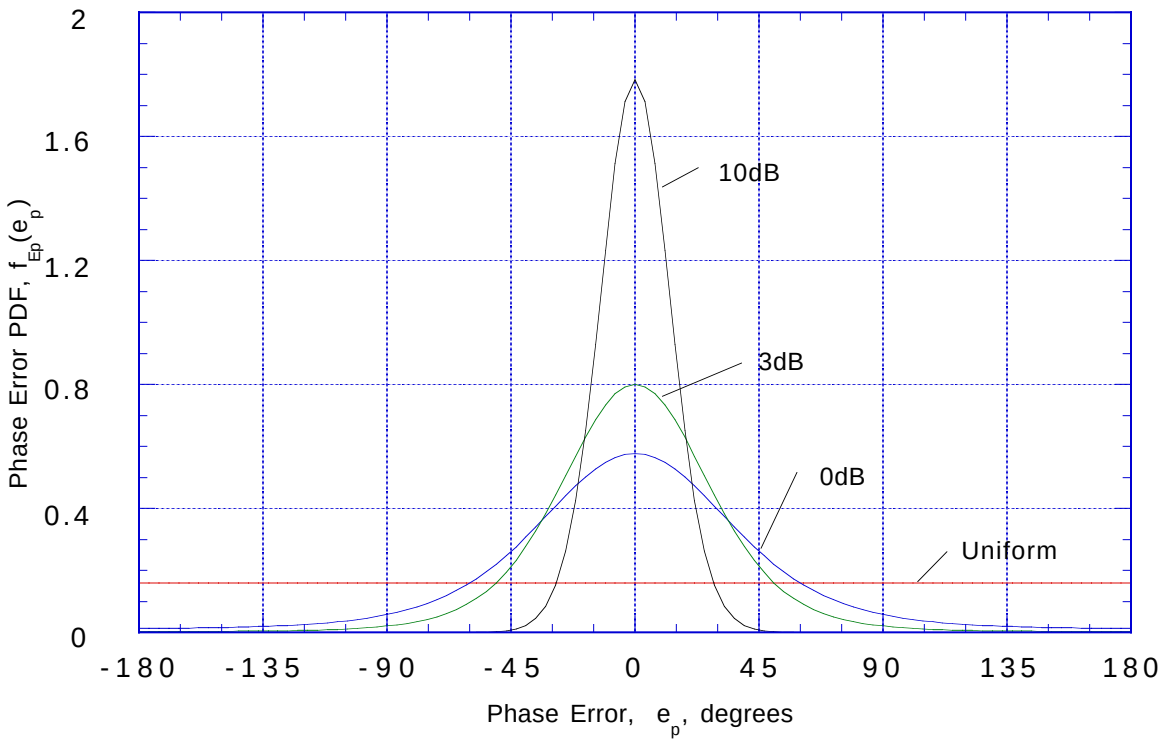


Figure 5.2: The phase error PDF of a signal in bandpass Gaussian noise.

The intuitive reason for the convergence to a Gaussian is seen by examining the transformation

$$E_P(t) = \tan^{-1}\left(\frac{\tilde{N}_Q(t)}{x_A(t) + \tilde{N}_I(t)}\right)$$

At high SNR the in-phase noise, $\tilde{N}_I(t)$, will be negligible compared to $x_A(t)$ and hence

$$E_P(t) \cong \tan^{-1}\left(\frac{\tilde{N}_Q(t)}{Z(t)}\right) \cong \frac{\tilde{N}_Q(t)}{Z(t)} \tag{5.27}$$

Equation (5.27) implies that at high SNR the phase error approximately becomes the quadrature noise scaled by the signal amplitude. Consequently at high SNR the phase of a signal in bandpass noise is approximately Gaussian with a variance of

$$E\left[E_P^2(t)\right]^2 = \frac{1}{2P}$$

Fig. 5.3 is a plot of the rms phase error obtained from (5.25) versus SNR that demonstrates this point.

It is interesting to note that at high SNR the bandpass noise and the resulting effects become separated into in-phase with the signal (amplitude) and in-quadrature with the signal (phase). [DR87] is a good source for a development of similar topics as presented in the last three sections. [Mid60] provides many generalizations to the simple results the last three sections.

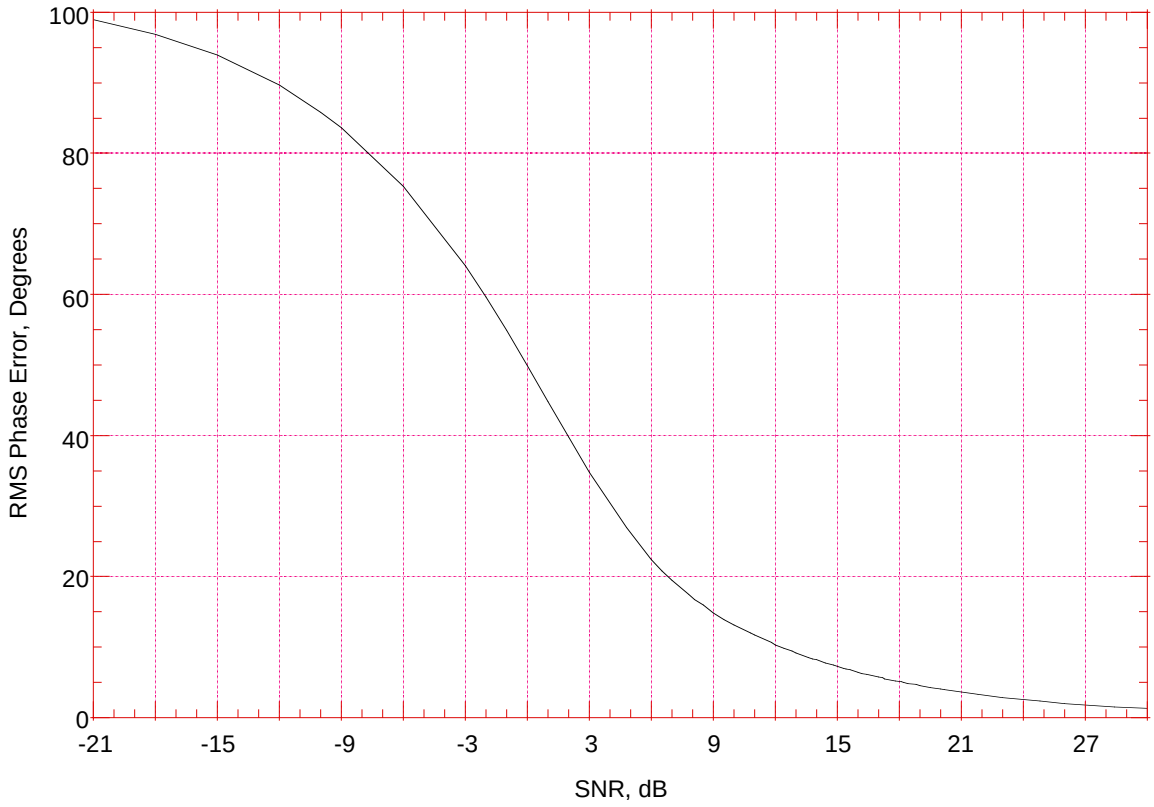


Figure 5.3: The RMS phase error of a signal in bandpass noise.

5.5 Complex Gaussian Random Processes

Complex Gaussian processes in communication system analysis result from examining the complex baseband representation of a bandpass random process. Good references for complex Gaussian random processes are [Mil64, Mil74].

Definition 5.2 *A complex process, $N_z(t) = N_I(t) + jN_Q(t)$, is a complex random process if*

1. $cov(N_z(t), N_z * (s)) = R_{N_z}(t, s)$
2. $cov(N_z(t), N_z(s)) = 0$.

This definition requires that $N_I(t)$ and $N_Q(t)$ have the same statistical description and that the cross correlation function is an "odd" function. This holds for the complex envelope of a process produced by bandpass filtering an AWGN.

Definition 5.3 *A complex process is a complex Gaussian process if $N_I(t)$ and $N_Q(t)$ are jointly Gaussian.*

5.5.1 Complex Gaussian Distribution

The density function of a complex Gaussian random vector (samples of a complex Gaussian process) has the form

$$f_{\vec{N}_z}(n_1, \ldots, n_N) = \frac{1}{\pi^N \det \mathbf{C}_{N_z}} \exp\left(-\left(\vec{n} - \vec{m}_{N_z}\right)^H \mathbf{C}_{N_z}^{-1} \left(\vec{n} - \vec{m}_{N_z}\right)\right) \tag{5.28}$$

where

$$\begin{aligned} \vec{N}_z &= [N_z(t_1) \ldots N_z(t_N)]^T \\ \vec{m}_{N_z} &= E\left[\vec{N}_z\right] = [m_{N_z}(t_1) \ldots m_{N_z}(t_N)]^T \\ \mathbf{C}_{N_z} &= E[(\vec{N}_z - \vec{m}_{N_z})(\vec{N}_z - \vec{m}_{N_z})^H] \end{aligned}$$

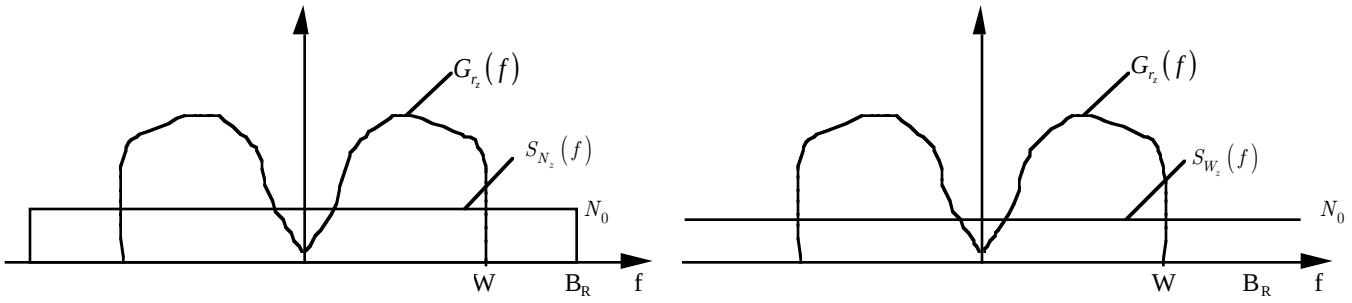
The PDF given in (5.28) for an N dimensional complex Gaussian vector is just a compact way of expressing the $2N$ dimensional real Gaussian vector PDF. The compact form for the PDF of samples of a complex Gaussian random process makes complex Gaussian random processes amenable to analysis and algorithm development.

5.5.2 Complex Additive White Gaussian Noise

A complex additive white Gaussian noise can be used to accurately model bandpass noise in communication systems. Consider a bandpass noise generated by an ideal bandpass filter, e.g.,

$$H_R(f) = \begin{cases} 1 & f_c - B_i \leq |f| \leq f_c + B_i \\ 0 & \text{elsewhere.} \end{cases} \tag{5.29}$$

The baseband (complex envelope) power spectrum of the bandpass noise, $S_{N_z}(f)$, and signal will look like that in Fig. 5.4-a). Since it is, in general, difficult to build a filter at a carrier frequency that is spectrally compact, the noise spectrum will typically be wider than the signal spectrum. The baseband processing will normally cut the noise bandwidth down to something close to W. A "white" noise that has the PSD as given in Fig. 5.4-b) will have the exact same output noise characteristics after baseband processing because



a) Signal and bandpass noise spectra b) Signal and complex white noise spectra

Figure 5.4: Exact and model power spectral densities.

1. the baseband processing will only pass the noise in the message bandwidth, W, and
2. the bandpass noise has a complex envelope that has a PSD that is constant, like the white noise, over this band.

Consequently, and accurate model for noise in a bandpass communication system is given as

$$Y_z(t) = r_z(t) + W_z(t) \tag{5.30}$$

where $r_z(t)$ is the received signal and $W_z(t)$ is the model for the corrupting noise where

$$S_{W_z}(f) = N_0 \qquad R_{W_z}(\tau) = N_0\delta(\tau). \tag{5.31}$$

This model will be denoted the complex additive white Gaussian noise in the sequel.

A complex AWGN is intended to model a bandpass noise whose bandwidth is larger than the received signal bandwidth. The advantage of the complex AWGN model is that the analysis and algorithm development can proceed in a much cleaner fashion than with a more realistic noise model while not introdcing significant modeling errors. This model will be used exclusively in the sequel and the accuracy of the approximation will be explored in the homework.

5.6 Quadratic Forms of Complex Gaussians

Will be needed to analyze noncoherent and transmitted reference demodulation structures. Not completed in this edition.

5.7 Homework Problems

Problem 5.1. As simple model of a radar receiver consider the following problem. A radar sends out a bandpass pulse and listens for the return at the receiver whose filtering is represented with an impulse response $h_R(t)$. The situation when no target is present is represented with the input being modeled as a real valued additive Gaussian noise, $W(t)$, being input into a linear time invariant bandpass filter with a real valued impulse response $h_R(t)$. $W(t)$, has a two sided spectral density $N_0/2$) Denote the baseband output as $Y_z(t) = N_z(t)$. The situation where a target is present can be represented as a tone plus white noise. The baseband output is then given as $Y_z(t) = A\exp[j\phi_p] + N_z(t)$. A target detection decision is given by a simple threshold test at time t

$$Y_A(t) = |Y_z(t)| \underset{\text{Target Absent}}{\overset{\text{Target Present}}{\gtrless}} \gamma \tag{5.32}$$

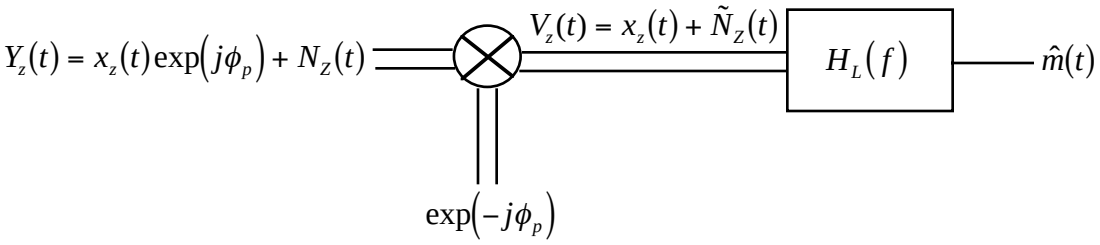


Figure 5.5: A demodulator block diagram.

a) For the case when no target is present, completely characterize a sample of the process $Y_A(t)$.

b) For a particular time t, find a value of the threshold, γ, where the probability that a target will be declared present when one is not present is 10^{-5}.

c) For the case when a target is present completely characterize a sample of the process $Y_A(t)$.

d) For the value of the threshold selected in b) find how large A should be to correctly detect a target present with a probability of detection greater than 0.9 for any value of t.

Problem 5.2. Prove that

$$\int_{-\pi/2}^{\pi/2} f_{Y_p(t)}(y)dy = \frac{1}{2}\text{erfc}\left(\sqrt{P}\right). \tag{5.33}$$

Problem 5.3. In a DSB-AM receiver modeled with Fig 5.5, the following complex envelope is received

$$Y_z(t) = \cos(200\pi t)\exp[j\phi_p] + N_z(t) \tag{5.34}$$

Assume $\hat{m}(t) = m_e(t) + \tilde{N}_I(t)$ and for simplicity that the bandpass noise, $N_z(t)$ is characterized as a white noise

$$S_{N_z}(f) = N_0. \tag{5.35}$$

The lowpass filter has a transfer function of

$$H_L(f) = \frac{1}{1 + j\frac{f}{f_l}} \tag{5.36}$$

a) Prove $\tilde{N}_I(t)$ is a stationary random process.

b) Compute $m_e(t)$ (the signal at the filter output).

c) Compute $E\left[\tilde{N}_I^2(t)\right]$ as a function of f_l. *Hint: power can be computed in the time and frequency domain.*

d) Compute the output SNR as a function of f_l.

e) What value of f_l optimizes the output SNR?

Problem 5.4. A computer can do a lot of tedious computations that you do not want to do by hand. To this end repeat the Problem 5.3 except with the bandpass noise characterized with

$$S_{N_z}(f) = \begin{cases} N_0 & |f| \leq 4000 \\ 0 & \text{elsewhere.} \end{cases} \tag{5.37}$$

Hint: the simplest approach may not be the same as in Problem 5.3.

Part II

Digital Communication Basics

Chapter 6

Digital Communication Basics

The problem that is of interest in the text is point–to–point binary data communications. The system model for such a communication system is given in Fig. 6.1. A source of binary encoded data with a total of K_b bit is present and it is desired to transmit this data to a binary data sink across a physical channel. The data output by the source is represented by a $K_b \times 1$ dimensional vector, $\vec{I}$, whose components take values 0,1. This vector is then mapped into one of 2^{K_b} analog waveforms represented by the baseband waveform $x_i(t)$. The transmitted waveform is put through a channel of some sort and corrupted by a noise or interference. The composite received signal, $Y_c(t)$, is then used to estimate which one of the possible vectors led to the transmitted waveform. This estimate is denoted $\hat{\vec{I}}$ and this estimate is passed to the data sink.

It should be noted that point-to-point data communications is but the simplest abstraction of the data communications reality. This point to point communication is often referred to as physical layer communications. There are many higher layers of abstraction in data communication. These higher layers deal with concepts like how do multiple users access a shared communication media, how is a communication path established in a multiple hop network, how do applications communication across the network. These issues are typically covered in the curriculum in a course on communication networks. Typical textbooks for communications networks are [KR04, LGW00, Tan02]. The student interested in the details of networked communication should take these courses or read these textbooks.

6.1 Digital Transmission

In exactly the same way as in analog communication, digital communication has a modulator and demodulator. The modulator produces an analog signal that depends on the digital data to be communicated. This analog signal is transmitted over a channel (a cable or radio propagation). The demodulator takes the received signal and constructs an estimate of the transmitted digital data. There

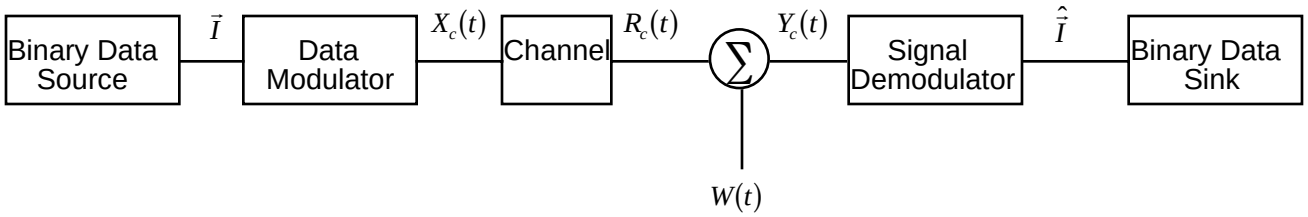


Figure 6.1: A model for point to point data communication.

are two types of digital communications that are implemented in practice: Baseband data communications as exemplified by compact disc recordings or magnetic recording and carrier modulated data communications as exemplified by telephone modems or radio modems. In this text we want to have a unifying framework that enables the material that is learned to apply to either baseband or bandpass data communications. As with analog communications, the complex envelope notation is used to achieve this goal. The only caveat that needs to be stated is that baseband data communication will always have a zero imaginary component, while for bandpass communications the imaginary component of the complex envelope might be nonzero.

6.1.1 Digital Modulation

Definition 6.1 *Digital modulation is a transformation of $\vec{I}$ into a complex envelope, $X_z(t)$.*

This transformation is equivalent to transforming $m(t)$ into a bandpass signal, $X_c(t)$. The digital modulation process, $X_z(t) = \Gamma_m\left(\vec{I}\right)$ is represented in Fig. 6.2. It should be emphasized that digital communication is achieved by producing and transmitting analog waveforms. There is no situation where communication takes place that this transformation from digital to analog does not occur. Similarly it should be emphasized that the modulator must be capapble of generating 2^{K_b} continuous time waveforms to represent each of the possible binary data vectors produced by the binary data source. When necessary, the possible data vector will be enumerated as $\vec{I} = i$. $i \in \left\{0, \ldots, 2^{K_b} - 1\right\}$ and the possible transmitted waveforms will be enumerated with $X_z(t) = x_i(t)$.

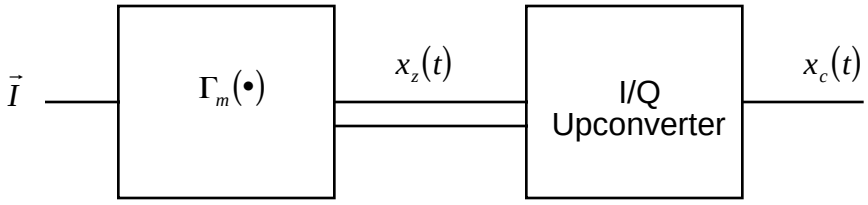


Figure 6.2: The digital modulation process. Note the IQ upconverter is given in Fig. 2.4

6.1.2 Digital Demodulation

Definition 6.2 *Digital demodulation is a transformation of $Y_c(t)$ into estimates of the transmitted bits, $\hat{\vec{I}}$. $X_z(t)$.*

Demodulation takes the received signal, $Y_c(t)$, and downconverts to the baseband signal, $Y_z(t)$. The baseband signal is then processed to produce an estimate of the transmitted data vector, $\vec{I}$. This estimate will be denoted $\hat{\vec{I}}$. Again, for this part of the text the channel output is always assumed to be

$$r_c(t) = L_p x_c\left(t - \tau_p\right)$$

where L_p is the propagation loss and τ_p is the propagation time delay. Define $\phi_p = -2\pi f_c \tau_p$ so that the channel output is given as

$$\begin{aligned} r_c(t) &= \sqrt{2} L_p x_A(t-\tau_p)\cos\left(2\pi f_c(t-\tau_p) + x_P(t-\tau_p)\right) \\ &= \Re\left[\sqrt{2} L_p x_z(t-\tau_p)\exp[j\phi_p]\exp[j2\pi f_c t]\right]. \end{aligned} \tag{6.1}$$

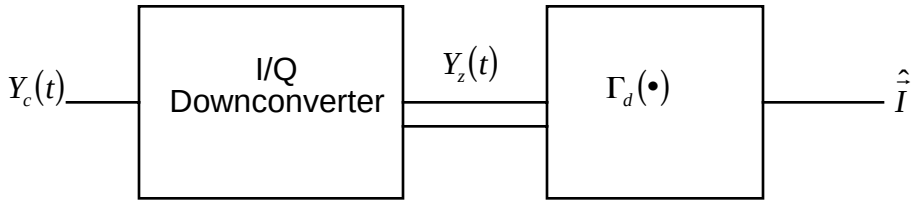


Figure 6.3: The digital demodulation process. Note the IQ downconverter is given in Fig. 2.4.

It is obvious from (6.1) that the received complex envelope is $r_z(t) = L_p x_z(t - \tau_p) \exp[j\phi_p]$. It is important to note that a time delay in a carrier modulated signal will produce a phase offset. The demodulation process conceptually is a down conversion to baseband and a reconstruction of the transmitted signal from $Y_z(t)$. The block diagram for the demodulation process is seen in Fig. 6.3.

Demodulation is the process of producing an $\hat{\vec{I}}$ from $Y_z(t)$ via a function $\Gamma_d\left(Y_z(t)\right)$. The remainder of the discussion on digital communications in this text will focus on identifying $\Gamma_m\left(\vec{I}\right)$ (modulators) and $\Gamma_d\left(Y_z(t)\right)$ (demodulators) and assessing the performance of these modulators and demodulators in the presence of noise. It is worth noting at the point that the word modem is actually an engineering acronym for a device that was both a *mo*dulator and a *demo*dulator that has become part of the English language. The word modem is now synonymous with any device that is used to transmit digital data (computer modem, cable modem, wireless modem, etc.)

6.2 Performance Metrics for Digital Communication

In evaluating the efficacy of various designs in this course the performance metrics commonly applied in engineering design must be examined. The most commonly used metrics for digital communications are

- Performance – This metric typically measures how often data transmission errors are made given the amount of transmitted power.
- Complexity – This metric almost always translates directly into cost.
- Spectral Efficiency – This metric measures how much bandwidth a modulation uses to implement the communication.

6.2.1 Performance

Performance in digital communication is reflected how often transmission errors occur as a function of the SNR. Transmission errors can be either bit errors (one bit in error) or frame errors (any error in a message or packet). The application often determines what the appropriate error metric is. With data transmission at a fixed transmit power, P_{x_c}, the reliability of any data communication can be increased by lowering the speed of the data communication. A lower speed transmission implies that the receiver bandwidth will be smaller and consequently the SNR can be made higher. In data communication, a transmission rate fair measure of SNR is the ratio of the average received energy per bit over the noise spectral density, E_b/N_0.

Definition 6.3 *The average energy per bit for a received signal, $R_z(t)$, where K_b bits are transmitted*

is

$$E_b = \frac{E\left[E_{R_z}\right]}{K_b} = \frac{L_p^2}{K_b} E\left[\int_{-\infty}^{\infty} |X_z(t)|^2 \, dt\right] \tag{6.2}$$

Hence throughout this text we shall parameterize performance by E_b/N_0 as that is the industry standard. It should be noted that the expectation or average in (6.2) is over the random transmitted data bits. Since the random transmitted bit are discrete random variables the expectation will be a summation versus the probability of each of the possible transmitted words. Consequently the general form for the average received energy per bit is

$$E_b = \frac{L_p^2}{K_b} \sum_{i=0}^{2^{K_b}-1} \pi_i \left[\int_{-\infty}^{\infty} |x_i(t)|^2 \, dt\right] = \frac{L_p^2}{K_b} \sum_{i=0}^{2^{K_b}-1} \pi_i E_i \tag{6.3}$$

where E_i is the energy of waveform $x_i(t)$.

6.2.2 Complexity

Complexity is a quantity that requires engineering judgment to estimate. Often the cost of a certain level of complexity changes over time. A good example of this tradeoff changing over a short period of time was seen in the land mobile telephony market in the 1990's when early in the decade many people resisted a move to a standard based on code division multiple access technologies based on the cost and complexity of the handheld phones. By the end of the decade the proposed telecommunication standards had become much more complex but the advances in circuit technology allowed low cost implementations.

6.2.3 Bandwidth Efficiency

The spectral efficiency of a communication system is typically a measure of how well the system is using the bandwidth resource. In this text the bit rate of communication is denoted W_b bits/second and the transmission bandwidth is denoted B_T. Bandwidth costs money to acquire and the owners of this bandwidth want to communicate at as high a data rate as possible. Examples are licenses to broadcast radio signals or the installation of copper wires to connect two points. Hence spectral efficiency is very important for people who try to make money selling communication resources. For instance if one communication system has a spectral efficiency that is twice the spectral efficiency of a second system then the first system can support twice the users on the same bandwidth. Twice the number of users implies twice the revenue. The measure of bandwidth efficiency that we will use in this class will be denoted spectral efficiency and is defined as

$$\eta_B = \frac{W_b}{B_T} \text{bits/second/Hz.}$$

The goal of this section is to associate a spectral characteristic or a signal bandwidth with a digital modulation. This spectral characteristic determines the bandwidth that a radio needs to have to support the transmission, as well as the spectral efficiency of a digital transmission scheme. The way this will be done in this text is to note that if the data being transmitted is known then the transmitted signal, $x_z(t)$, is a deterministic energy signal. The spectral characterization of deterministic energy signals is given by the energy spectrum

$$G_{x_z}(f) = |X_z(f)|^2 = \mathcal{F}\left\{R_{x_z}(\tau)\right\} = \mathcal{F}\left\{\int_{-\infty}^{\infty} x_z(t) x_z^*(t-\tau) dt\right\}. \tag{6.4}$$

Building upon this spectral characterization in the same way as was done for Gaussian random processes, the function that will be used throughout these notes to describe the spectral characteristics of a transmitted signal is the average energy spectrum per bit.

Definition 6.4 *The average energy spectrum per bit for a transmitted signal, $X_z(t)$, where K_b bits are transmitted is*

$$D_{X_z}(f) = \frac{E\left[G_{X_z}(f)\right]}{K_b}. \tag{6.5}$$

This definition of average energy spectrum per bit (measured in Joules per bit per Hertz) is consistent with the definition of power spectral density for random processes as given in 3.4 (measured in Joules per second per Hertz). It should be noted that the expectation or average in (6.5) is over the random transmitted data bits. Since the random transmitted bit are discrete random variables the expectation will be a summation versus the probability of each of the possible transmitted words. Consequently the general form for the average energy spectrum per bit is

$$D_{X_z}(f) = \frac{1}{K_b} \sum_{i=0}^{2^{K_b}-1} \pi_i G_{x_i}(f). \tag{6.6}$$

The transmission bandwidth, B_T, of a digital communications signal can be obtained from the average energy spectrum per bit, $D_{X_z}(f)$.

6.2.4 Other Important Characteristics

Many times in communication applications other issues besides bandwidth efficiency, complexity, and performance are important. For example for a handheld mobile device the size, weight, and battery usage are important for the user. Often in wireless communications energy efficiency of the algorithms are of paramount importance. Often to increase the talk time for a mobile phone an algorithm will give up performance for using less energy. One important issue in mobile devices is the linearity of final amplifier before the antenna. A high power linear amplifier is both expensive and consumes larger amounts of current, so this is not a desirable characteristic for a mobile device. On the other hand, high power is often needed to communicate with a remote basestation or a satellite. Likewise nonlinear amplifiers, which are more energy efficient, often produce unacceptable distortion or spectral regrowth. In many handheld devices modulations are chosen to minimize the requirement on the linearity of the power amplifier. This will be discussed in some detail in Chapter **??**. These issues will not be a major focus of this text like the bandwidth efficiency, complexity, and performance, but are mentioned as they are often involved in performance tradeoffs for system design.

6.3 Some Limits on Performance of Digital Communication Systems

Digital communications is a relatively unique field in engineering in that there is a theory that gives some performance limits for data transmission. The body of work that provides us with these fundamental limits is information theory and the founder of information theory was Claude Shannon [Sha48]. While this text cannot derive all the important results from information theory it will attempt to highlight the important results in information theory that relate to the material in this text. A course in information theory is highly recommended [CT92].

An important contribution of Claude Shannon was to identify that every channel had an associated capacity, C and reliable (in fact error free) transmission is possible on the channel when $W_b < C$. A channel of significant interest for a majority of this text is the channel which experiences an additive

white Gaussian noise (AWGN) distortion. For this AWGN channel when the signal uses a transmission bandwidth B_T, Shannon identified the capacity as [Sha48]

$$C = B_T \log_2 (1 + SNR) . \tag{6.7}$$

This immediately leads to a constraint on the spectral efficiency that can be reliably achieved

$$\eta_B < \log_2 (1 + SNR) . \tag{6.8}$$

Equation (6.8) unfortunately states that to achieve a linear increase in spectral efficiency a communication engineer must provide exponentially greater received SNR. Hence in most communication system applications the spectral efficiencies achieved are usually less than 15bits/s/Hz (often much less).

Further insight into the problem is gained by reformulating (6.8). First off the received noise power is directly a function of the transmission bandwidth. If we assume an ideal bandpass filter of bandwidth B_T the results of Chapter 3 give the noise power and SNR as

$$P_N = N_0 B_T \qquad SNR = \frac{P_S}{N_0 B_T} \tag{6.9}$$

Recall that most communication system engineers like to quantify performance with E_b/N_0 and that $P_s = E_b W_b$ so that (6.8) becomes

$$\eta_B < \log_2 \left(1 + \frac{E_b}{N_0} \frac{W_b}{B_T} \right) = \log_2 \left(1 + \frac{E_b}{N_0} \eta_B \right) . \tag{6.10}$$

The achievable spectral efficiency versus E_b/N_0 is represented in Fig. 6.4. The line in Fig. 6.4 represents the solutions to the equation

$$\eta_B = \log_2 \left(1 + \frac{E_b}{N_0} \eta_B \right) . \tag{6.11}$$

For a given E_b/N_0, information theory indicates that reliable communication at spectral efficiencies below the line in Fig. 6.4 are achievable while spectral efficiencies above the line are not achievable. Throughout the remainder of the text the goal will be to give an exposition on how to design communication systems that have operating points which can approach this ultimate performance given in Fig. 6.4.

The results in Fig. 6.4 provide some interesting insights for how communication systems should be designed. In situations where bandwidth is the most restricted resource the goal then is to drive the received E_b/N_0 to as large a value as possible. For example many telecommunication systems have designed operating points where $E_b/N_0 > 10$dB. In situations where E_b is the most restricted resource it is possible to still achieve reliable communication by reducing the spectral efficiency. For example communications with deep space probes is limited by the amount of power that can be received. Communication systems for deep space communication are designed most often to have relatively low bit rates and by setting $\eta_B < 1$. Also you can see from Fig. 6.4 that there is a limit on how small E_b/N_0 can be made and still maintain reliable communications. This minimum is $E_b/N_0 = \ln 2 = -1.59$dB. These results from information theory provide benchmarks by which we can calibrate performance as we progress in our understanding of digital communication theory.

6.4 Homework Problems

Problem 6.1. If the telephone network was well modeled by an AWGN channel with $B_T = 3.2$kHz and an $E_b/N_0 = 30$dB what bandwidth of information transmission could reliably be supported?

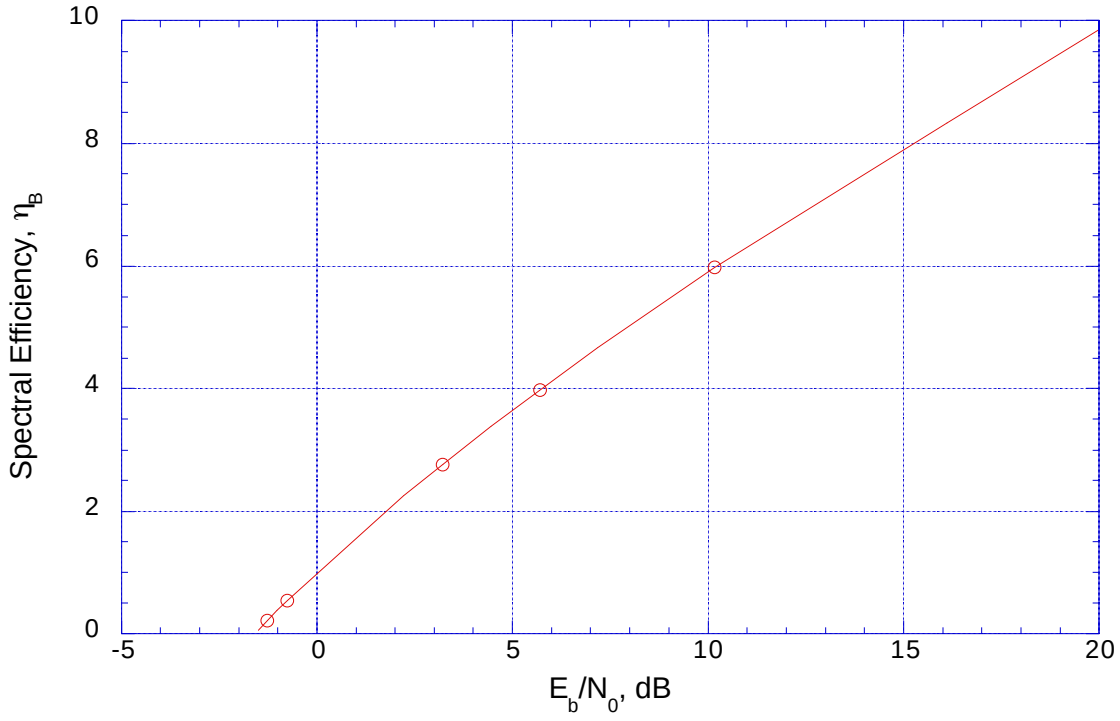


Figure 6.4: Maximum achievable spectral efficiency.

Chapter 7

Optimal Single Bit Demodulation Structures

7.1 Introduction

In this chapter we consider the transmission and demodulation of one bit of information transmitted on an all-pass channel. This chapter will demonstrate how statistical decision theory is useful in the design of digital communications. The treatment here is less rigorous than one might see in an engineering book on detection (e.g., [Poo88]). The treatment given here is my synthesis of ideas already presented by many authors before. A particular favorite of mine is [Web87].

Here a set of notation is formulated. The bit of information will be denoted I with $I = 0, 1$. The discussion will assume that the bit to be sent is a random variable with $P(I = 0) = \pi_0$ and $P(I = 1) = \pi_1$. The transmitted waveform is $X_z(t)$. If $I = 0$ then $X_z(t) = x_0(t)$ and if $I = 1$ then $X_z(t) = x_1(t)$. The support[1] of $x_0(t)$ and $x_1(t)$ is in the interval $[0, T_p]$. In other words the signals $x_0(t)$ and $x_1(t)$ are both energy waveforms no longer than T_p in length. Consequently the bit rate is defined to be $W_b = 1/T_p$. Fig. 7.1 shows an example of two waveforms for one bit transmission. These two waveforms are explicitely given as

$$x_0(t) = \begin{cases} \sin\left(\frac{4\pi t}{T_p}\right) & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \qquad x_1(t) = \begin{cases} 1 & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{7.1}$$

For the channel that was considered in the analog portion of the text the output received signal can be accurately modeled with

$$Y_z(t) = L_p X_z(t - \tau_p) \exp[j\phi_p] + W_z(t) \tag{7.2}$$

where $W_z(t)$ is a **complex** additive white Gaussian noise (AWGN) with $S_{W_z}(f) = N_0$ that models a thermal noise at the receiver frontend[2], L_p is the propagation loss, τ_p is the propagation delay, and ϕ_p is a propagation induced phase shift. The majority of this book will be focused on **coherent** digital communications. Coherent communications implies that the receiver knows entirely the distortion that can occur in transmission (with the model in (7.2) the exact value of L_p, τ_p, and ϕ_p). For coherent communication, the model in (7.2) can complicate the ideas of digital modulation and demodulation design so we will adopt a simpler one, i.e.,

$$Y_z(t) = X_z(t) + W_z(t). \tag{7.3}$$

[1] The support of a function is the domain of the function where the range is nonzero valued.

[2] See Section 5.5 for a motivation of this noise model.

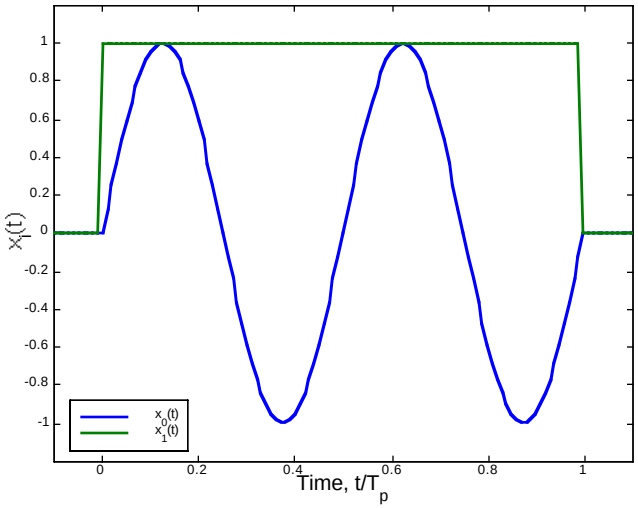


Figure 7.1: An example of two possible waveforms used to transmit one bit of information.

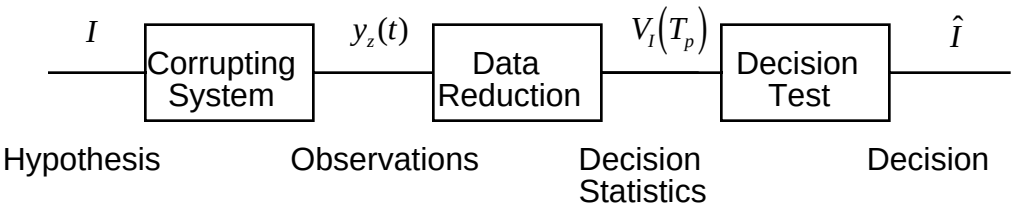


Figure 7.2: The problem formulation for hypothesis testing.

The model in (7.3) assumes the transmitted signal is received undistorted. There is little loss in generality in considering this model for coherent communications since each of the parameters L_p, τ_p, and ϕ_p in (7.2) has a simple separable change for the optimal demodulator. A homework problem explores this difference. The energy per bit in this simple binary transmission example is

$$E_b = \pi_0 E_{x_0} + \pi_1 E_{x_1}. \tag{7.4}$$

7.1.1 Statistical Hypothesis Testing

Digital communications borrows a vast majority of it's theory from the well developed theory of statistical hypothesis testing. A hypothesis test come about when a person is faced with the problem of making a definite decision with respect to an uncertain hypothesis which is known only through its observable consequences. A statistical hypothesis test is an algorithm based on a set of observations to decide on the alternative (for or against the hypothesis) which minimizes certain risks [Leh86].

The general statistical hypothesis testing problem formulation is given in Fig. 7.2. In general there is some system that disrupts the ability to make decisions. The output of that system is the raw observations. These raw observations are processes in some way to make the decision. The decision making process is usually formulated by getting either an analytical or an empirical understanding of how likely each hypothesis is and how the observations relate to the hypothesis.

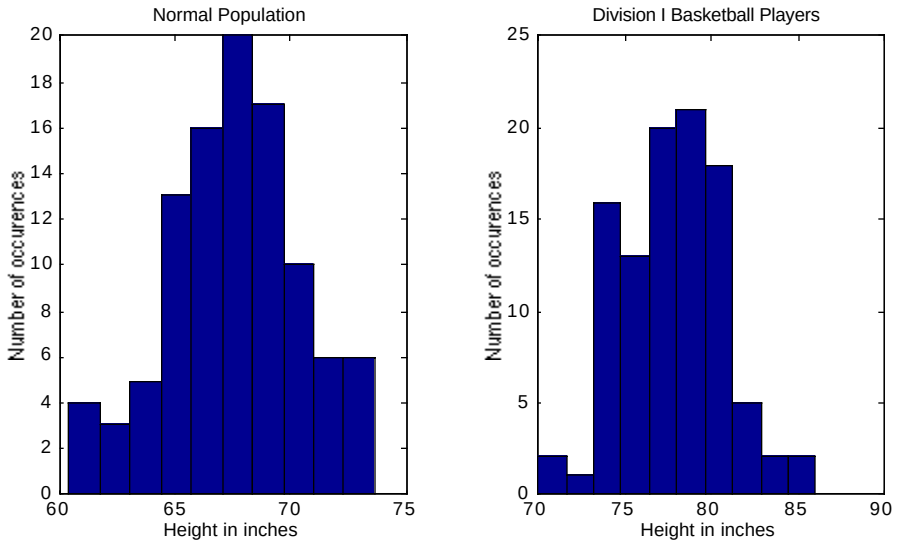


Figure 7.3: Histograms of the height of nonbasketball players and basketball players.

Example 7.1: Imagine you have just been hired at Westwood University as the men's basketball coach. Half the team, upset about the firing of the previous coach, quit the team and you have a big game against your archrival in three days. You are desperate to find Division I basketball players and have started searching around campus (hypothesis: Basketball player or no basketball player). Your assistant coach had the brilliant insight that basketball players tend to be taller than average and you decided to base your entire decision on the height of individuals (decision statistic: height). The assistant went out and measured 100 basketball players and 100 non-basketball players and gave you the histograms in Fig. 7.3. Given the problem formulation three questions can be asked

1. If a person is 5'4" tall (64 inches) could he be a basketball player? The answer is yes but it is highly unlikely so you would probably decide against asking a 5'4" tall person to join the team.

2. If a person is 7'0" tall (84 inches) could he not be a basketball player? Again the answer is yes and but it is highly unlikely so you would probably invite all 7'0" people to join the team.

3. If a person is 6'0" tall (72 inches) would you invite him to be on the team? There seems to be about an equal probability, judging from Fig. 7.3, that a nonbasketball playing person would be 6'0" tall as there is that a basketball player would be 6'0" tall. The best decision would likely be that 6'0" tall person not be invited since there are many more nonbasketball playing people than basketball players and hence it is more probable that a 6'0" tall person is not a basketball player.

This example shows intuitively how decisions are made, how statistics are formed and how the probability of each hypothesis should impact the decision. These elements are all in the digital communications problem.

It is clear from the previous example of trying to decide on basketball players that both what is observed in a statistical test and what are the prior distributions on the possible outcomes can both significantly impact the decision that is made. To capture these characteristics statisticians have defined two important quantities.

Definition 7.1 *The a priori probability is the probability associated with a possible hypothesis before*

any experiments are completed.

Example 7.2: A rough estimate that the probability a randomly chosen person on the Westwood University campus is a Division I basketball player is roughly 0.001.

Definition 7.2 *The a posteriori probability (APP) is the probability associated with a possible hypothesis which takes into account any observed experimental outcomes and the a priori probability.*

A priori probabilities and a posteriori probabilities have a prominent role in the theory of digital communications and in fact form the basis of most modern data modem technology.

7.1.2 Statistical Hypothesis Testing in Digital Communications

The optimum structure for demodulation of known transmitted signals in the presence of the white noise (i.e., (7.3)) is shown in Fig. 7.4. The structure consists of a linear filter where the real output of this linear filter is sampled and subject to a threshold test. The justification for why this structure is optimum will be left to a course on detection theory (e.g., [Poo88]). A more general and mathematically rigorous derivation is provided in Appendix A to complement the discussion in this chapter. Linear filters, samplers, and threshold tests are some of the most common electrical components available. The fact that the combination of a linear time-invariant filter, a sampler, and a threshold test forms the best single bit digital communications demodulator is quite striking. The sample time can be arbitrary but here is chosen to be at T_p, the end of the transmitted pulse. Hopefully after the entire optimum demodulation structure is examined in detail the reasons for this selection will become clear. For clarity, a threshold test is a component that provides the following logical operation: if

$$V_I(T_p) > \gamma \qquad \text{then} \qquad \hat{I} = 1 \tag{7.5}$$

else

$$V_I(T_p) \leq \gamma \qquad \text{then} \qquad \hat{I} = 0. \tag{7.6}$$

Threshold tests or comparators are easily implemented in electronic circuits.

Definition 7.3 *A statistic is any processing of data to produce a number that represents the data but is of lower dimensionality than the original data.*

This definition matches with how we use statistics in a wide variety of applications (e.g., mean and variance of grades). In the case considered here $V_I(T_p)$ is one number that represents the entire continuous time received waveform, $Y_z(t)$ and is a statistic for deciding about I.

Definition 7.4 *A sufficient statistic is statistic where making a decision or estimate based only on the statistic results in no loss in information/performance in estimation or detection compared to using the full data record.*

In the binary detection problem an optimal decision about I can be made by only considering $V_I(T_p)$, consequently $V_I(T_p)$ is a sufficient statistic for the entire observed waveform, $Y_z(t)$.

7.1.3 Digital Communications Design Problem

The digital communications design problem consists of identifying the components of the optimal demodulator such that a good tradeoff can be achieved between performance, complexity, and spectral efficiency. The approach taken in this set of notes to understanding this tradeoff for single bit demodulation is to work through five simple design tasks

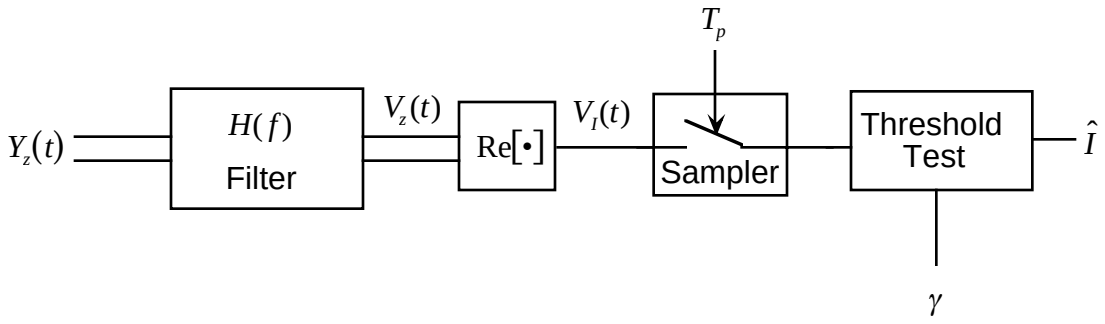


Figure 7.4: The optimum demodulator for a single bit transmission in AWGN.

1. Given $x_0(t), x_1(t)$ and $H(f)$, design the optimum threshold test (find γ).
2. Given $x_0(t)$ and $x_1(t)$ and $H(f)$ with the optimum threshold test, compute the performance of the demodulator. The performance metric of interest will be the probability of making a bit decision error.
3. Given $x_0(t)$ and $x_1(t)$, design the filter, $H(f)$ that minimizes the bit error probability (BEP).
4. Given the optimum demodulator, design $x_0(t)$ and $x_1(t)$ to optimize performance.
5. Design $x_0(t)$ and $x_1(t)$ to have desired spectral characteristics.

7.2 Minimum Probability of Error Bit Demodulation

The goal in this section is to address Design Task 1. We are looking to fix $x_0(t)$, $x_1(t)$, and $H(f)$ and find a threshold, γ, that gives the optimum performance. Our criterion for optimal demodulation will be based on which bit is more probable given an observed output from the receiver, $V_I(T_p) = v_I$. Recall this text maintains the convention that capital letters denote random variables and lower case letters will denote observed realizations of those random variables. This situation of trying to map a random variable into one of two possible decisions on a bit is a situation where this convention is necessary to clarify the actual underlying mathematics. $V_I(T_p)$ and $\hat{I}$ are random variables but the optimum mapping from the observations $V_I(T_p) = v_I$ to the decision $\hat{I}(v_I)$ is a deterministic function.

Example 7.3: Consider the signals shown in Fig. 7.1 with a truncated ideal lowpass filter with design bandwidth of $B_T = 4/T_p$, i.e.,

$$h(t) = \begin{cases} 2B_T \text{sinc}\left(2B_T(t-\tau_h)\right) & 0 \le t \le 2\tau_h \\ 0 & \text{elsewhere} \end{cases} \tag{7.7}$$

where $\tau_h = 3/T_p$. Note for simplicity of displaying the signals at various points of the demodulator we have chosen both the transmitted signals and the filter to be real valued. The signals and the noiseless output of the filter are plotted in the Fig. 7.5. The vertical line represents a sampling time when the waveforms are relatively easy to distinguish and this sample time will be used in this example as we examine Design Task 1.

The decision will be made based on which value of the transmitted bit is more probable (likely) given the observed value v_I. This type of demodulation is known as maximum *a posteriori* bit demodulation (MAPBD). A MAPBD computes the a posteriori probability (APP) for each possible hypothesis and

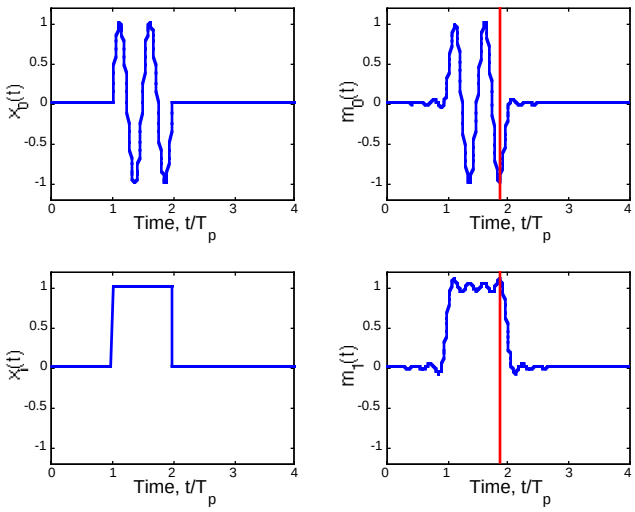


Figure 7.5: The transmitted signals and the noiseless output of $H(f)$ for Example 7.3.

then chooses the hypothesis whose APP is largest. The MAPBD chooses the most likely hypothesis after both the a priori information and the observations are analyzed. MAPBD can be shown to be the minimum probability of error demodulation scheme by using the theory of Bayes detection [Leh86][3]. Mathematically this can be stated as

$$P\left(I=1\,|v_I\right) \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} P\left(I=0\,|v_I\right). \tag{7.8}$$

This decoding rule first computes the *a posteriori* probability (APP) of each possible value of the transmitted bit, $P\left(I=i\,|v_I\right)\ i=0,1$, and then makes a decision based on which APP is largest. This procedure of computing an APP based on the observed channel outputs is a common theme in modern digital communications.

These probabilities can be computed with Bayes rule (mixed form)

$$\begin{aligned} P(I=1|v_I) &= \frac{f_{v_I|1}(v_I|I=1)P(I=1)}{f_{v_I}(v_I)} \\ P(I=0|v_I) &= \frac{f_{v_I|0}(v_I|I=0)P(I=0)}{f_{v_I}(v_I)} \end{aligned} \tag{7.9}$$

When discussing detection problems the notation $\pi_0 = P(I=0)$ and $\pi_1 = P(I=1)$ is often used [Poo88] and this text will adopt this notation as well. When the common term present in (7.9) is cancelled from both sides of (7.8) the decoding rule becomes

$$f_{V_I|1}(v_I|I=1)\pi_1 \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} f_{V_I|0}(v_I|I=0)\pi_0. \tag{7.10}$$

Fig. 7.6 shows the block diagram for this MAPBD threshold test. The forming of the threshold test is essentially the finding of which of two decision statistics are larger. This is a common theme that will

[3]Bayes detection theory is briefly overviewed in Appendix A

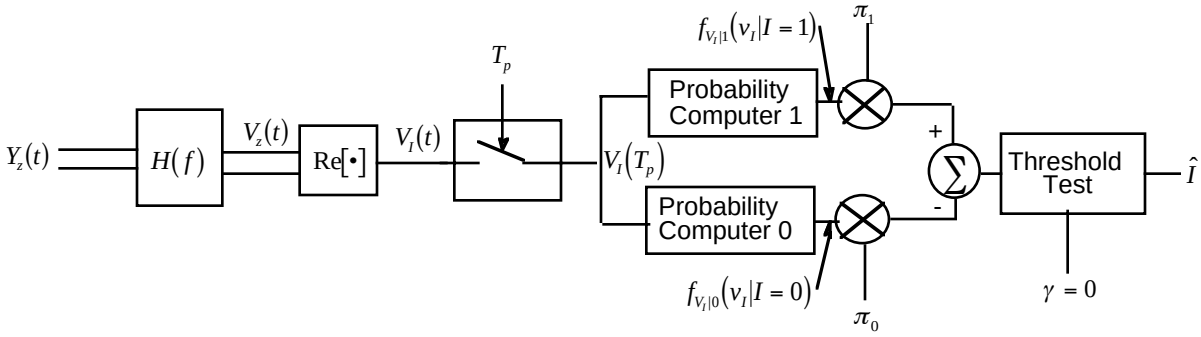


Figure 7.6: The maximum a posteriori bit demodulator.

weave itself through the remainder of the text. To prepare the reader for this thread, the decision test in (7.106) can be restated as

$$\hat{I} = \arg\max_{i=0,1} f_{V_I|i}(v_I|I=i)\pi_i. \tag{7.11}$$

The optimum test can often be simplified from the form given in (7.106) but the discussion throughout the course will often return to a demodulator format based on the form in (7.106). Another form of the test that is often used is known as the likelihood ratio test [Poo88] given as

$$\frac{f_{V_I|1}(v_I|I=1)}{f_{V_I|0}(v_I|I=0)} \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \frac{\pi_0}{\pi_1}. \tag{7.12}$$

From a practical point of view this maximum a posteriori bit demodulator is not interesting if the computation of $f_{V_I|i}(v_I|I=i),\ i=0,1$ is not simple. Fortunately the form of the receiver for the single bit demodulator can be significantly simplified.

7.2.1 Characterizing the Filter Output

The filter output is straightforward to characterize since conditioned on the transmitted signal, the input is a known signal plus white Gaussian noise. If $I=0$ and denoting $h(t)$ as the filter impulse response, then

$$\begin{aligned} V_{z,0}(t) &= \int_{-\infty}^{\infty} (x_0(\tau) + W_z(\tau))\, h(t-\tau)d\tau \\ &= \int_{-\infty}^{\infty} x_0(\tau)h(t-\tau)d\tau + \int_{-\infty}^{\infty} W_z(\tau)h(t-\tau)d\tau \\ &= m_0(t) + N_z(t). \end{aligned} \tag{7.13}$$

The term $m_0(t)$ represents the output of the filter when $x_0(t)$ is input. Recall $N_z(t)$ is a Gaussian random process that is characterized with

$$R_{N_z}(\tau) = N_0 V_h(\tau) = N_0 \int_{-\infty}^{\infty} h(t)h^*(t-\tau)dt \qquad \text{and} \qquad S_{N_z}(f) = N_0\,|H(f)|^2 \tag{7.14}$$

Likewise if $I = 1$ then

$$\begin{aligned} V_{z,1}(t) &= \int_{-\infty}^{\infty} (x_1(\tau) + W_z(\tau))h(t-\tau)d\tau \\ &= \int_{-\infty}^{\infty} x_1(\tau)h(t-\tau)d\tau + \int_{-\infty}^{\infty} W_z(\tau)h(t-\tau)d\tau \\ &= m_1(t) + N_z(t) \end{aligned} \tag{7.15}$$

where $m_1(t)$ represents the output of the filter when $x_1(t)$ is input.

Output time samples of this filter are easily characterized. If $I = 0$, the sample taken at time T_p is given as

$$V_{0,z}(T_p) = m_0(T_p) + N_z(T_p)$$

and the real part of this sample is given as

$$V_{0,I}(T_p) = m_{o,I}(T_p) + N_I(T_p).$$

$V_I(T_p)$ is a Gaussian random variable (RV) with

$$f_{V_I|0}\left(v_I|I=0\right) = \frac{1}{\sqrt{2\pi\sigma_{N_I}^2}} \exp\left[-\frac{\left(v_I - m_{0,I}(T_p)\right)^2}{2\sigma_{N_I}^2}\right] \tag{7.16}$$

where $\sigma_{N_I}^2 = \mathrm{var}\left(N_I(T_p)\right)$. Since $N_I(T_p)$ is the output of a linear filter it is easy to show using the results from Chapters 3 and 4 that $2\sigma_{N_I}^2 = \int_{-\infty}^{\infty} N_0 \left|H(f)\right|^2 df$. In a similiar fashion

$$f_{V_I|1}\left(v_I|I=1\right) = \frac{1}{\sqrt{2\pi\sigma_{N_I}^2}} \exp\left[-\frac{\left(v_I - m_{1,I}(T_p)\right)^2}{2\sigma_{N_I}^2}\right] \tag{7.17}$$

Consequently the MAPBD given in (7.106) requires the computation of two Gaussian probability density functions (PDF) have the same variance but different means. The means of these two PDFs are a function of the demodulation filter and the two possible transmitted signals representing the two possible values of the bit.

Example 7.4: Continuing with Example 7.3, Fig. 7.7 shows a sample path of $V_0(t)$ (top left) and a sample path of $V_1(t)$ (bottom left) with $E_b/N_0 = 10$dB. These sample paths clearly behave like a signal, $m_i(t)$, plus a noise. The same sample point as considered in Example 7.3 is again shown with the vertical line. Histograms of the sampled filter output for 4000 random trials of this experiment when each of the two hypotheses are true ($I = 0$ at the top right and $I = 1$ at the bottom right) are also shown Fig. 7.7. This histogram shows empirically that the Gaussian PDF is correct model for $f_{V_I|i}\left(v_I|I=i\right)$.

7.2.2 Uniform A Priori Probability

There are many ways to simplify the receiver structure from this point but the one that provides the most intuition and is most practical is to assume uniform *a priori* probability (i.e., $\pi_0 = \pi_1 = 0.5$). Uniform *a priori* probabilities often arise in practice since this is the goal of source coding [CT92]. Uniform *a priori* probabilities also produce a threshold test which is not a function of the *a priori* probabilities. The form of the optimal test when the *a priori* probability is not uniform is explored

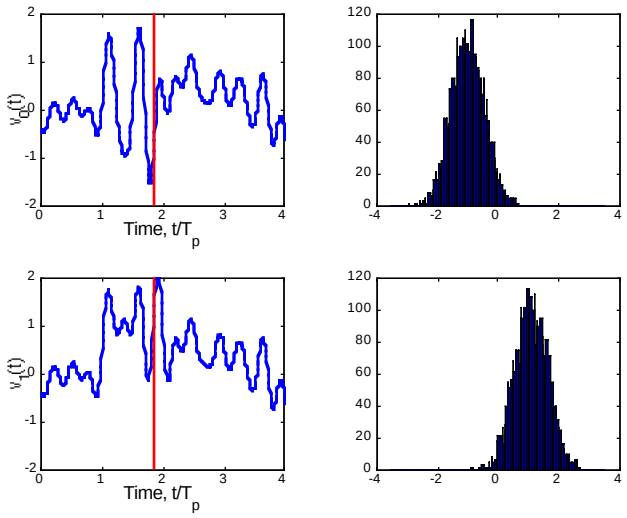


Figure 7.7: Sample paths of the filter outputs for each of the possible transmitted signals and histograms of the filter output samples.

in the homework. When $\pi_0 = 0.5$, the optimum decoder, often denoted the maximum likelihood bit demodulator (MLBD), is

$$\frac{1}{\sqrt{2\pi\sigma_{N_I}^2}} \exp\left[-\frac{(v_I - m_{1,I}(T_p))^2}{2\sigma_{N_I}^2}\right] \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \frac{1}{\sqrt{2\pi\sigma_{N_I}^2}} \exp\left[-\frac{(v_I - m_{0,I}(T_p))^2}{2\sigma_{N_I}^2}\right]. \tag{7.18}$$

Consequently this MLBD computes two Gaussian PDFs (having the same variance and different means) and makes a decision based on which of the two PDFs takes a higher value. This is the maximum likelihood principle [Poo88]. The monotonic nature of the Gaussian PDF implies that this results in a simple threshold test on $v_I(T_p)$. Fig. 7.8 shows pictorially how this decoder operates.

The MLBD has a very interesting geometric interpretation. Cancelling common terms and taking log[4] gives

$$(v_I - m_{1,I}(T_p))^2 \underset{\hat{I}=1}{\overset{\hat{I}=0}{\gtrless}} (v_I - m_{0,I}(T_p))^2 \tag{7.19}$$

Recall the definition of Euclidean distance in an N dimensional vector space.

Definition 7.5 *Euclidean distance in an N dimensional vector space between two points $\vec{x}$ and $\vec{y}$ is*

$$\delta_E(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^{N} (x_i - y_i)^2} \tag{7.20}$$

For simplicity of notation the squared Euclidean distance is also defined.

[4]Taking the log does not change the decision rule as $\log(x)$ is a monotonic function.

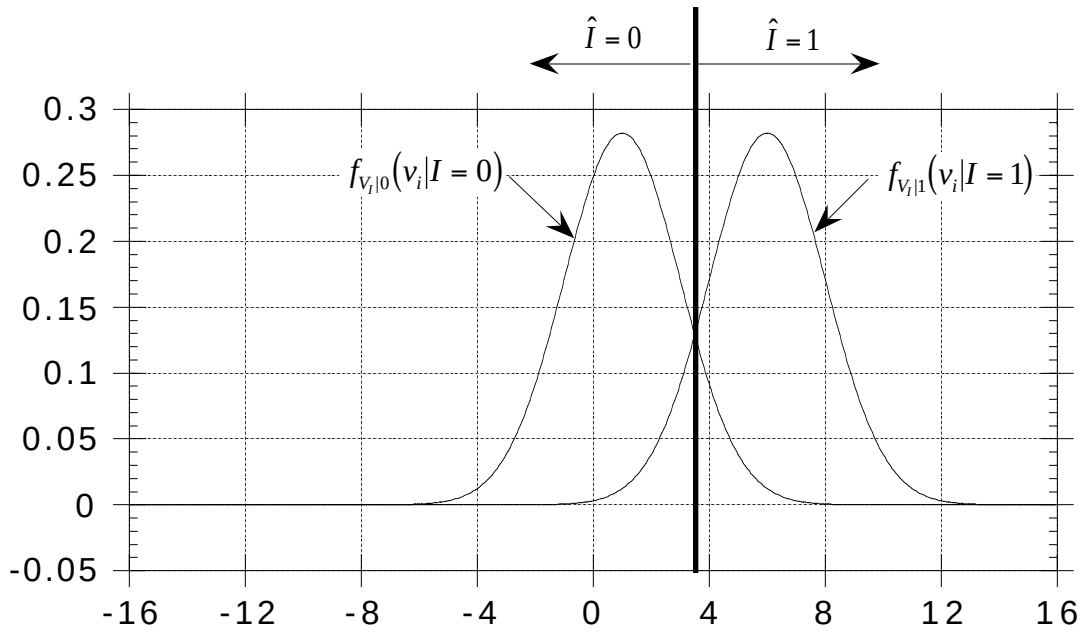


Figure 7.8: An example of the resulting test given by (7.18). $m_{0,I}(T_p) = 1$, $m_{1,I}(T_p) = 6$, and $\sigma^2_{N_I} = 2$.

Definition 7.6 *The squared Euclidean distance in an N dimensional vector space between two points $\vec{x}$ and $\vec{y}$ is*

$$\Delta_E(\vec{x}, \vec{y}) = \sum_{i=1}^{N} (x_i - y_i)^2 \tag{7.21}$$

Consequently the MLBD can be rewritten as

$$\Delta_E(v_I, m_{1,I}(T_p)) \underset{\hat{I}=1}{\overset{\hat{I}=0}{\gtrless}} \Delta_E(v_I, m_{0,I}(T_p)) \tag{7.22}$$

This implies that the maximum likelihood bit decision is based on whether v_I is geometrically closer to $m_{1,I}(T_p)$ or $m_{0,I}(T_p)$. This geometric interpretation of decision rules will arise several more times during in the context of demodulation for digital communications.

The decoder can further be simplified to a very simple threshold test. Assuming $m_{1,I}(T_p) > m_{0,I}(T_p)$, completing square, and doing some algebra reduces the optimum decision rule to

$$v_I \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \gamma = \frac{m_{1,I}(T_p) + m_{0,I}(T_p)}{2} \tag{7.23}$$

The MLBD compares the observed filter output, v_I with a threshold which is the arithmetic average of the two filter outputs corresponding to the possible transmitted signals in the absence of noise. For the example considered in Fig 7.8 where $m_{0,I}(T_p) = 1$, $m_{1,I}(T_p) = 6$, and $\sigma^2_{N_I} = 2$ the optimum threshold is $\gamma = 3.5$. It is interesting to note that the noise variance does not enter into the MLBD structure. It

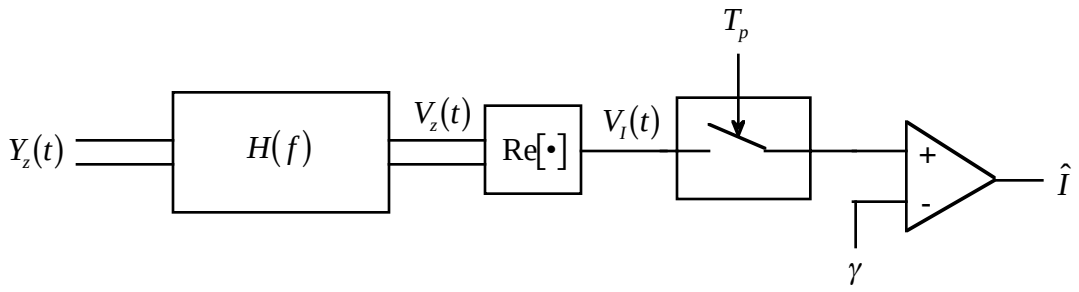


Figure 7.9: The simple form of the MLBD.

will be left as a student exercise to compute the threshold test for the case $m_{1,I}(T_p) < m_{0,I}(T_p)$. It is equally simple. The block diagram for the simplest form for the MLBD is shown in Fig. 7.9.

Design Task 1 is complete; we have found the optimum threshold test given $x_0(t)$, $x_1(t)$ and $H(f)$ for the uniform prior distribution case. Results for a nonuniform prior distribution are similiar.

7.3 Performance Analysis

Recall Design Task 2 is given $x_0(t)$ and $x_1(t)$ and $H(f)$ with the MLBD threshold test, γ, compute the performance of the demodulator. The performance metric of interest will be the probability of making a bit decision error. In digital communications the simplest performance metric is the bit error probability (BEP).

Definition 7.7 *Bit error probability (BEP) is*

$$P_B(E) = P\left(\hat{I} \neq I\right).$$

The BEP can be computed using total probability as

$$P_B(E) = P\left(\hat{I} = 1 \,|I = 0\right)\pi_0 + P\left(\hat{I} = 0 \,|I = 1\right)\pi_1 \tag{7.24}$$

Recall that when $m_{1,I}(T_p) > m_{0,I}(T_p)$, the MAPBD has the form

$$v_I \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \gamma \tag{7.25}$$

so that

$$P\left(\hat{I} = 1 \,|I = 0\right) = P\left(V_I(T_p) > \gamma \,|I = 0\right) \tag{7.26}$$

and

$$P\left(\hat{I} = 0 \,|I = 1\right) = P\left(V_I(T_p) < \gamma \,|I = 1\right). \tag{7.27}$$

Since conditioned on I, $V_I(T_p)$ is a Gaussian random variable with known mean and variance, the probabilities in (7.26) and (7.27) are simple to compute. These probabilities in (7.26) and (7.27) are simply the area under the tails of two Gaussian PDFs. Examples of the tails are illustrated in Figure 7.10.

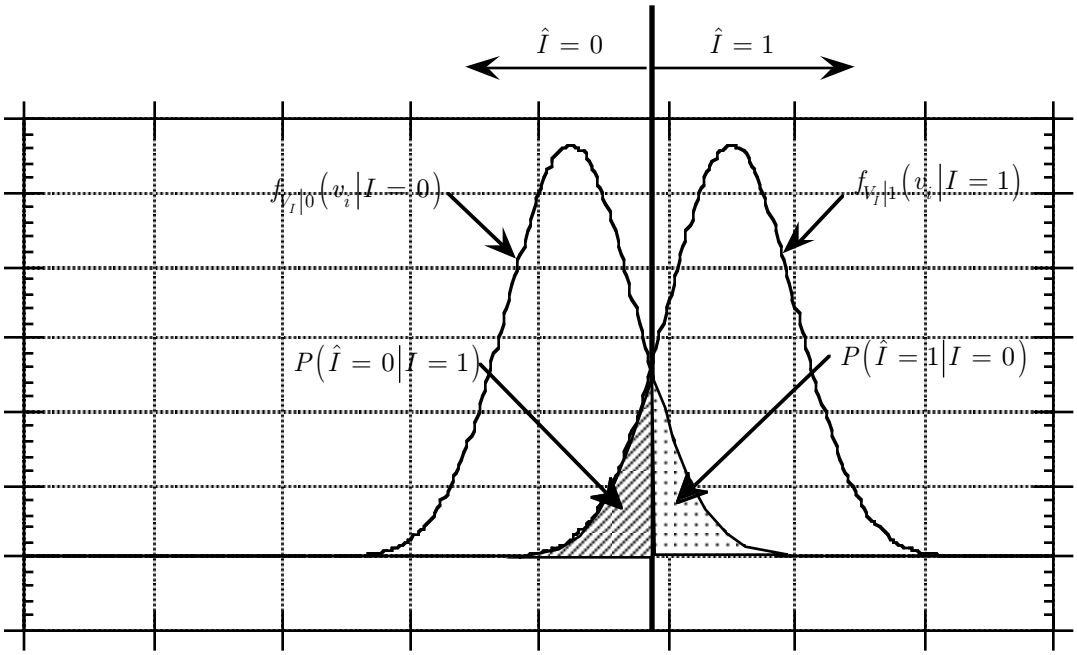


Figure 7.10: The tails of the Gaussian distribution that are the conditional bit error probability. This is for a test given by (7.18) with $m_{0,I}(T_p) = 1$, $m_{1,I}(T_p) = 6$, and $\sigma^2_{N_I} = 2$.

7.3.1 Erf Function

This text will compute the probability that a Gaussian RV lies in an interval using the erf() and erfc() functions [Ae72]. While this probability is expressed using different functions by various authors, this text uses the erf() and erfc() because it is commonly available in math software packages (e.g., Matlab).

Definition 7.8 *The erf function is*

$$erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt.$$

The cumulative distribution function of a Gaussian RV, X, with mean m_X and variance σ_X^2 is then given as

$$F_X(x) = \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{x - m_X}{\sqrt{2}\sigma_X}\right).$$

While the CDF of the Gaussian random variables is defined using different functions by various authors, this function is used in these notes because it is commonly used in math software packages (e.g., Matlab). Three properties of the erf function important for finding probabilities associated with Gaussian random variables are given as

$$\begin{aligned} \text{erf}(\infty) &= 1 \\ \text{erfc}(z) &= 1 - \text{erf}(z) \\ \text{erf}(-z) &= -\text{erf}(z). \end{aligned}$$

Some approximations to the $\text{erfc}(x)$ are available and they provide insight into the probability of error in digital communications. Two upperbounds are

$$\text{erfc}(x) \leq \exp\left[-x^2\right] = a_1(x) \tag{7.28}$$

$$\text{erfc}(x) \leq \frac{1}{\sqrt{\pi}x}\exp\left[-x^2\right] = a_2(x). \tag{7.29}$$

These approximations and $\text{erfc}(x)$ are plotted in Fig. 7.11.

7.3.2 Uniform A Priori Probability

Considering the case of the uniform probabilities with the MLBD demonstrates the simplicity of this calculation. Examining (7.26) and using (7.16) gives

$$\begin{aligned} P\left(\hat{I} = 1 \,|I = 0\right) &= \int_\gamma^\infty f_{V_I|0}\left(v_I|I = 0\right) dv_I \\ &= \int_\gamma^\infty \frac{1}{\sqrt{2\pi\sigma_{N_I}^2}} \exp\left[-\frac{\left(v_I - m_{0,I}(T_p)\right)^2}{2\sigma_{N_I}^2}\right] dv_I \end{aligned} \tag{7.30}$$

This probability can be simplified using the following change of variables

$$t = \frac{v_I - m_{0,I}(T_p)}{\sqrt{2}\sigma_{N_I}} \qquad dt = \frac{dv_I}{\sqrt{2}\sigma_{N_I}} \tag{7.31}$$

to give

$$P\left(\hat{I} = 1 \,|I = 0\right) = \frac{1}{2}\text{erfc}\left(\frac{\gamma - m_{0,I}(T_p)}{\sqrt{2}\sigma_{N_I}}\right) = \frac{1}{2}\text{erfc}\left(\frac{m_{1,I}(T_p) - m_{0,I}(T_p)}{2\sqrt{2}\sigma_{N_I}}\right) \tag{7.32}$$

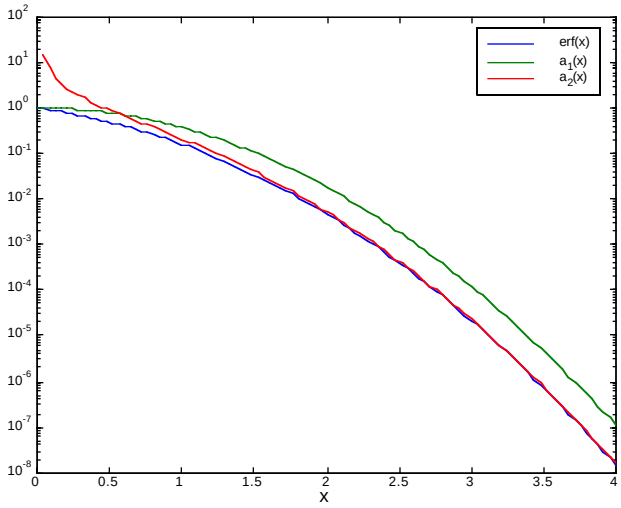


Figure 7.11: The erfc() and two approximations.

where the value for γ given in (7.23) is used to get the last expression. Similiar manipulations give

$$P\left(\hat{I}=0\,|I=1\right)=\frac{1}{2}\text{erfc}\left(\frac{m_{1,I}(T_p)-m_{0,I}(T_p)}{2\sqrt{2}\sigma_{N_I}}\right) \tag{7.33}$$

and consequently the BEP is also

$$P_B(E)=\frac{1}{2}\text{erfc}\left(\frac{m_{1,I}(T_p)-m_{0,I}(T_p)}{2\sqrt{2}\sigma_{N_I}}\right) \tag{7.34}$$

Design Task 2 is complete; we have found the performance of the MLBD given $x_0(t)$, $x_1(t)$ and $H(f)$ for the uniform prior distribution case. Results for a nonuniform prior distribution are similiar.

Some observations about the MLBD performance are in order here.

Definition 7.9 *The effective signal–to–noise (SNR) ratio for MLBD is*

$$\eta=\left(\frac{m_{1,I}(T_p)-m_{0,I}(T_p)}{2\sqrt{2}\sigma_{N_I}}\right)^2$$

Note the following

1. $P_B(E)=\frac{1}{2}\text{erfc}\left(\sqrt{\eta}\right)$,
2. $P_B(E)$ is monotone decreasing in η, in fact (7.28) indicates $P_B(E)$ behaves like $\frac{1}{2}\exp\left[-\eta\right]$ for moderately large η
3. Maximizing η will minimize $P_B(E)$,
4. For a fixed $x_0(t)$ and $x_1(t)$ the only quantity that effects the value of η which is under control of the communication system designer is the filter response ($h(t)$ or $H(f)$).

7.4 Filter Design

Considering the performance characteristic highlighted in the previous section leads to Design Task 3. This task is stated; given $x_0(t)$ and $x_1(t)$ design the filter, $H(f)$, that minimizes the BEP or equivalently maximizes η.

7.4.1 Maximizing Effective SNR

To solve this problem η must be expressed in terms of $H(f)$. To this end note

$$2\sigma_{N_I}^2 = N_0 \int_{-\infty}^{\infty} |H(f)|^2 \, df \tag{7.35}$$

and

$$\begin{aligned} m_1(T_p) - m_0(T_p) &= \int_{-\infty}^{\infty} (x_1(\lambda) - x_0(\lambda))\, h(T_p - \lambda) d\lambda \\ &= \mathcal{F}^{-1}\left\{ (X_1(f) - X_0(f))\, H(f) \right\}\big|_{t=T_p} \\ &= \int_{-\infty}^{\infty} (X_1(f) - X_0(f))\, H(f) \exp\left[j2\pi f T_p \right] df. \end{aligned} \tag{7.36}$$

To get a more compact notation we define

$$b_{10}(t) = x_1(t) - x_0(t) \qquad B_{10}(f) = X_1(f) - X_0(f).$$

Consequently

$$\eta = \left(\frac{m_{1,I}(T_p) - m_{0,I}(T_p)}{2\sqrt{2}\sigma_{N_I}} \right)^2 = \frac{\left(\Re\left\{ \int_{-\infty}^{\infty} B_{10}(f) H(f) \exp\left[j2\pi f T_p \right] df \right\} \right)^2}{4N_0 \int_{-\infty}^{\infty} |H(f)|^2 \, df}. \tag{7.37}$$

The form of $H(f)$ which maximizes the effective SNR is not readily apparent in examining (7.37) but interesting results can be obtain by using Schwarz's Inequality.

Theorem 7.1 (Schwarz's Inequality) *For two functions $X(f)$ and $Y(f)$ where $\int_{-\infty}^{\infty} |X(f)|^2 \, df < \infty$ and $\int_{-\infty}^{\infty} |Y(f)|^2 \, df < \infty$.*

$$\left| \int_{-\infty}^{\infty} X(f) Y^*(f) df \right|^2 \leq \int_{-\infty}^{\infty} |X(f)|^2 \, df \int_{-\infty}^{\infty} |Y(f)|^2 \, df$$

where equality holds only if $X(f) = AY(f)$ where A is a complex constant.

Making the assignment

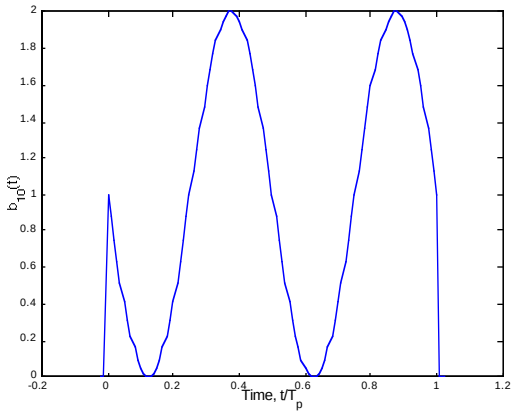
$$X(f) = H(f) \qquad Y(f) = B_{10}^*(f) \exp[-j2\pi f T_p] \tag{7.38}$$

and noting

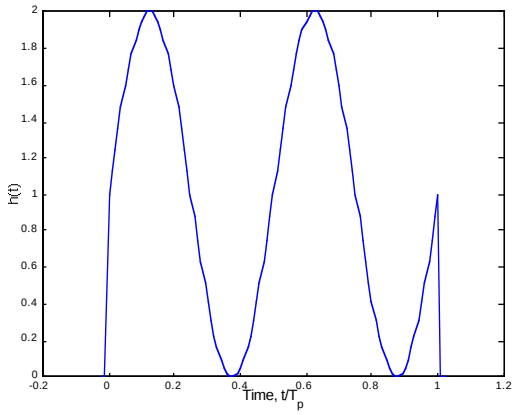
$$\left(\Re(X) \right)^2 \leq |X|^2 = \left(\Re(X) \right)^2 + \left(\Im(X) \right)^2 \tag{7.39}$$

Schwarz's inequality results in the following inequality for the effective SNR

$$\eta = \frac{\left(\Re\left\{ \int_{-\infty}^{\infty} B_{10}(f) H(f) \exp\left[j2\pi f T_p \right] df \right\} \right)^2}{4N_0 \int_{-\infty}^{\infty} |H(f)|^2 \, df} \leq \frac{1}{4N_0} \int_{-\infty}^{\infty} |B_{10}(f)|^2 \, df. \tag{7.40}$$



a) The effective signal, $b_{10}(t)$



b) The matched filter impulse response, $h(t) = b_{10}^*(T_p - t)$

Figure 7.12: The effective signal, $b_{10}(t)$ and the matched filter impulse response, $b_{10}^*(T_p - t)$ for the two example waveforms considered in Fig. 7.1.

η can be maximized by selecting

$$H(f) = CB_{10}^*(f)\exp[-j2\pi fT_p] \tag{7.41}$$

where C is a real constant. It should be noted that the constant now needs to be real so that

$$\Im\left\{\int_{-\infty}^{\infty} B_{10}(f)H(f)\exp\left[j2\pi fT_p\right]df\right\} = 0 \tag{7.42}$$

such that

$$\left(\Re\left\{\int_{-\infty}^{\infty} B_{10}(f)H(f)\exp\left[j2\pi fT_p\right]df\right\}\right)^2 = \left|\left\{\int_{-\infty}^{\infty} B_{10}(f)H(f)\exp\left[j2\pi fT_p\right]df\right\}\right|^2. \tag{7.43}$$

Schwarz's inequality has provided two powerful results: the optimum filter (7.41) and the maximum effective signal to noise ratio (7.40). From the perspective of performance the constant C makes no difference so without loss of generality the remainder of the discussion will assume $C = 1$.

7.4.2 The Matched Filter

Since the optimum filter has the form in the frequency domain given in (7.41) the impulse response is given as

$$h(t) = \mathcal{F}^{-1}\left\{B_{10}^*(f)\exp[-j2\pi fT_p]\right\} = b_{10}^*(T_p - t) = x_1^*(T_p - t) - x_0^*(T_p - t). \tag{7.44}$$

For MLBD the effective signal is $b_{10}(t) = x_1(t) - x_0(t)$ and the optimum filter impulse response is effectively a time reversed, time shifted, conjugate of this effective signal. This filter is known as the matched filter since it is matched to the effective signal and the form for this SNR maximizing filter was first identified by North [Nor43]. An example with the waveforms considered in Fig. 7.1 is shown in Fig. 7.12. It should be noted that having the sample time at T_p makes the matched filter a causal filter. Design Task 3 is complete; given $x_0(t)$ and $x_1(t)$, the filter, $H(f)$ that minimizes the bit error probability (BEP) has been found.

Example 7.5: Consider the signals shown in Fig. 7.1 with a matched filter. The signals and the noiseless output of the matched filter are plotted in the Fig. 7.13. The vertical line represents $t = T_p$.

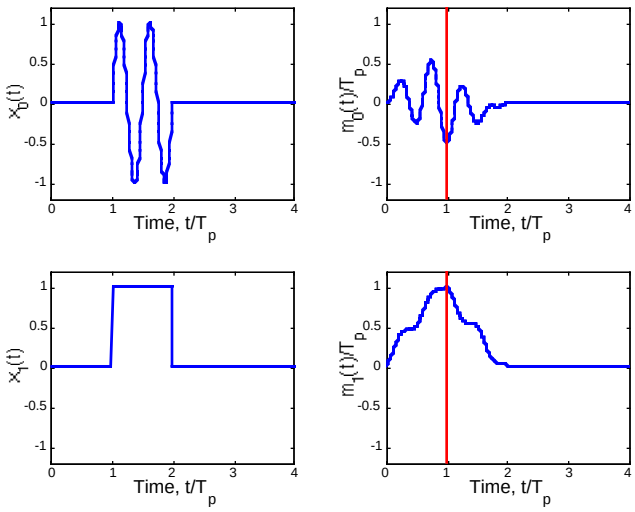


Figure 7.13: The transmitted signals and the noiseless output of matched filter for Example 7.5.

7.4.3 MLBD with the Matched Filter

More insight into digital communication receiver design can be achieved by looking in detail at the resulting demodulation structure when using a matched filter demodulator. First examine the matched filter output,

$$V_z(T_p) = \int_{-\infty}^{\infty} Y_z(\tau)h(T_p-\tau)d\tau = \int_{-\infty}^{\infty} Y_z(\tau)\left(x_1^*(\tau)-x_0^*(\tau)\right)d\tau. \tag{7.45}$$

Examining (7.45) shows that the optimum filter correlates the input signal plus noise, $Y_z(t)$, with the conjugate of the difference between the possible transmitted signals. Consequently the optimum filter is often denoted and implemented as a correlation operation.

The output noise, $N_z(t)$, is the same regardless of which bit is transmitted so it will be characterized first. Recall the output noise is given as

$$N_z(T_p) = \int_{-\infty}^{\infty} W_z(\tau)h(T_p-\tau)d\tau = \int_{-\infty}^{\infty} W_z(\tau)\left(x_1^*(\tau)-x_0^*(\tau)\right)d\tau. \tag{7.46}$$

Since $W_z(t)$ is a complex AWGN, the output noise, $N_z(T_p)$ will be a complex Gaussian RV. Using standard theory the ouput variance of this RV is given as

$$\begin{aligned} 2\sigma_{N_I}^2 &= N_0\int_{-\infty}^{\infty}|h(\tau)|^2\,d\tau = N_0\int_{-\infty}^{\infty}|x_1^*(T_p-\tau)-x_0^*(T_p-\tau)|^2\,d\tau \\ &= N_0\left(\int_{-\infty}^{\infty}|x_1(\tau)|^2\,d\tau+\int_{-\infty}^{\infty}|x_0(\tau)|^2\,d\tau-2\Re\left\{\int_{-\infty}^{\infty}x_1(\tau)x_0^*(\tau)d\tau\right\}\right) \\ &= N_0\left(E_1+E_0-2\Re\left\{\int_{-\infty}^{\infty}x_1(\tau)x_0^*(\tau)d\tau\right\}\right). \end{aligned} \tag{7.47}$$

In a similiar fashion the conditional means of the matched filter outputs for the two possible trans-

mitted signals, $m_1(T_p)$ and $m_0(T_p)$, are easily derived. If $I = 1$ then

$$\begin{aligned} m_1(T_p) &= \int_{-\infty}^{\infty} x_1(\tau)h(T_p-\tau)d\tau = \int_{-\infty}^{\infty} x_1(\tau)\left(x_1^*(\tau)-x_0^*(\tau)\right)d\tau \\ &= E_1 - \int_{-\infty}^{\infty} x_1(\tau)x_0^*(\tau)d\tau. \end{aligned} \tag{7.48}$$

The conditional mean of the decision statistic when $I = 1$ is

$$m_{1,I}(T_p) = E_1 - \Re\left\{\int_{-\infty}^{\infty} x_1(\tau)x_0^*(\tau)d\tau\right\}. \tag{7.49}$$

Likewise If $I = 0$ then

$$\begin{aligned} m_0(T_p) &= \int_{-\infty}^{\infty} x_0(\tau)h(T_p-\tau)d\tau = \int_{-\infty}^{\infty} x_0(\tau)\left[x_1^*(\tau)-x_0^*(\tau)\right]d\tau \\ &= \int_{-\infty}^{\infty} x_0(\tau)x_1^*(\tau)d\tau - E_0 = \left(\int_{-\infty}^{\infty} x_1(\tau)x_0^*(\tau)d\tau\right)^* - E_0. \end{aligned} \tag{7.50}$$

The conditional mean of the decision statistic when $I = 0$ is

$$m_{0,I}(T_p) = \Re\left\{\int_{-\infty}^{\infty} x_1(\tau)x_0^*(\tau)d\tau\right\} - E_0. \tag{7.51}$$

The optimum decision threshold with matched filter processing is

$$\gamma = \frac{m_{1,I}(T_p)+m_{0,I}(T_p)}{2} = \frac{E_1-E_0}{2}. \tag{7.52}$$

Consequently when matched filter processing is utilized all important quantities in the receiver are a function of the energy of the two signals used to represent the bit values, E_1 and E_0, and $\int_{-\infty}^{\infty} x_1(\tau)x_0^*(\tau)d\tau$.

Since the quantity $\int_{-\infty}^{\infty} x_1(\tau)x_0^*(\tau)d\tau$ appears many times throughout discussions of digital communications systems a definition will be introduced to simplify the notation.

Definition 7.10 *The signal correlation coefficient between two deterministic signals, $x_i(t)$ and $x_j(t)$ is*

$$\rho_{ij} = \frac{\int_{-\infty}^{\infty} x_i(\tau)x_j^*(\tau)d\tau}{\sqrt{E_iE_j}}.$$

Using this definition gives

$$2\sigma_{N_I}^2 = N_0\left(E_1+E_0-2\sqrt{E_1E_0}\Re\left\{\rho_{10}\right\}\right) \tag{7.53}$$

$$m_{1,I}(T_p) = E_1 - \sqrt{E_1E_0}\Re\left\{\rho_{10}\right\} \tag{7.54}$$

$$m_{0,I}(T_p) = \sqrt{E_1E_0}\Re\left\{\rho_{10}\right\} - E_0. \tag{7.55}$$

Using these simplified forms for the important parameters in the MLBD with matched filter processing allows us to gain some insight into the signal design problem.

Example 7.6: Continuing with Example 7.5, Fig. 7.14 shows a sample path of $V_0(t)$ (top left) and a sample path of $V_1(t)$ (bottom left) for the matched filter with $E_b/N_0 = 10$dB. These sample paths clearly behave like a signal, $m_i(t)$, plus a noise. A histogram of 4000 samples of the filter output are also shown Fig. 7.14. This histogram shows imperically that the Gaussian PDF is correct model for $f_{V_I|i}(v_I|I=i)$. Since the decision threshold is halfway between the two means of the Gaussian PDF this figure also shows that the matched filter produces a better bit error performance than the non-optimal filter considered in Example 7.4. It should also be noted that $E_0 = 0.5T_p$, $E_0 = T_p$, and $\rho_{10} = 0$ and consequently $m_{1,I}(T_p) = E_1$ and $m_{0,I}(T_p) = -E_0$ and these results are reflected in the matched filter output signals shown in Fig. 7.14.

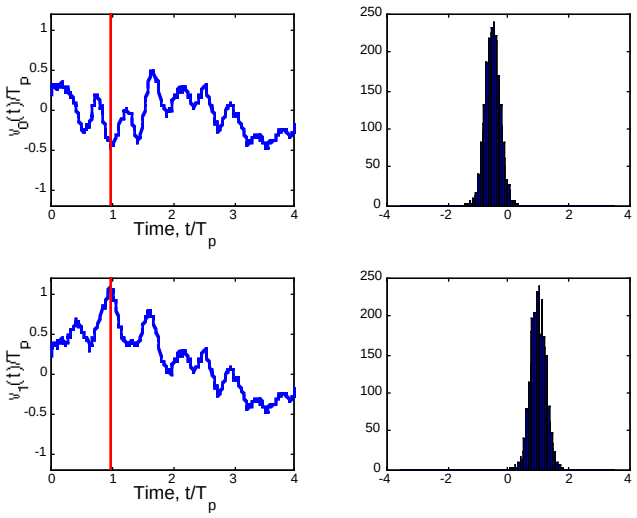


Figure 7.14: Sample paths of the matched filter outputs for each of the possible transmitted signals and histograms of the matched filter output samples.

7.4.4 More Insights on the Matched Filter

Two questions often arise with students after they read the previous sections. The two questions are

1. Why is the $\Re[\bullet]$ operator needed in the demodulator?
2. Could you do better by looking at more than one sample?

Understanding the answer to these questions and the optimal demodulator is obtained by looking at the characteristics of the signals at the matched filter outputs. First let's break the matched filter into two filters as shown in Fig. 7.15 and ignore the noise to gain some insight. The two filters are now matched to each of the two possible transmitted waveforms and the output signals when $x_j(t)$ is transmitted are

$$v_{0,j}(t) = \int_0^{T_p} x_j(\lambda) x_0^*(T_p - t + \lambda) d\lambda \tag{7.56}$$

$$v_{1,j}(t) = \int_0^{T_p} x_j(\lambda) x_1^*(T_p - t + \lambda) d\lambda. \tag{7.57}$$

Note that at $t = T_p$

$$v_{j,j}(T_p) = \int_0^{T_p} |x_j(\lambda)|^2 d\lambda \tag{7.58}$$

so that the signal received at each time is derotated into a positive real signal before the integration is performed. So in the filter matched to the signal sent it is clear the output sample will be $v_{jj}(T_p) = E_j + j0$, while in the filter not matched to the signal sent the output sample will be $v_{\bar{j}j}(T_p) = \rho_{j\bar{j}}\sqrt{E_j E_{\bar{j}}}$. Recall that $|\rho_{j\bar{j}}| \leq 1$ so $v_{jj}(T_p)$ can be viewed as the signal of interest out of the matched filter and $v_{\bar{j}j}(T_p)$ can be viewed as an interference. The signal of interest is a real valued signal since the matched filter perfectly derotates the signal to which it is matched. This perfect derotation characteristic of the matched filter is why only the real part of the output of the matched filter is sufficient for optimum demodulation.

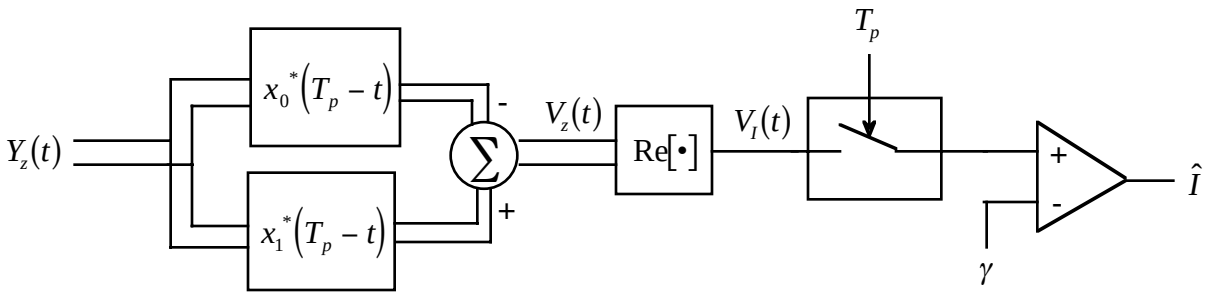


Figure 7.15: The two filter representation of the matched filter to the effective signal.

It is also clear in looking at (7.56) that the matched filter uses the entire transmitted signal to produce the decision statistic. Since the received signal is correlated with a version of the transmitted signal, the matched filter output is essentially a combination of the received signal over the entire support of the transmitted signal. This optimal filtering produces out a signal proportional to the energy of the transmitted signal and obviates the need to take more than one sample. Consequently using the matched filter demodulator makes the real part of the matched filter output sampled at $t = T_p$ a sufficient statistic for demodulation. The simplicity and optimality of the demodulator are striking.

7.5 Signal Design

Design Task 4, the signal design task, can now be addressed. This signal design task is stated as: given the optimum demodulator, design $x_0(t)$ and $x_1(t)$ to optimize performance. The results of the previous section indicate that performance of the binary detector is a function of E_1, E_0, and ρ_{10}. Performance of digital communication systems is most often quantified by the average energy per transmitted bit given as

$$E_b = \frac{E_1 + E_0}{2}.$$

Throughout the remainder of the course E_b will be used in characterizing performance.

To bring further insight to the discussion it is useful to introduce a definition

Definition 7.11 *The Euclidean square distance between two continuous time signals, $x_i(t)$ and $x_j(t)$ is*

$$\Delta_E(i,j) = \int_{-\infty}^{\infty} |x_i(t) - x_j(t)|^2 \, dt$$

It should be noted that

$$\begin{aligned} \Delta_E(i,j) &= \int_{-\infty}^{\infty} |x_i(t)|^2 \, dt + \int_{-\infty}^{\infty} |x_j(t)|^2 \, dt - \int_{-\infty}^{\infty} x_i(t) x_j^*(t) dt - \int_{-\infty}^{\infty} x_j(t) x_i^*(t) dt \qquad (7.59) \\ &= E_i + E_j - 2\sqrt{E_i E_j} \Re\{\rho_{ij}\}. \qquad (7.60) \end{aligned}$$

The performance and signal design task can now be formulated using the Euclidean square distance. The effective SNR with matched filter processing can be expressed with the Rayleigh energy theorem as

$$\eta = \frac{1}{4N_0} \int_{-\infty}^{\infty} |B_{10}(f)|^2 \, df = \frac{1}{4N_0} \int_{-\infty}^{\infty} |b_{10}(t)|^2 \, dt \qquad (7.61)$$

Using the concept of the Euclidean square distance this effective SNR reduces to

$$\eta = \frac{1}{4N_0}\int_{-\infty}^{\infty} |x_1^*(T_p - t) - x_0^*(T_p - t)|^2\, dt = \frac{\Delta_E(1,0)}{4N_0}. \tag{7.62}$$

Consequently the BEP for matched filtering processing is given as

$$P_B(E) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\Delta_E(1,0)}{4N_0}}\right) \tag{7.63}$$

The preceeding results demonstrate that the only parameter necessary to consider to optimize a binary signal set is the square Euclidean distance between the two possible transmitted signals. Equivalently, the square Euclidean distance is determined by the signal energies and the signal correlation coefficient.

Some examples will help illustrate the ideas of signal design. The "uneducated" masses typically think that digital communications occurs with the following waveforms

$$x_{1A}(t) = \begin{cases} \sqrt{\dfrac{2E_b}{T_p}} & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{7.64}$$

$$x_{0A}(t) = 0 \tag{7.65}$$

This particular signal set has $E_1 = 2E_b$, $E_0 = 0$ and $\rho_{10} = 0$. The resulting squared Euclidean distance is $\Delta_E(1,0) = 2E_b$ and the performance is given as

$$P_B(E) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right). \tag{7.66}$$

Performance can be improved by increasing the energy in $x_0(t)$ and simultaneously decreasing the ρ_{10}. For example the binary set of waveforms

$$x_{0B}(t) = \begin{cases} \sqrt{\dfrac{E_b}{T_p}} & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{7.67}$$

$$x_{1B}(t) = \begin{cases} -\sqrt{\dfrac{E_b}{T_p}} & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{7.68}$$

has $E_1 = E_b$, $E_0 = E_b$ and $\rho_{10} = -1$ and $\Delta_E(1,0) = 4E_b$. For the same average, energy signal set B provides the same performance as signal set A in an environment where the noise spectral density is twice as high. As communication engineers we are interested in the performance of the systems we design for varying noise levels hence signal set B is denoted as having a 3dB performance gain over signal set A. Fig. 7.16 shows the resulting square Euclidean distance versus signal energy for a fixed ρ_{10} and with $E_b = 1$. It is clear the signal set considered in (7.67) and (7.68), $E_1 = E_0 = E_b$ and $\rho_{10} = -1$ and $\Delta_E(1,0) = 4E_b$, provides the best performance. The situation where $x_0(t) = -x_1(t)$ is common in digital communication and it often referred to as **antipodal** signaling. Also, in general, higher energy and lower signal correlation will produce better performance.

Design Task 4 is complete; we have found the best signal designs for matched filter MLBD. The interesting characteristics of the signal design problem is that performance is only a function of the squared Euclidean distance and that the resulting optimum signal design is not unique. There are a infinite number of signal sets that achieve the same performance. Further optimization will have to be accomplished considering other parameters besides performance.

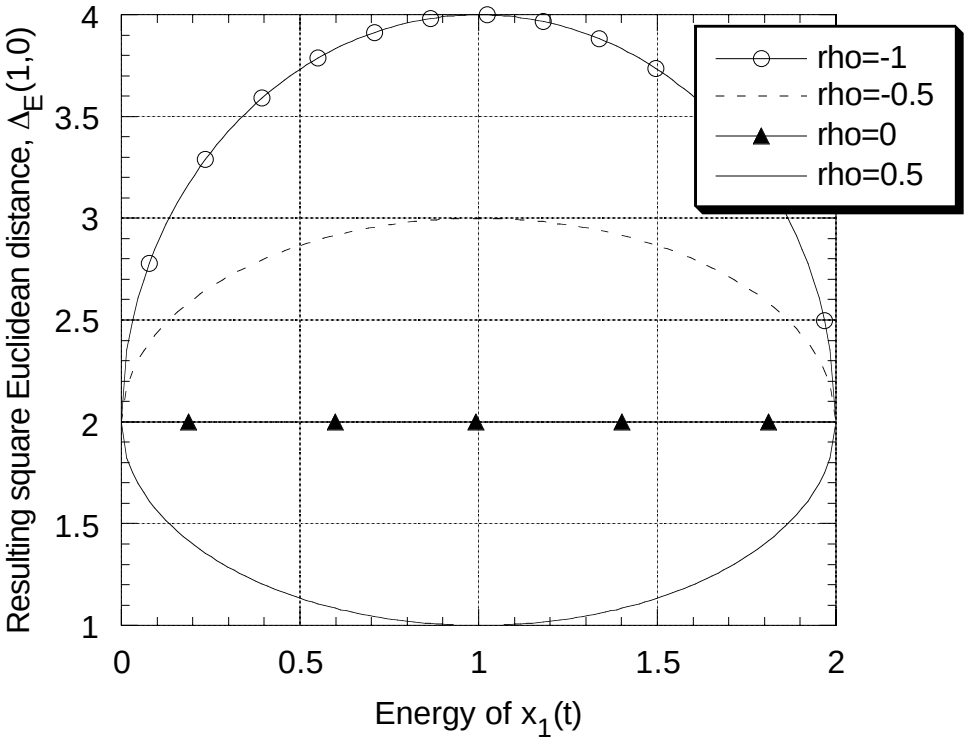


Figure 7.16: Square Euclidean distance versus E_1. E_b=1.

7.6 Spectral Characteristics

For single bit transmission the average energy spectral density is straightforward to compute. Given the framework used here and the definition of the average energy spectrum density given in (6.5) it is apparent that

$$D_{X_z}(f) = \pi_0 G_{x_0}(f) + \pi_1 G_{x_1}(f) \tag{7.69}$$

For example, for the signal set considered in the last section given as

$$x_0(t) = \begin{cases} \sqrt{\frac{E_b}{T_p}} & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{7.70}$$

$$x_1(t) = \begin{cases} -\sqrt{\frac{E_b}{T_p}} & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{7.71}$$

has

$$G_{x_1}(f) = E_b T_p \left(\frac{\sin(\pi f T_p)}{\pi f T_p} \right)^2 \qquad G_{x_0}(f) = E_b T_p \left(\frac{\sin(\pi f T_p)}{\pi f T_p} \right)^2 \tag{7.72}$$

The resulting energy spectrum for this case is plotted in Fig. 7.17 for $E_b T_p = 1$. It should be noted that the bandwidth in any definition is inversely proportional to the length of the pulse. Engineering intuition says that faster transmission rates (smaller T_p) requires more bandwidth. The average energy spectrum gives a way to quantify engineering intuition.

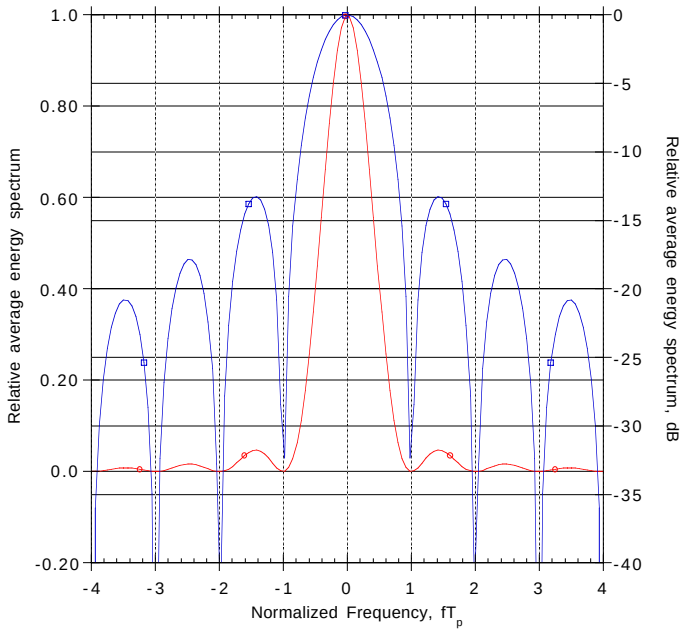


Figure 7.17: The average energy spectrum. $\pi_0 = \pi_1 = 0.5$.

7.7 Examples

The chapter is concluded by considering two obvious and classical examples of carrier modulated digital communication: binary frequency shift keying (BFSK) and binary phase shift keying (BPSK).

7.7.1 Frequency Shift Keying

BFSK modulation sends the bit of information by transmitting a carrier pulse of one of two frequencies. This is an obvious simple signalling scheme and one used in many early modems. The signal set is given as

$$x_0(t) = \begin{cases} \sqrt{\frac{E_b}{T_p}} \exp\left[j2\pi f_d t\right] & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{7.73}$$

$$x_1(t) = \begin{cases} \sqrt{\frac{E_b}{T_p}} \exp\left[-j2\pi f_d t\right] & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{7.74}$$

where f_d is known as the frequency deviation. The frequency difference between the two carrier pulses is $2f_d$. It is apparent that each waveform in a BFSK signal set has equal energy that has here been set to $E_1 = E_0 = E_b$. The vector diagram for the two waveforms for BFSK with $f_d = 0.25/T_p$ and $E_b = T_p$ is shown in Fig. 7.18. The plots of the bandpass waveforms for $f_d = 1/T_p$ and $f_c = 3/T_p$ and $E_b = T_p$ is shown in Fig. 7.19. This subsection will consider the optimum demodulator and the design of optimum signal sets for BFSK.

BFSK is interesting to investigate as the resulting characteristics are counter to many engineer's

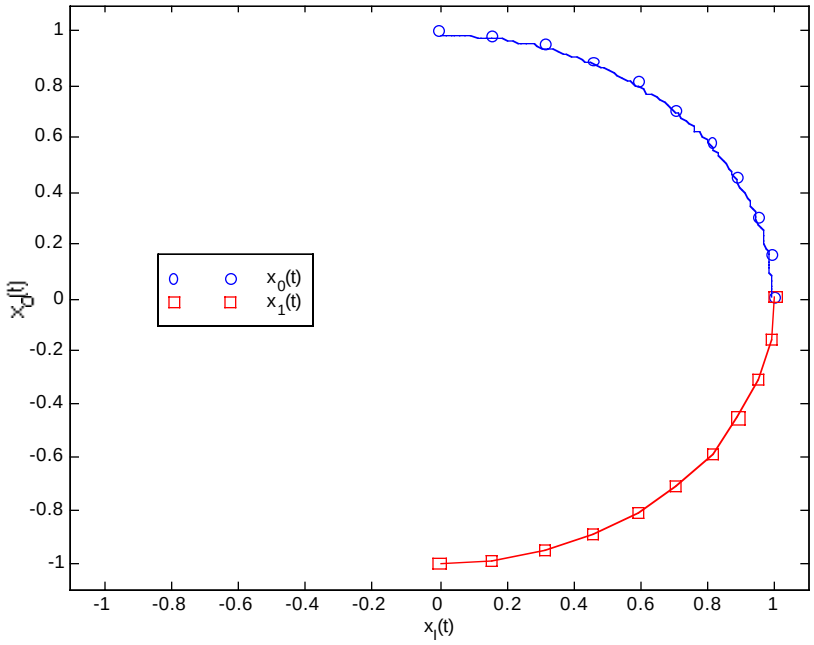


Figure 7.18: The vector diagram for the two BFSK waveforms. $f_d T_p = 0.25$ and $E_b = T_p$.

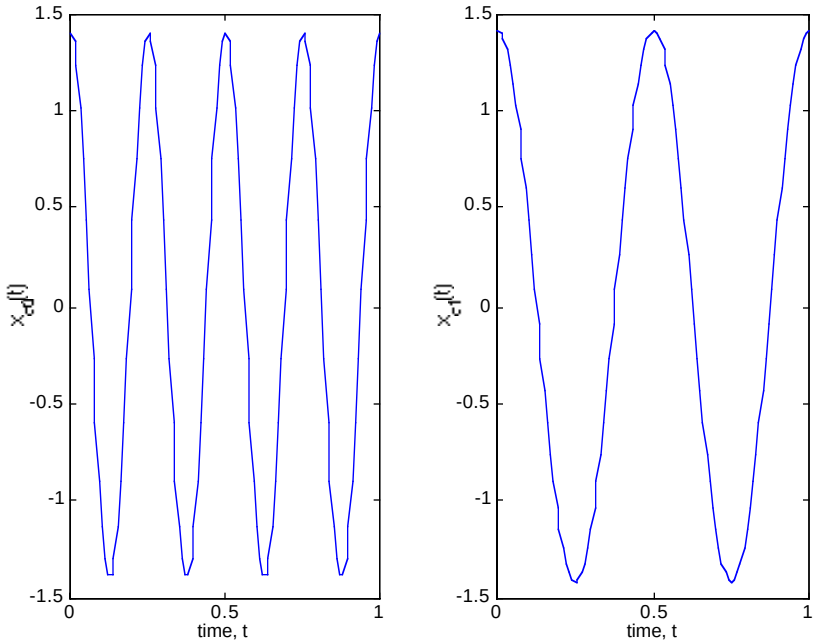


Figure 7.19: The bandpass BFSK waveforms. $f_d T_p = 1$, $f_c T_p = 3$, and $E_b = T_p$.

intuition. The matched filter impulse response is given as

$$\begin{aligned} h(t) &= x_1^*(T_p - t) - x_0^*(T_p - t) \\ &= j2\sqrt{\frac{E_b}{T_p}} \sin\left(2\pi f_d (T_p - t)\right) \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{elsewhere} \end{aligned} \tag{7.75}$$

Note since $\Re\{x_1(t)\} = \Re\{x_0(t)\}$ the optimum demodulator only uses the imaginary part of the received signal to differentiate between the two bits, i.e.,

$$v_I(T_p) = -2\sqrt{\frac{E_b}{T_p}} \int_0^{T_p} Y_Q(t) \sin\left(2\pi f_d t\right) dt. \tag{7.76}$$

For $\pi_0 = \pi_1 = 0.5$ (MLBD), the optimum threshold is given as

$$\gamma = \frac{E_1 - E_0}{2} = 0. \tag{7.77}$$

Consequently the threshold test is very simple for the MLBD of a BFSK modulation, i.e.,

$$\int_0^{T_p} Y_Q(t) \sin\left(2\pi f_d t\right) dt \underset{\hat{I}=1}{\overset{\hat{I}=0}{\gtrless}} 0. \tag{7.78}$$

The important parameter for determining the performance of the optimum demodulator is the square Euclidean distance. The only parameter that is not specfied in the BFSK model is the frequency difference between the two tones in the modulation. The parameter f_d determines this difference and also will determine the Euclidean distance. In general,

$$\Delta_E(1,0) = E_1 + E_0 - 2\Re\left\{\sqrt{E_1 E_0}\rho_{10}\right\}.$$

For BFSK $E_0 = E_1 = E_b$ and

$$\begin{aligned} \rho_{10} &= \frac{\int_{-\infty}^{\infty} x_1(t) x_0^*(t) dt}{E_b} = \frac{1}{T_p}\int_0^{T_p} \exp\left[-j4\pi f_d t\right] dt. \\ &= \frac{\sin\left(4\pi f_d T_p\right)}{4\pi f_d T_p} - j\frac{\cos\left(4\pi f_d T_p\right) - 1}{4\pi f_d T_p}. \end{aligned} \tag{7.79}$$

Consequently the Euclidean distance is given as

$$\Delta_E(1,0) = 2E_b\left(1 - \frac{\sin\left(4\pi f_d T_p\right)}{4\pi f_d T_p}\right) \tag{7.80}$$

and the BEP performance is

$$P_B(E) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\Delta_E(1,0)}{4N_0}}\right) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{2N_0}\left(1 - \frac{\sin\left(4\pi f_d T_p\right)}{4\pi f_d T_p}\right)}\right). \tag{7.81}$$

Fig. 7.20 has a plot of the resulting correlation and the Euclidean distance as a function of f_d.

The frequency offset significantly affects the performance. The best performance is obtained for $f_d \approx \frac{3}{8T_p}$ which produces $\Re\{\rho_{10}\} \approx -0.21$. In other words the performance of coherent BFSK is

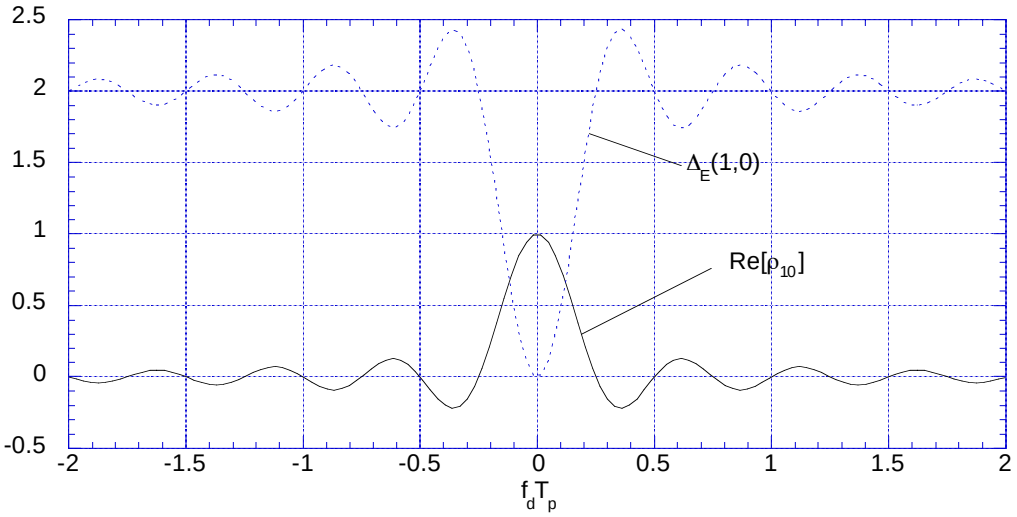


Figure 7.20: Correlation coefficient and Euclidean squared distance for BFSK.

optimized when the two frequencies used in the modulation are different by approximately $0.75/T_p$ Hz. This performance is 2.2dB from the optimal modulation $\Delta_E(1,0) = 4E_b$. Hence BFSK is not a modulation of choice when high performance is the driving constraint. Often engineering intuition leads one to believe that performance is optimized when the frequency difference is large. At large frequency offsets the signal set converges to being orthogonal ($\Re\{\rho_{10}\} = 0$) and this produces 0.8dB degradation in the performance compared to the optimal BFSK. Orthogonal BFSK is often used for a variety of practical reasons. Coherent demodulation for $\rho_{10} = 0$ produces a BEP of

$$P_B(E) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right). \tag{7.82}$$

The bit error probability is plotted versus E_b/N_0 in Fig. 7.21 for the case of optimal signaling (denoted BPSK[5]), optimal BFSK and orthogonal BFSK. The minimum f_d that still achieves an orthogonal signal set is $f_d = 0.25/T_p$. This form of BFSK modulation has some practical advantages when transmitting more than one bit of information and hence has been dubbed minimum shift keying (MSK) modulation. This modulation will be examined in more detail in the sequel.

The spectral efficiency of BFSK can be evaluated by looking at the average energy spectral density per bit. Recall the average energy spectral density per bit is given for binary modulations as

$$D_{X_z}(f) = \pi_0 G_{x_0}(f) + \pi_1 G_{x_1}(f) \tag{7.83}$$

The energy spectrum of the two individual waveforms is given as

$$G_{x_0}(f) = E_b T_p \left(\frac{\sin\left(\pi(f-f_d)T_p\right)}{\pi(f-f_d)T_p}\right)^2 \qquad G_{x_1}(f) = E_b T_p \left(\frac{\sin\left(\pi(f+f_d)T_p\right)}{\pi(f+f_d)T_p}\right)^2 \tag{7.84}$$

The resulting energy spectrum for BFSK is plotted in Fig. 7.22 for various f_d for $E_b T_p = 1$. The bandwidth of the BFSK signal is a function of both f_d and T_p. A smaller T_p produces a wider bandwidth while a larger f_d produces a wider bandwidth. It should be noted that with this form of BFSK the

[5]Why, will be apparent next section.

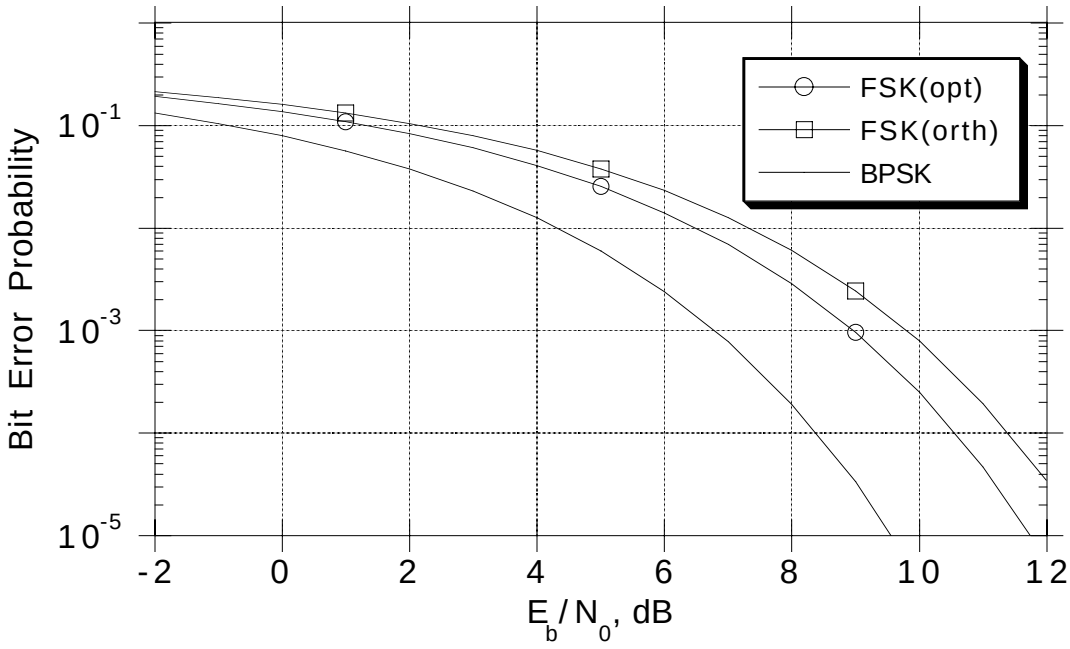


Figure 7.21: Probablity of error for some example binary modulations.

transmission rate is $W_b = 1/T_p$. In examining Fig. 7.22 the bandwidth of BFSK for the optimal performance is about $B_T = 1.5/T_p$ (i.e., 3dB bandwidth) so that the spectral efficiency of BPSK is about $\eta_B = 0.67$ bits/s/Hz.

The advantages of BFSK are summarized as

- Very simple to generate. Simply gate one of two oscillators on depending on the bit to be sent.
- Simple demodulation structure.

The disadvantages of BFSK are summarized as

- Performance is not optimum.
- Occupies a relatively wide bandwidth.

7.7.2 Phase Shift Keying

BPSK modulation sends the bit of information by transmitting a carrier pulse of one of two phases. The signal set is given as

$$\begin{aligned} x_0(t) &= \sqrt{\frac{E_b}{T_p}} \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{7.85}$$

$$\begin{aligned} x_1(t) &= \sqrt{\frac{E_b}{T_p}} \exp\left[j\theta\right] \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{7.86}$$

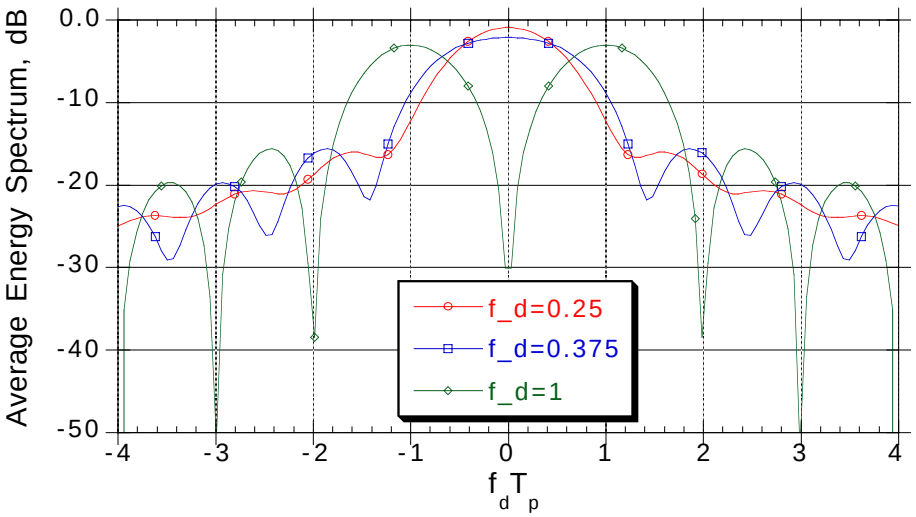


Figure 7.22: The average energy spectral density of BFSK for various f_d. $\pi_0 = \pi_1 = 0.5$.

where θ is known as the phase deviation. It is apparent that each waveform in a BPSK signal set has equal energy that has here been set to $E_1 = E_0 = E_b$. The vector diagram for the two waveforms for BPSK with $\theta = \pi$ and $E_b = T_p$ is shown in Fig. 7.23. The plots of the bandpass waveforms for $\theta = \pi$, $E_b = T_p$ and $f_c = 3/T_p$ is shown in Fig. 7.24.

This subsection will consider the optimum demodulator and the design of optimum signal sets for BPSK. The matched filter impulse response is given as

$$\begin{aligned} h(t) &= x_1^*(T_p - t) - x_0^*(T_p - t) \\ &= \sqrt{\frac{E_b}{T_p}}\,(\exp(-j\theta) - 1) \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{7.87}$$

The voltage for the threshold test is

$$\begin{aligned} v_I(T_p) &= \Re\left\{(\exp(-j\theta) - 1)\sqrt{\frac{E_b}{T_p}}\int_0^{T_p} Y_z(t)dt\right\} \\ &= (\cos(\theta) - 1)\sqrt{\frac{E_b}{T_p}}\int_0^{T_p} Y_I(t)dt + \sin(\theta)\sqrt{\frac{E_b}{T_p}}\int_0^{T_p} Y_Q(t)dt \end{aligned} \tag{7.88}$$

For $\pi_0 = \pi_1 = 0.5$ (MLBD), the optimum threshold is given as

$$\gamma = \frac{E_1 - E_0}{2} = 0. \tag{7.89}$$

Conseqeuntly the threshold test is very simple (compare to zero) for the MLBD of a BPSK modulation.

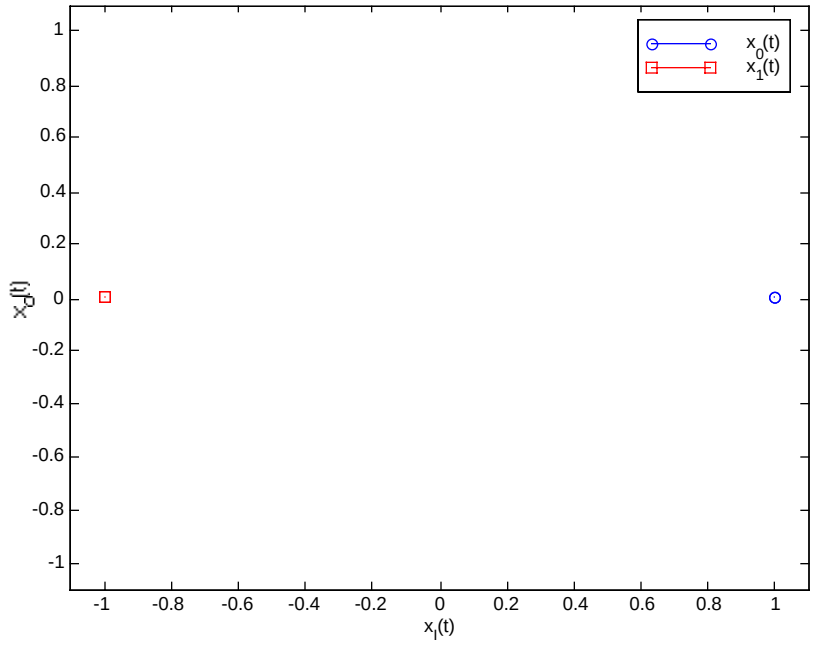


Figure 7.23: The vector diagram for the two BPSK waveforms. $\theta = \pi$ and $E_b = T_p$.

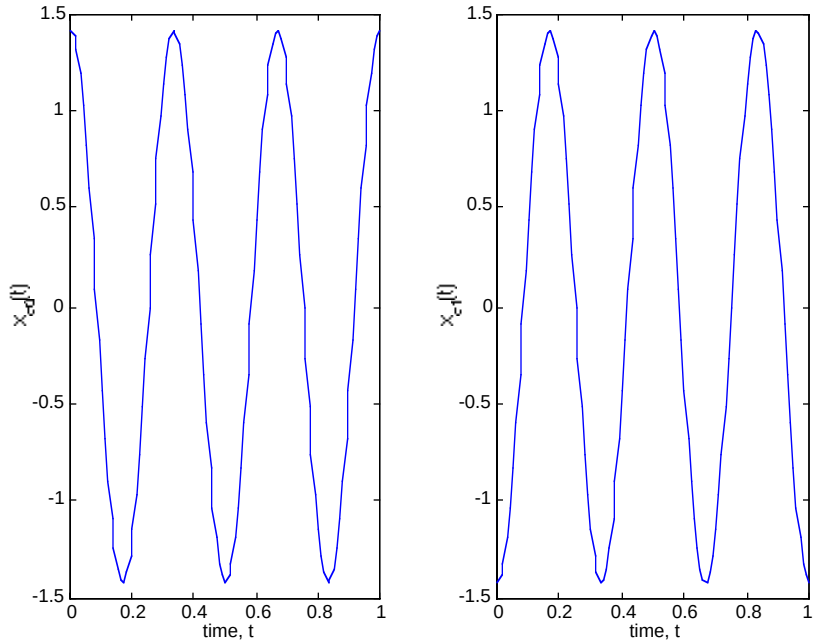


Figure 7.24: The bandpass BPSK waveforms. $\theta = \pi$, $f_c T_p = 3$, and $E_b = T_p$.

The important parameter for determining the performance of the optimum demodulator is the square Euclidean distance. The only parameter that is not specfied in the BPSK model is the phase difference between the two signals. The parameter θ determines the Euclidean distance. Recall again,

$$\Delta_E(1,0) = E_1 + E_0 - 2\Re\left\{\sqrt{E_1 E_0}\rho_{10}\right\}.$$

For BPSK $E_0 = E_1 = E_b$ and

$$\rho_{10} = \exp(j\theta). \tag{7.90}$$

Consequently the Euclidean distance is given as

$$\Delta_E(1,0) = 2E_b\left(1 - \cos(\theta)\right)$$

and the BEP performance is

$$P_B(E) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\Delta_E(1,0)}{4N_0}}\right) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{2N_0}\left(1 - \cos(\theta)\right)}\right). \tag{7.91}$$

It is obvious that optimum performance is obtained for $\theta = \pi$. The optimum signal set has $x_1(t) = -x_0(t)$. Frankly, in BPSK modulation there is no reason to choose anything other than $\theta = \pi$ so in the sequel this particular modulation will be denoted as BPSK. This signal set gives optimum BEP performance. The optimum threshold test simplifies to

$$\int_0^{T_p} Y_I(t)dt \underset{\hat{I}=1}{\overset{\hat{I}=0}{\gtrless}} 0. \tag{7.92}$$

The BEP is

$$P_B(E) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right). \tag{7.93}$$

The bit error probability is plotted versus E_b/N_0 in Fig. 7.21. The spectral efficiency of BPSK can be evaluated by looking at the average energy spectral density and this is plotted in Fig. 7.17. It should be noted that with this form of BPSK the transmission rate is $W_b = 1/T_p$. In examining Fig. 7.17 the bandwidth of BPSK is about $B_T = 1/T_p$ (i.e., 3dB bandwidth) so that the spectral efficiency of BPSK is about $\eta_B = 1$ bits/s/Hz..

The advantages of BPSK are summarized as

- Very simple to generate.
- Simple demodulation structure.
- Optimum BEP performance

The one disadvantage of the BPSK signal set discussed in this section is that the spectral characteristics might not be all that is desired. Method to improve the spectral characteristics will be discussed in the homework.

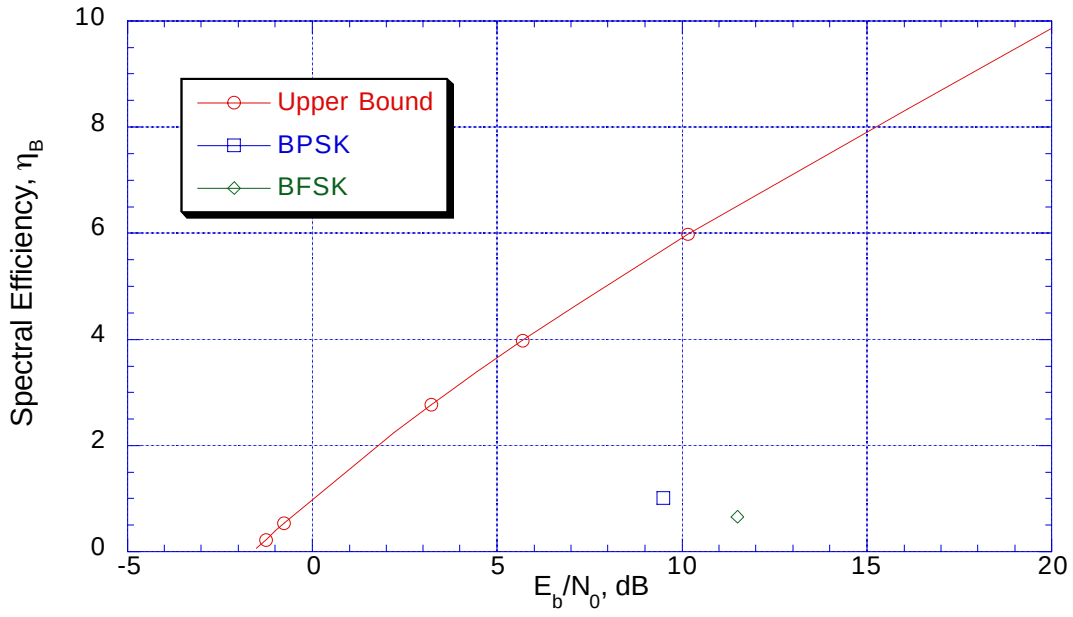


Figure 7.25: A comparison of the spectral efficiency of BPSK and BFSK with the upper bound.

7.7.3 Discussion

This section introduced two example modulations to transmit one bit of information: BFSK and BPSK. These two modulations are the most common used binary modulations in engineering practice. BPSK has the advantages in performance and spectral efficiency. BFSK has some advantages in complexity that become more striking when synchronization issues are considered (which we will do in the sequel). As a final point it is worth comparing the spectral efficiency performance of these two modulations with the upperbounds provided by information theory (see Section 6.3). For the discussion in this text we will denote reliable communication as being an error rate of 10^{-5}. BPSK requires approximately E_b/N_0=9.5dB to achieve reliable communication while BFSK requires approximately E_b/N_0=11.5dB. These two operating points and the upper bound on the possible performance are plotted in Fig. 7.25. It is clear from this figure that BPSK and BFSK result in performance far from the upper bound but that should not concern the reader as there is still a lot of communication theory to explore.

7.8 Homework Problems

Problem 7.1. Consider a binary digital communication system with equally likely bits which is corrupted by an additive white Gaussian noise of two-sided spectral density of $N_0/2$. The two possible transmitted signals are of the form

$$\begin{aligned} x_0(t) &= \sqrt{\frac{E_b}{T_p}}\exp\left[j(2\pi f_0 t+\theta_0)\right] \qquad 0\le t\le T_p \\ &= 0 \qquad \text{elsewhere} \\ x_1(t) &= \sqrt{\frac{E_b}{T_p}}\exp\left[j(2\pi f_1 t+\theta_1)\right] \qquad 0\le t\le T_p \\ &= 0 \qquad \text{elsewhere} \end{aligned} \tag{7.94}$$

a) For $f_0 = 10$Hz, $\theta_0 = 0$ and $f_c = 50$Hz plot $x_c(t)$ when $I = 0$.

b) Give the average energy spectrum per bit $D_{x_z}(f)$ (as a function of f_0 and f_1).

c) Detail out the optimum demodulator. Give impulse responses for any filters and specify any threshold tests.

d) Compute ρ_{01}.

e) If $f_0 = f_1 = 0$, choose values for θ_0 and θ_1 to optimize the performance.

f) If $\theta_0 = 0$ and $\theta_1 = 90^\circ$ choose values for f_0 and f_1 to optimize the performance.

Problem 7.2. For MAPBD with $m_{0,I}(T_p) = 1$, $m_{1,I}(T_p) = 6$, and $\sigma^2_{N_I} = 2$ (see Fig. 7.8)

a) Find the simplest form for MAPBD as a function of π_0. (Hint: it is also a threshold test).

b) Compute and plot the $P_B(E)$ as a function of π_0.

c) Note that $\pi_0 = 0.5$ corresponds to MLBD. Plot the BEP performance if the MLBD was used for $\pi_0 \neq 0.5$ and compare to the results in b).

d) Let $\sigma^2_{N_I}$ be arbitrary and interpret the resulting MAPBD derived in a) when $\sigma^2_{N_I}$ gets large (large noise in observation) and when $\sigma^2_{N_I}$ gets small (little noise in observation). Does the detector characteristics follow intuition? Why?

Problem 7.3. Consider BFSK considered in Section 7.7.1, and find the exact frequency separation which optimizes performance. Note it is not exactly $f_d = \frac{3}{8T_p}$.

Problem 7.4. Consider BFSK considered in Section 7.7.1 on an AWGN channel, and specify for $\pi_0 \neq 0.5$

a) The optimum demodulator.

b) The optimum demodulator performance.

c) The spectral characteristics of the transmitted signal.

Consider the special case of $\pi_0 = 0.25$ in detail and produced plots for parts b) and c).

Problem 7.5. In Problem 2.12 three pulse shapes were presented. Assume one bit is to be transmitted on an AWGN channel and π_0=0.5.

a) Pick two out of the three pulse shapes to represent the possible value of the bit being transmitted in a single shot binary system such that the probability of bit error is minimized. Give the resulting correlation coefficient and probability of error when the one sided noise spectral density is N_0.

b) Antipodal signalling is the optimal binary communications design. Choose one of the three pulse shapes to use in an antipodal binary communication system which would make the most efficient use of bandwidth? Justify your answer.

Problem 7.6. Ethernet uses a biphase modulation to transmit data bits. Bi-phase modulation is a baseband linear modulation where each bit, I, is transmitted with $x_0(t) = u(t)$ or $x_1(t) = -u(t)$ where the pulse shape, $u(t)$ given in Fig.7.26. Assume an AWGN channel and equally likely data bits.

a) Compute the resulting BEP.

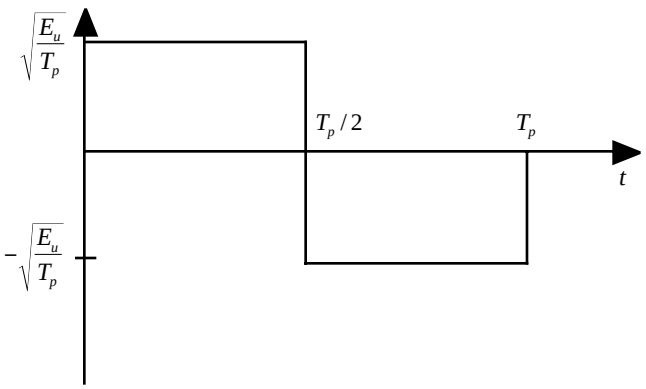


Figure 7.26: The pulse used in Ethernet.

b) If $1/T_p$=10MHz and the Ethernet signal at the transmitter has a peak value of 1 volt (assume a 1 Ω system) and the value of N_0=-160dBm/Hz, find the amount of cable loss, L_c, that can be tolerated and still achieve a 10^{-9} error rate. Note these conditions imply that in Figure 7.26 $\sqrt{E_u/T_p} = L_c$.

c) Compute the average energy spectrum per bit for the Ethernet modulation.

d) Postulate why this pulse shape was chosen for transmission over cables.

Problem 7.7. Recall that VSB modulation has the form $x_z(t) = x_I(t) + j\,(h_v(t) \otimes x_I(t))$. Assume that $x_I(t)$ is chosen from a BPSK signal set where the duration of the carrier pulse is given as T and that the support for the quadrature modulation $x_Q(t) = h_v(t) \otimes x_I(t)$ is given as $T_p > T$. Find the simplest form of the optimum demodulator and compute the resulting BEP.

Problem 7.8. A bandpass binary digital communication system with equally likely bits was designed to have $x_z(t) = d_i u(t)$ where $d_0 = 1$ and $d_1 = -1$. The hardware implementation resulted in the actual transmitted modulation symbols being $d_0 = 1$ and $d_1 = \exp\left[j\frac{7\pi}{8}\right]$.

a) If the demodulator was built to be optimum for the design waveforms, what would be the resulting probability of error.

b) Give the optimum demodulation structure for the actual transmitted waveforms.

c) Give the performance for the optimum demodulator for the actual transmitted waveforms.

Problem 7.9. It is possible to construct a binary signaling set with $\rho_{01} \neq -1$ which has better performance than a signal set with $\rho_{01} = -1$. Construct two signal sets (one with $\rho_{01} \neq -1$ and one with $\rho_{01} = -1$) with the same E_b such that the $\rho_{01} \neq -1$ signal set has a better performance when the corrupting noise in an AWGN than the $\rho_{01} = -1$ signal set.

Problem 7.10. In this problem we consider a form of binary combined phase and frequency modulation. Assume equally likely data bits and an AWGN at the receiver that corrupts the transmission. The two possible baseband transmitted signals are

$$\begin{aligned} x_0(t) &= \sqrt{\frac{E_b}{T_p}} \exp\left[\frac{j\pi t}{4T_p}\right] \qquad 0 \leq t \leq T_p \\ &= 0 \qquad \text{elsewhere} \end{aligned} \tag{7.95}$$

$$\begin{aligned} x_1(t) &= \sqrt{\frac{E_b}{T_p}} \exp\left[-\frac{j\pi t}{4T_p} + j\theta\right] \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{elsewhere} \end{aligned} \tag{7.96}$$

a) For $f_c = 4/T_p$ plot the bandpass signal corresponding to $x_0(t)$.

b Find the demodulator that minimizes the probability of bit error for a single bit transmission given

$$Y_z(t) = X_z(t) + W_z(t)$$

where $W_z(t)$ is a complex additive white Gaussian noise of one-sided spectral density N_0.

c) Give an exact expression for the bit error probability in terms of E_b, θ, and N_0 for this signal set and optimum demodulator. What is the value of θ that minimizes the probability of error?

d) Compute the average energy spectrum per bit. What is the spectral efficiency using the 3dB bandwidth?

Problem 7.11. You are called in by Nokey to be a consultant to the design of a binary digital communications system for use in an AWGN channel. With the assumpion of equally likely bits, one waveform chosen was

$$\begin{aligned} x_0(t) &= \sqrt{2}\sin\left(\frac{2\pi t}{T_p}\right) \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{elsewhere} \end{aligned}$$

The engineers at Nokey are hotly debating between the following two waveforms as the second waveform.

$$\begin{aligned} x_{1a}(t) &= \sqrt{2}\cos\left(\frac{2\pi t}{T_p}\right) \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{elsewhere} \end{aligned}$$

and

$$\begin{aligned} x_{1b}(t) &= -1 \qquad 0 \le t \le \frac{T_p}{2} \\ &= 0 \qquad \text{elsewhere} \end{aligned}$$

a) Select the better of the two waveforms for bit error performance and give the reasons for your selection.

b) Give the optimum demodulation structure for both pairs of waveforms. Detail out the matched filter response(s) and threshold tests for each possible pair of waveforms.

c) Which set of waveforms would have better spectral characteristics? Explain why.

Problem 7.12. Show that optimum demodulation for the signal model given in (7.2) is given by the system shown in Fig. 7.27. Compute the value of $\gamma(L_p)$.

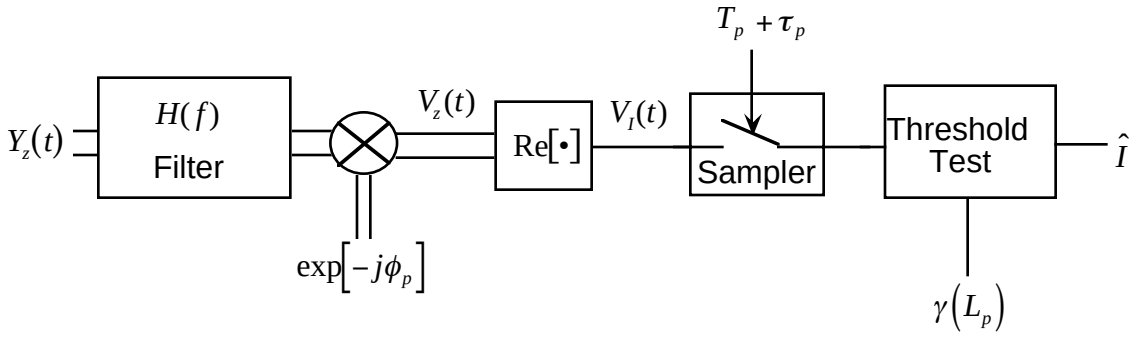


Figure 7.27: The optimal demodulator for the signal model in (7.2).

Problem 7.13. In certain situations it is useful to hop to many different frequencies during transmission of a bit (e.g., to avoid intentional or unintentional jamming). Consider such a communication system where two hops are made during one bit transmission so that the transmitted signal has the form

$$X_z(t) = \begin{cases} D_z\sqrt{\dfrac{E_b}{T_p}}\exp\left[j2\pi f_1 t\right] & 0 \le t \le \frac{T_p}{2} \\ D_z\sqrt{\dfrac{E_b}{T_p}}\exp\left[j2\pi f_2 t\right] & \frac{T_p}{2} < t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{7.97}$$

where $D_z = d_i$ for $I = i$. Assume each bit value is equally likely and that the received signal is of the form

$$Y_z(t) = X_z(t) + W_z(t) \tag{7.98}$$

where $W_z(t)$ is a complex additive white Gaussian noise of one-sided spectral density N_0.

a) Detail out the optimum demodulator and simplify as much as possible.

b) Give the bit error probability as a function of d_0, d_1, f_1, and f_2.

c) Optimize the performance with a selection of d_0, d_1, f_1, and f_2 under the constraint that $E_b = (E_0 + E_1)/2$.

d) Is it possible to select frequencies f_1 and f_2, $f_1 \neq f_2$ such that $X_P(t)$ is continuous. If yes give an example.

e) Find and plot $D_{x_z}(f)$ for $f_1 = -2/T_p$ and $f_2 = 1.5/T_p$.

Problem 7.14. In this problem we consider a form of baseband binary modulation where the signals are restricted to be real and positive valued. Assume equally likely data bits and an additive white Gaussian noise at the receiver that corrupts the transmission. Two possible sets of baseband transmitted signals are signal set A

$$\begin{aligned} x_{0A}(t) &= C_A \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{elsewhere} \end{aligned} \tag{7.99}$$

$$x_{1A}(t) = 0 \tag{7.100}$$

and signal set B

$$\begin{aligned} x_{0B}(t) &= C_B \cos\left(\frac{\pi t}{2T_p}\right) \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{elsewhere} \end{aligned} \tag{7.101}$$

$$\begin{aligned} x_{1B}(t) &= C_B \sin\left(\frac{\pi t}{2T_p}\right) \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{elsewhere} \end{aligned} \tag{7.102}$$

a) What is the average energy per bit for these two signal sets (e.g., $E_b(A)$ and $E_b(B)$) as a function of C_A and C_B?

b) Find the demodulator that minimizes the probability of bit error for each signal set for $E_b = 1$ in a single bit transmission given

$$Y_z(t) = X_z(t) + W_z(t)$$

where $W_z(t)$ is an additive white Gaussian noise of one-sided spectral density N_0.

c) Which signaling scheme produces better performance for $E_b(A) = E_b(B)$=1. Quantify the increase in SNR which would be necessary in the worse performing scheme to give the two schemes equal performance.

d) How much loss in performance is incurred for the best of these two schemes compared to the case of optimum signaling with E_b=1 where the signals are allowed to take positive and negative values. Give an example set of two waveforms that achieves this optimum performance.

e) Find the average energy spectrum per bit for signal set A. Give a signal set defined over the same interval $[0, T_p]$ with the same performance as signal set A but with better spectral characteristics. Be specific about exactly which spectral characteristics are better.

Problem 7.15. This problem is concerned with the transmission of one bit of information, I, by the waveforms $x_0(t)$ and $x_1(t)$ that have energy E_{x0} and E_{x1} respectively during the time interval $[0, T_p]$. Assume demodulation is to be done in the presence of an additive white Gaussian noise, $W_z(t)$, with one-sided spectral density of N_0. A common impairment in radio circuits is a DC offset added to the signal (most often in either analog up or down conversion). Consequently the received signal has the form

$$Y_z(t) = C + x_i(t) + W_z(t) \tag{7.103}$$

where C is a known complex constant.

a) If $C = 0$ what are the optimum waveforms when the received average energy per bit is constrained to be E_b? Sketch the maximum likelihood bit demodulator. Give the probability of bit error as a function of E_b and N_0.

b) For $C \neq 0$ find E_1 and E_0 and E_b as a function of C.

c) Consider a maximum likelihood bit demodulator. For $C \neq 0$ what are the optimum waveforms when the received energy per bit is constrained to be E_b? Sketch the simplest maximum likelihood bit demodulator. Detail out the filters and the threshold test.

d) Show a judicious choice of transmitted waveforms leads to a demodulator that is not a function of C.

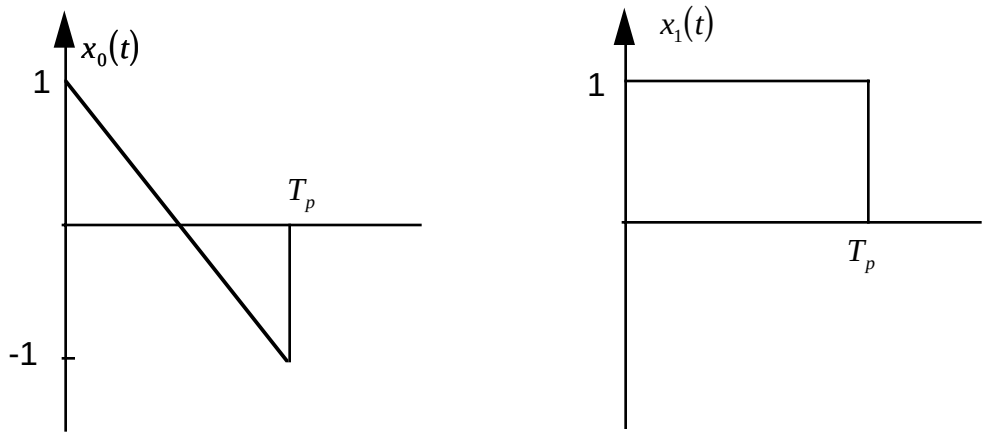


Figure 7.28: A binary signal set.

e) For $C \neq 0$ and the optimum waveforms from c) give the probability of bit error as a function of E_b, C, and N_0.

f) For $C \neq 0$ formulate the maximum a posteriori bit demodulator.

g) For $C \neq 0$ specify a set of waveforms $x_0(t)$ and $x_1(t)$ such that the spectral efficiency can be argued to be near optimum.

Problem 7.16. A binary digital communication system uses the two waveforms in Fig. 7.28 to communicate a bit of information. Assume these signals are corrupted by an additive white Gaussian noise with a two sided spectral density of $N_0/2$ and that the bits are equally likely.

a) Sketch the simplest form for optimum bit demodulator.

b) Calculate E_0, E_1, and ρ_{10}.

c) Calculate the Euclidean square distance between $x_0(t)$ and $x_1(t)$, $\Delta_E(1,0)$.

d) What is the resulting bit error probability

e) Keep one of either $x_0(t)$ or $x_1(t)$ and select a third signal $x_2(t)$ such that $|x_2(t)| \leq 1$ and $x_2(t)$ is of length T_p to be used to transmit one bit of information such that the resulting bit error probability for the optimum demodulation is as low as possible

Problem 7.17. Digital communication waveforms can often be viewed as analog communication waveforms with specific message waveforms.

a) To this end show that BFSK is equivalent to FM modulating one of two waveforms, i.e.,

$$x_i(t) = A_c \exp\left[j2\pi f_k \int_{-\infty}^{t} m_i(\lambda)d\lambda\right] \tag{7.104}$$

where

$$m_0(t) = \begin{cases} 1 & 0 \leq t \leq T_p \\ 0 & \text{elsewhere} \end{cases} \qquad m_1(t) = \begin{cases} -1 & 0 \leq t \leq T_p \\ 0 & \text{elsewhere.} \end{cases} \tag{7.105}$$

Identify f_k and A_c.

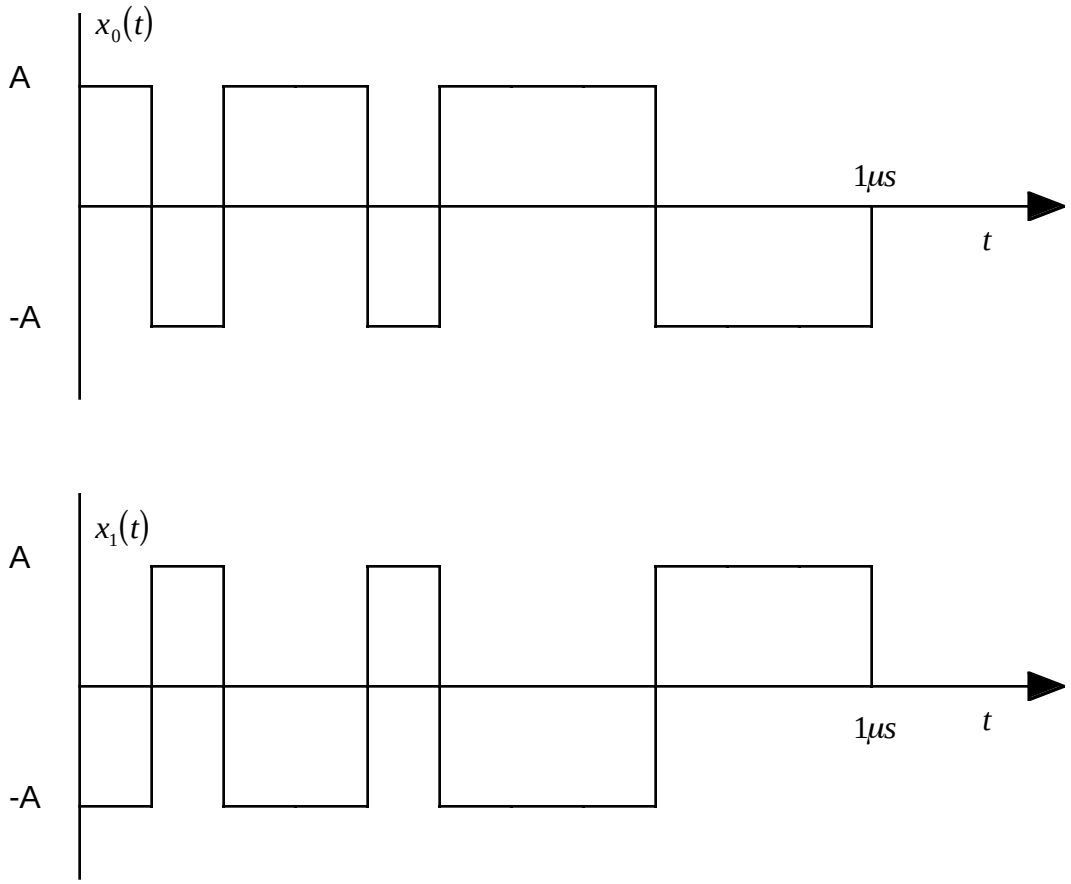


Figure 7.29: The two signals used in a binary modulation for IEEE 802.11b.

b) To this end show that BPSK is equivalent to DSB-AM modulating one of the two waveforms given in (7.105). Identify A_c.

Problem 7.18. In the wireless local network protocol denoted IEEE 802.11b the lowest rate modulation is a binary modulation using the two waveforms given in Fig. 7.29. Assume $\pi_0 = 0.5$ and that $\tau_1 = 1/11\mu$s, $\tau_2 = 2/11\mu$s, and $\tau_3 = 3/11\mu$s .

a) Choose A such that the energy per bit is E_b.

b) What is the bit rate?

c) Keeping $x_0(t)$, can you choose a new waveform different than $x_1(t)$ that will have the same energy and will improve the performance? If you can then show an example.

d) Assume equally likely bits and a complex additive white Gaussian noise with a one-sided power spectral density N_0 corrupts the received signal. Find the simplest form of the optimum demodulator.

e) Find $D_{x_z}(f)$. Using the 3dB bandwidth what is the spectral efficiency of this modulation?

Problem 7.19. You have decided to try to predict the outcome of the next presidential race by using your electrical engineering knowledge learned your communication theory class. You decide to use one outcome of an approval rating poll for the current president in the summer before the election as the

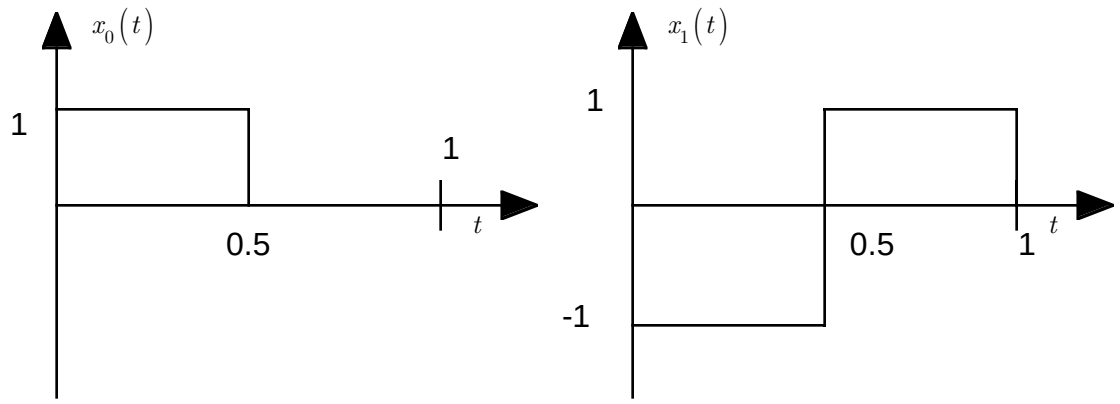


Figure 7.30: Two waveforms to send one bit.

observation. This observation is a random variable V. The minimum error rate decision rule was given as

$$f_{V|1}(v|I=1)\pi_1 \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} f_{V|0}(v|I=0)\pi_0. \tag{7.106}$$

where π_0 is the a priori probability that the challenging party wins and π_1 is the a priori probability the incumbent party wins. After careful analysis of the polling data you realize that

$$\begin{aligned} V &= 53+N \qquad I=1 \text{ (when the incumbent party wins)} \\ V &= 49+N \qquad I=0 \text{ (when the challenging party wins)} \end{aligned} \tag{7.107}$$

where N is a zero mean Gaussian random variable with $E\left[N^2\right]=4$.

a) If $\pi_0 = 0.4$ find π_1.

b) If $\pi_0 = 0.4$ and $v = 51$ is observed, make the minimum probability of error decision.

c) If $\pi_0 = 0.5$ and $v = 50$ is observed, make the minimum probability of error decision.

d) If $\pi_0 = 0.5$ find the probability of error in the optimum test when the incumbent wins.

e) If $v = 50$ find the region for the values of π_0 where the decision will be opposite of that found in c).

Problem 7.20. The two waveforms in Fig. 7.30 are to be used to send one bit. Assume that $\pi_0 = 0.5$.

a) Compute E_0 and E_1.

b) Specify the optimum demodulator structure. Clearly specify any threshold test, filter impulse responses, and/or sample times.

c) Compute the Euclidean distance between $x_0(t)$ and $x_1(t)$.

Problem 7.13. This signalling scheme is known as frequency hopping.

a) This modulation has the form $X_i(t) = d_i u(t)$.

$$\begin{aligned} Q &= \int Y_z(t)u^*(t)dt \\ &= \int_0^{\frac{T_p}{2}} Y_z(t)\sqrt{\frac{E_b}{T_p}}\exp\left[-j2\pi f_1 t\right]dt + \int_{\frac{T_p}{2}}^{T_p} Y_z(t)\sqrt{\frac{E_b}{T_p}}\exp\left[-j2\pi f_2 t\right]dt \end{aligned}$$

$$\begin{aligned} E_i &= |d_i|^2 E_u = |d_i|^2 E_b \\ T_0 &= \Re[d_0^* Q] - \frac{|d_0|^2 E_b}{2} \\ T_1 &= \Re[d_1^* Q] - \frac{|d_1|^2 E_b}{2} \end{aligned}$$

b)

$$\begin{aligned} P_B(E) &= \frac{1}{2}\text{erfc}\left[\sqrt{\frac{\Delta_E(1,0)}{4N_0}}\right] \\ \text{where} \quad \Delta_E(1,0) &= |d_0 - d_1|^2 E_b \end{aligned}$$

c)

$$\begin{aligned} d_0 &= 1 \\ d_1 &= -1 \end{aligned} \quad \Longrightarrow \quad \Delta_E(1,0) = 4E_b$$

d) Only point of possible discontinuity is at $t = T_p/2$.

$$\begin{aligned} &2\pi f_1 \frac{T_p}{2} = 2\pi f_2 \frac{T_p}{2} \\ &2\pi f_1 \frac{T_p}{2} = 2\pi f_2 \frac{T_p}{2} + 2\pi n \qquad \text{where } n \text{ is an integer} \\ &f_2 = f_1 + \frac{2n}{T_p} \end{aligned}$$

e)

$$
\begin{aligned}
&G_{X_0}(f) = |X_0(f)|^2 = |d_0|^2|U(f)|^2 \\
&u(t) = u_1(t) + u_2(t) \\
&u_1(t) = \begin{cases} \sqrt{\frac{E_b}{T_p}} \exp[j2\pi f_1 t] & 0 \le t \le \frac{T_p}{2} \\ 0 & \text{elsewhere} \end{cases} \\
&u_2(t) = \begin{cases} \sqrt{\frac{E_b}{T_p}} \exp[j2\pi f_2 t] & \frac{T_p}{2} \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \\
&U_1(f) = \sqrt{E_b T_p} \exp\left[-j\frac{\pi f T_p}{2}\right] \operatorname{sinc}\left(\frac{(f-f_1)T_p}{2}\right) \\
&U_2(f) = \sqrt{E_b T_p} \exp\left[-j\frac{3\pi f T_p}{2}\right] \operatorname{sinc}\left(\frac{(f-f_2)T_p}{2}\right) \\
&G_{X_0}(f) = |d_0|^2|U_1(f) + U_2(f)|^2 \\
&G_{X_1}(f) = |d_1|^2|U_1(f) + U_2(f)|^2 \\
&D_{X_z}(f) = \frac{|d_0|^2 + |d_1|^2}{2}|U_1(f) + U_2(f)|^2
\end{aligned}
$$

See Matlab code and plot in Fig. 7.31.

```
%
% Generating the spectrum of FH
% Author: M. Fitz
% Last modified: 1/28/04
%
close all
clear all
numpts=2048;
f_1=2;
f_2=-2;
freq=linspace(-4,4, numpts+1);
g_pr=zeros(1,numpts+1);
dum1=exp(-j*pi*freq/2).*sinc(freq-f_1*ones(1,numpts+1));
dum2=exp(-j*3*pi*freq/2).*sinc(freq-f_2*ones(1,numpts+1));
g_pr=(dum1+dum2).*conj(dum1+dum2);
g_prdb=10*log10(g_pr);
figure(1)
plot(freq,g_prdb)
axis([-4 4 -50 5])
xlabel('Normalized frequency, fT_p')
ylabel('Normalized energy spectrum per bit')
hold off
```

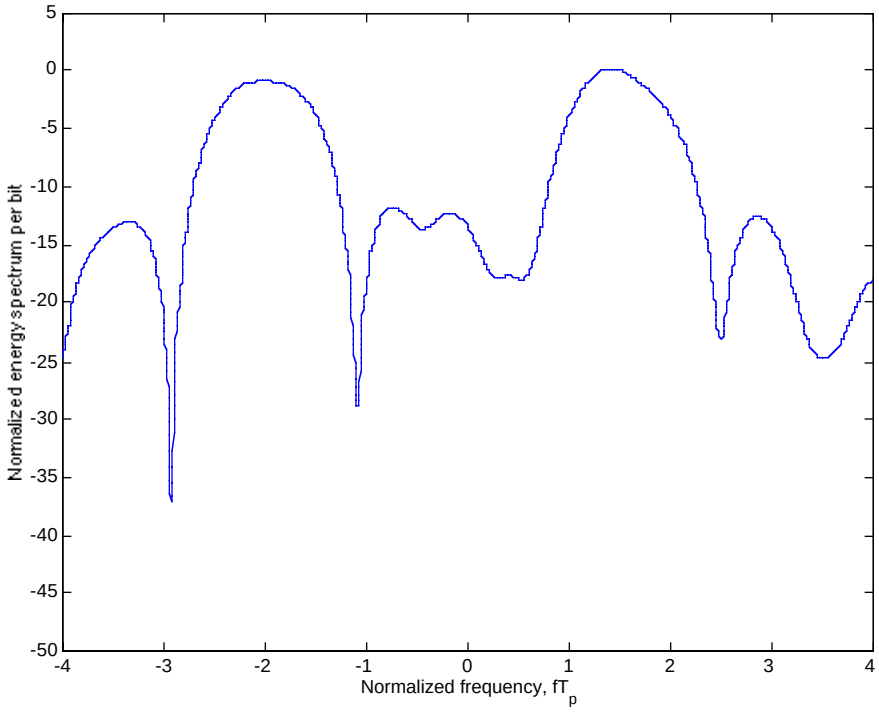


Figure 7.31: Spectrum of a frequency hopped signal.

7.9 Projects

7.9.1 Project 1

Project Goals

1. Design a pulse shape to meet a typical spectral emissions mask.
2. Compute the probability of error for both bandpass and baseband matched filtering in the presence of sample time error to show the advantages of processing signals at baseband.

Design a pulse shape that can be used with binary pulse amplitude modulation having the form

$$x_c(t) = \pm u(t)\sqrt{2}\cos(2\pi f_c t)$$

that will meet the spectral emissions mask given in Fig. 7.32 but the pulse must not extend for longer than 80μ s in time. The carrier frequency is 1MHz. This spectral emissions mask is the one used for GSM handsets. GSM is a second generation cellular telephony standard. So this problem is one of great practical interest.

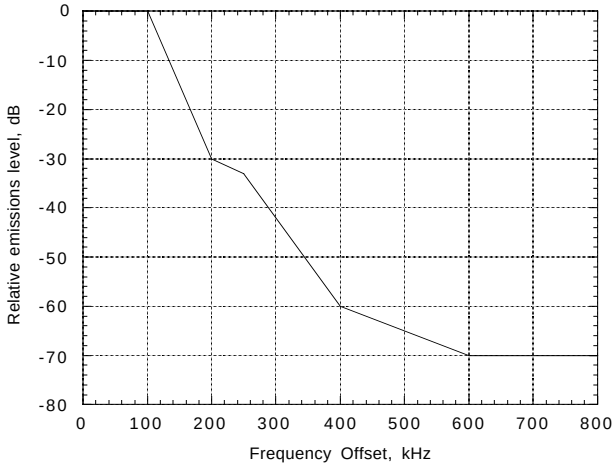


Figure 7.32: Spectral emissions mask.

We have discussed one type of demodulator structures for bandpass signals; baseband matchd filters. This structure match filters to the bandpass signal after creating the complex envelope of the received signal (see Fig. 7.33). It is also possible to build a matched filter to the bandpass signal itself ($h(t) = x_{1c}(T_p - t) - x_{0c}(T_p - t)$)

a) Derive the form for and plot the output of the bandpass matched filter for all time when $x_c(t) = u(t)\sqrt{2}\cos(2\pi f_c t)$.

b) Derive the form for and plot the output of the baseband matched filter when $x_c(t) = u(t)\sqrt{2}\cos(2\pi f_c t)$.

c) Compute the E_b/N_0 required to produce a 10^{-5} error probability for the optimum demodulator.

d) If the bandpass matched filter is sampled at $T_p + 2 \times 10^{-7}$ instead of T_p compute the E_b/N_0 required to produce a 10^{-5} error probability.

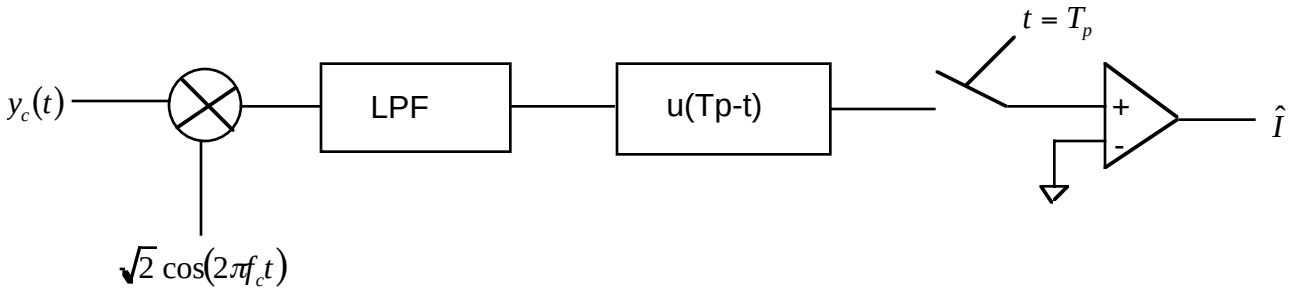


Figure 7.33: The baseband demodulator for binary pulse amplitude modulation.

e) If the baseband matched filter is sampled at $T_p + 2 \times 10^{-7}$ instead of T_p compute the E_b/N_0 required to produce a 10^{-5} error probability.

f) How big would the timing error have to be for the baseband matched filter to achieve the same answer as in d)?

Chapter 8

Transmitting More Than One Bit

People who buy communication systems rarely want to only transmit one bit. Consequently there is a practical need to generalize the results from the previous chapter to the transmission of multiple bits. The problem considered here is that we transmit K_b bits by selecting one of $M = 2^{K_b}$ waveforms. In the literature this is often denoted M-ary digital communications. The demodulator has the task of deciding which of these waveforms was sent in the presence of an additive noise. This chapter will consider a variety of problems associated with transmitting more than one bit.

8.1 A Reformulation For One Bit Transmission

The optimum demodulator can be reformulated in several ways. Many of these reformulations give insight for receiver structures we will consider in the sequel. Recall from Chapter 7 that the MLBD, when $Y_z(t) = y_z(t)$ is observed, has the form

$$v_I(T_p) = \Re\left[\int_{-\infty}^{\infty} y_z(\tau)\left(x_1^*(\tau) - x_0^*(\tau)\right)d\tau\right] \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \frac{E_1 - E_0}{2}. \tag{8.1}$$

Rearranging (8.1) gives

$$\Re\left[\int_{-\infty}^{\infty} y_z(\tau)x_1^*(\tau)d\tau\right] - \frac{E_1}{2} \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \Re\left[\int_{-\infty}^{\infty} y_z(\tau)x_0^*(\tau)d\tau\right] - \frac{E_0}{2}. \tag{8.2}$$

The form of the MLBD in (8.2) which forms a decision statistic for each signal by computing the output of a matched filter to the transmitted signal and adding an energy correction term will frequently appear in the sequel. In fact to simplify the discussion the following definition is useful.

Definition 8.1 *The maximum likelihood metric for $I = i$ is*

$$T_i = \Re\left[\int_{-\infty}^{\infty} y_z(\tau)x_i^*(\tau)d\tau\right] - \frac{E_i}{2}. \tag{8.3}$$

Using this definition allows a reformulation of the MLBD as

$$T_1 \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} T_0. \tag{8.4}$$

Equation (8.4) gives a particularly insightful form of the decision process: form the maximum likelihood metric for each bit value and choose to decode the bit value corresponding to the largest metric.

The MAPBD can similarly take other useful forms. Multiplying both sides of (8.2) by $2/N_0$ and taking the exponential gives another form for the MLBD, i.e.,

$$\exp\left[\frac{2}{N_0}\Re\left[\int_{-\infty}^{\infty} y_z(\tau)x_1^*(\tau)d\tau\right] - \frac{E_1}{N_0}\right] \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \exp\left[\frac{2}{N_0}\Re\left[\int_{-\infty}^{\infty} y_z(\tau)x_0^*(\tau)d\tau\right] - \frac{E_0}{N_0}\right]. \tag{8.5}$$

The justification of why the constant term is $2/N_0$ is provided in a book on detection theory [Poo88] and briefly highlighted in Appendix A. The generalization to unequal priors has the MAPBD taking the form

$$\exp\left[\frac{2}{N_0}\Re\left[\int_{-\infty}^{\infty} y_z(\tau)x_1^*(\tau)d\tau\right] - \frac{E_1}{N_0}\right]\pi_1 \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \exp\left[\frac{2}{N_0}\Re\left[\int_{-\infty}^{\infty} y_z(\tau)x_0^*(\tau)d\tau\right] - \frac{E_0}{N_0}\right]\pi_0. \tag{8.6}$$

Using the maximum likelihood metric gives the MAPBD as

$$\exp\left[\frac{2T_1}{N_0}\right]\pi_1 \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} \exp\left[\frac{2T_0}{N_0}\right]\pi_0. \tag{8.7}$$

In fact this demodulator structure can be viewed as being equivalent to computing the APP for each of the two possible bits given the receiver input and then selecting the bit value that corresponds to this maximum, i.e.,

$$P\left(I=1\,|y_z(t)\right) \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} P\left(I=0\,|y_z(t)\right). \tag{8.8}$$

where $y_z(t)$ is the observed receiver input. Forms for the optimal detector given in (8.5) and (8.6) will be used often in the sequel to form optimum detectors in some special cases.

8.2 Optimum Demodulation

Now we want to consider the digital communication system design problem of transmitting more than one bit of information. In the sequel K_b shall denote the number of bits to be transmitted and the bit sequence is denoted

$$\vec{I} = [I(1)\ I(2)\ \ldots\ I(K_b)]^T \tag{8.9}$$

where $I(k)$ take values 0,1. For simplicity of notation a particular sequence of bits can be designated by the numeric value the bits represent, i.e., $\vec{I} = i,\ i \in \{0,\ldots,M-1\}$ where $M = 2^{K_b}$. To represent the M values the bit sequence can take, M different analog waveforms should be available for transmission. Denote by $x_i(t)$, $i \in \{0,\ldots,M-1\}$ the waveform transmitted when $\vec{I} = i$ is to be transmitted. Here again we will assume the analog waveforms have support on $t \in [0, T_p]$. Given this problem formulation we want to extend the results obtained in the previous chapter. It is worth noting that this problem formulation gives a bit rate of $W_b = K_b/T_p$ bits per second. The average energy per bit is given as

$$E_b = \frac{1}{K_b}\sum_{i=0}^{M-1}\pi_i E_i \tag{8.10}$$

where again E_i is the energy of waveform $x_i(t)$.

Example 8.1: Consider the case of $K_b = 3$, the bits and the words are enumerated as

$I(1)$	$I(2)$	$I(3)$	i
0	0	0	0
0	0	1	1
0	1	0	2
0	1	1	3
1	0	0	4
1	0	1	5
1	1	0	6
1	1	1	7

Example 8.2: Here we consider a modulation known as pulse width modulation with $K_b = 2$. This modulation embeds information in the width of the pulse that is transmitted. This is a popular baseband modulation in certain applications where positive logic level signals are desired in the transmitter. The transmitted signals are given as

$$x_0(t) = \begin{cases} \sqrt{\frac{32E_b}{10T_p}} & 0 \le t \le T_p/4 \\ 0 & \text{elsewhere} \end{cases} \qquad x_1(t) = \begin{cases} \sqrt{\frac{32E_b}{10T_p}} & 0 \le t \le T_p/2 \\ 0 & \text{elsewhere} \end{cases} \tag{8.11}$$

$$x_2(t) = \begin{cases} \sqrt{\frac{32E_b}{10T_p}} & 0 \le t \le 3T_p/4 \\ 0 & \text{elsewhere} \end{cases} \qquad x_3(t) = \begin{cases} \sqrt{\frac{32E_b}{10T_p}} & 0 \le t \le T_p \\ 0 & \text{elsewhere.} \end{cases} \tag{8.12}$$

8.2.1 Optimum Word Demodulation Receivers

A first receiver to be considered is the maximum a posteriori word demodulator (MAPWD). This MAPWD can again be shown to be the minimum word error probability receiver using Bayes detection theory. A straightforward generalization of (8.8) leads to

$$\hat{\vec{I}} = \arg \max_{i \in \{0,\ldots,M-1\}} P\left(\vec{I} = i \,|\, y_z(t)\right) \tag{8.13}$$

where the arg max notation refers to the particular binary word that has the maximum APP. This is the obvious generalization of binary MAP detection to the M-ary decision problem. Using the results of the previous section this MAP decoding rule becomes

$$\begin{aligned} \hat{\vec{I}} &= \arg \max_{i \in \{0,\ldots,M-1\}} \exp\left[\frac{2}{N_0} \Re\left[\int_{-\infty}^{\infty} y_z(\tau) x_i^*(\tau) d\tau\right] - \frac{E_i}{N_0}\right] \pi_i \\ &= \arg \max_{i \in \{0,\ldots,M-1\}} \exp\left[\frac{2}{N_0} \Re\left[v_i(T_p)\right] - \frac{E_i}{N_0}\right] \pi_i \qquad (8.14) \\ &= \arg \max_{i \in \{0,\ldots,M-1\}} \exp\left[\frac{2T_i}{N_0}\right] \pi_i \qquad (8.15) \end{aligned}$$

where

$$V_i(t) = \int_{-\infty}^{\infty} Y_z(\tau) x_i^*(T_p - t + \tau) d\tau \tag{8.16}$$

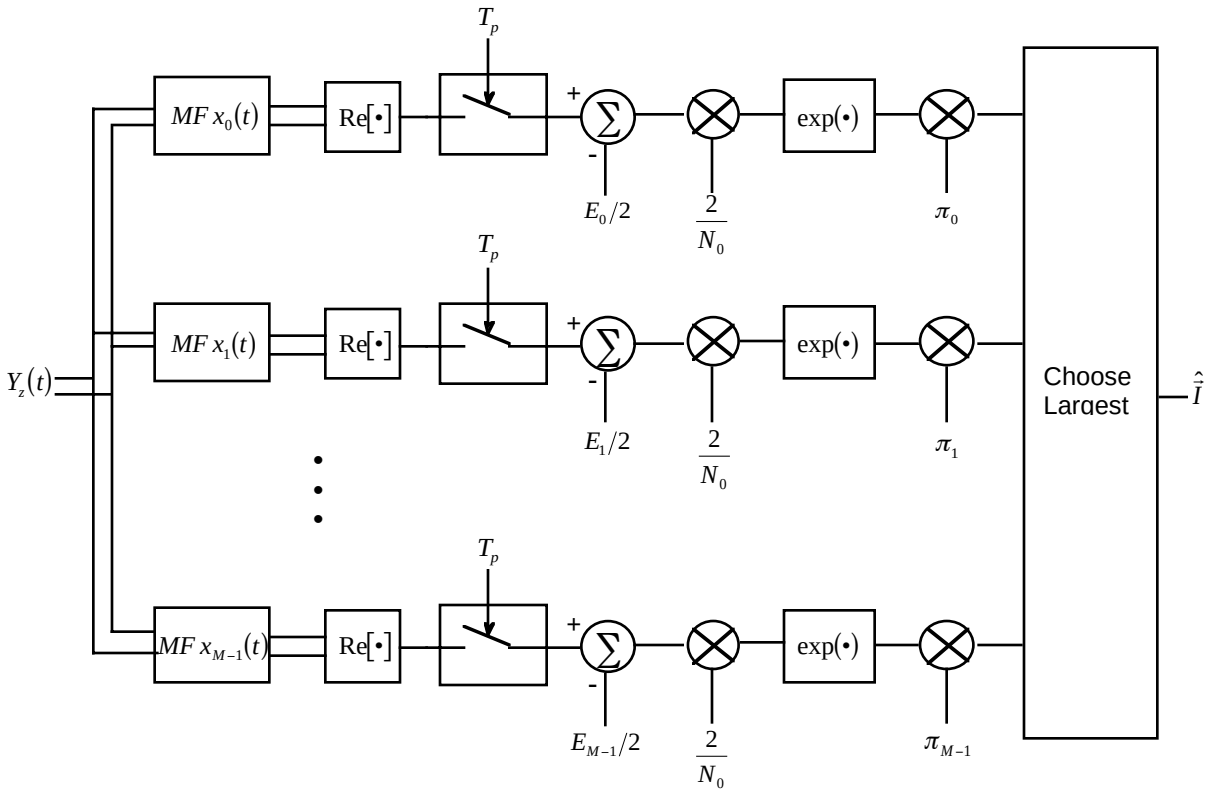


Figure 8.1: The block diagram for the M-ary MAP demodulator.

is denoted the i^{th} matched filter output and

$$E_i = \int_{-\infty}^{\infty} |x_i(t)|^2 \, dt \tag{8.17}$$

is the energy of the i^{th} analog waveform. The i^{th} matched filter output when sampled at $t = T_p$ again will give a correlation of the received signal with the i^{th} possible transmitted signal. Here T_i represents the maximum likelihood metric for the word $\vec{I} = i$. For future use it is important to realize that

$$\exp\left[\frac{2T_i}{N_0}\right]\pi_i = \exp\left[\frac{2}{N_0}\Re\left[v_i(T_p)\right] - \frac{E_i}{N_0}\right]\pi_i = CP\left(\vec{I} = i\,|y_z(t)\right) \tag{8.18}$$

where C is a constant that is not a function of $\vec{I}$. The block diagram for the optimum MAPWD is shown in Fig. 8.1.

The demodulator in the case of equal priors, i.e., $\pi_i = 1/M,\ \forall i$, can be greatly simplified. We will denote this demodulator as the maximum likelihood word demodulator (MLWD). Since all terms have an equal π_i this common term can be canceled from each term in the decision rule and since the $\ln(\bullet)$ function is monotonic the MLWD is given as

$$\begin{aligned} \hat{\vec{I}} &= \arg \max_{i\in\{0,\ldots,M-1\}} \Re\left[v_i(T_p)\right] - \frac{E_i}{2} \\ &= \arg \max_{i\in\{0,\ldots,M-1\}} T_i. \end{aligned} \tag{8.19}$$

The maximum likelihood metric is computed for each of the possible transmitted signals. Decoding is accomplished by selecting the binary word associated with the largest decision metric. The block

diagram for the optimum MLWD is shown in Fig. 8.2.

Example 8.3: Returning to the pulse width modulation example the four matched filter outputs are given as

$$\begin{aligned} v_0(T_p) &= \sqrt{\frac{32E_b}{10T_p}}\int_0^{T_p/4} y_z(\tau)d\tau \qquad v_1(T_p) = \sqrt{\frac{32E_b}{10T_p}}\int_0^{T_p/2} y_z(\tau)d\tau \\ v_2(T_p) &= \sqrt{\frac{32E_b}{10T_p}}\int_0^{3T_p/4} y_z(\tau)d\tau \qquad v_3(T_p) = \sqrt{\frac{32E_b}{10T_p}}\int_0^{T_p} y_z(\tau)d\tau. \end{aligned} \tag{8.20}$$

The energy of each of the transmitted signals is given as

$$E_0 = \frac{8E_b}{10} \qquad E_1 = \frac{16E_b}{10} \qquad E_2 = \frac{24E_b}{10} \qquad E_3 = \frac{32E_b}{10}. \tag{8.21}$$

The maximum likelihood metrics for this modulation are given as

$$\begin{aligned} T_0 &= \sqrt{\frac{32E_b}{10T_p}}\int_0^{T_p/4} y_I(\tau)d\tau - \frac{4E_b}{10} \qquad T_1 = \sqrt{\frac{32E_b}{10T_p}}\int_0^{T_p/2} y_I(\tau)d\tau - \frac{8E_b}{10} \\ T_2 &= \sqrt{\frac{32E_b}{10T_p}}\int_0^{3T_p/4} y_I(\tau)d\tau - \frac{12E_b}{10} \qquad T_3 = \sqrt{\frac{32E_b}{10T_p}}\int_0^{T_p} y_I(\tau)d\tau - \frac{16E_b}{10}. \end{aligned} \tag{8.22}$$

The important thing to notice for both MAPWD and MLWD is that the optimal demodulator complexity increases exponentially with the number of bits transmitted. The number of matched filters required in each demodulator is $M = 2^{K_b}$ consequently the complexity of these demodulation schemes is $O(2^{K_b})$. The notation $O(x)$ implies the complexity of the algorithm is proportional to x, i.e., a constant times x. This complexity is obviously unacceptable if large files of data are to be transmitted. To make data communications practical, ways will have to be developed that make the complexity linear in the number of bits sent, i.e., $O(K_b)$. This chapter will continue to look at M-ary signalling schemes as they often are used in practice (in conjunction with other complexity reduction techniques) to achieve either higher performance or higher bandwidth efficiency.

8.2.2 Performance Analysis

The performance analysis of an M-ary demodulator is much more complicated than that of the binary detector but many of the ideas are similar. Two important quantities should be related at this point: the average symbol energy is denoted E_s and this is given as

$$E_s = \sum_{i=0}^{M-1} E_i \pi_i. \tag{8.23}$$

The case of equal priors reduces to

$$E_s = \frac{1}{M}\sum_{i=0}^{M-1} E_i. \tag{8.24}$$

The average energy per bit is then given as

$$E_b = \frac{E_s}{K_b}. \tag{8.25}$$

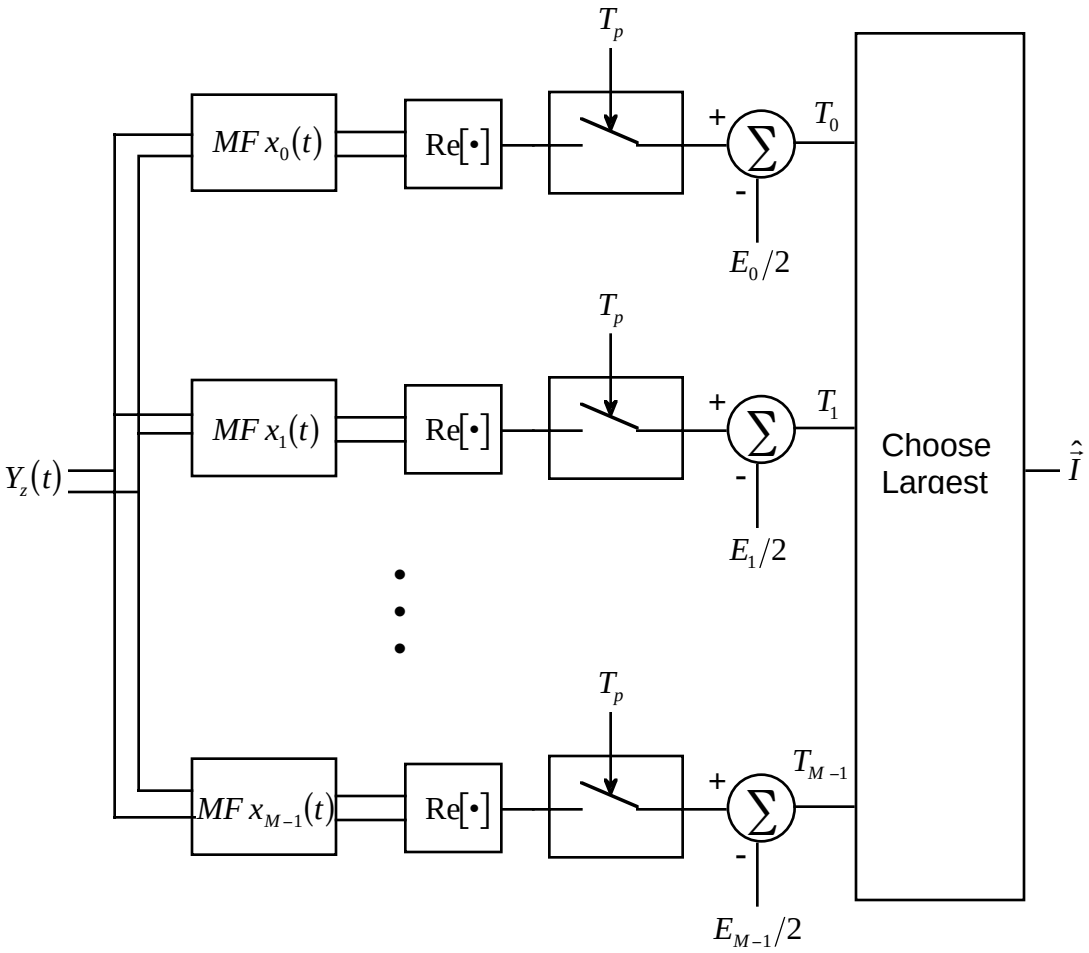


Figure 8.2: The block diagram for the MLWD demodulator.

Example 8.4: Returning again to the pulse width modulation example we have

$$E_b = \frac{E_0 + E_1 + E_2 + E_3}{8} = \left(\frac{8E_b}{80} + \frac{16E_b}{80} + \frac{24E_b}{80} + \frac{32E_b}{80} \right). \tag{8.26}$$

We again start the process of understanding the performance by characterizing the likelihood metrics.

Likelihood Metrics

The likelihood metrics, $T_i, i \in \{0, \ldots, M-1\}$ are matched filter outputs with an energy correction term. Conditioned on the transmitted signal the likelihood metrics are Gaussian random variables. For instance, if $x_j(t)$ is assumed transmitted then denote the i^{th} conditional likelihood metric as

$$\begin{aligned} T_{i|j} &= \Re\left[\int_{-\infty}^{\infty} Y_z(t) x_i^*(t) dt\right] - \frac{E_i}{2} \\ &= \Re\left[\int_{-\infty}^{\infty} x_j(t) x_i^*(t) dt\right] + N_I^{(i)} - \frac{E_i}{2} \end{aligned} \tag{8.27}$$

where $N_I^{(i)}$ is a real zero mean Gaussian random variable given as

$$N_I^{(i)} = \Re\left[\int_{-\infty}^{\infty} W_z(t) x_i^*(t) dt\right] \tag{8.28}$$

with

$$\text{var}\left(N_I^{(i)}\right) = \frac{E_i N_0}{2}. \tag{8.29}$$

It should be noted that

$$\begin{aligned} T_{i|j} &= \Re\left[\rho_{ji}\sqrt{E_i E_j}\right] - \frac{E_i}{2} + N_I^{(i)} \\ T_{j|j} &= \frac{E_j}{2} + N_I^{(j)}. \end{aligned} \tag{8.30}$$

Finally if two conditional likelihood metrics, $T_{i|j}$ and $T_{k|j}$ are considered jointly, they will be jointly Gaussian random variables with a correlation coefficient of (see Problem 8.14)

$$\rho = \frac{E\left[N_I^{(i)} N_I^{(k)}\right]}{\frac{N_0}{2}\sqrt{E_i E_k}} = \Re\left[\rho_{ik}\right]. \tag{8.31}$$

This correlation between the matched filer outputs will be explored in more depth in the homework

Probability of Word Error

The probability of making a demodulation error is denoted the word error probability, $P_W(E)$. This can be written via total probability as

$$P_W(E) = \sum_{j=0}^{M-1} P\left(\hat{\vec{I}} \neq j \,\middle|\, \vec{I} = j\right) \pi_j \tag{8.32}$$

When the priors are unequal the nonlinear nature of MAPWD in (8.14) makes the computation of $P\left(\hat{\vec{I}} \neq j \left| \vec{I} = j \right.\right)$ very difficult. Not many results exist in the literature discussing $P_W(E)$ calculations for MAPWD.

The case of equal priors has a demodulator with a more exploitable structure. For MLWD the word error probability is

$$P_W(E) = \frac{1}{M} \sum_{j=0}^{M-1} P\left(\hat{\vec{I}} \neq j \left| \vec{I} = j \right.\right). \tag{8.33}$$

The conditional probability of word error is given as

$$\begin{aligned} P\left(\hat{\vec{I}} \neq j \left| \vec{I} = j \right.\right) &= P\left(\max_{j \neq i} T_i > T_j \left| \vec{I} = j \right.\right) \\ &= P\left(\max_{j \neq i} T_{i|j} > T_{j|j}\right) \\ &= P\left(\bigcup_{\substack{i=0 \\ j \neq i}}^{M-1} \left\{T_{i|j} > T_{j|j}\right\}\right). \end{aligned} \tag{8.34}$$

In general the probability in (8.34) is quite tough to compute. While each of the $T_{i|j},\ i = 1, M$ is a Gaussian random variable, computing the distribution of the maximum of correlated Gaussian random variables in nontrivial. For a significant number of modulations with small M the probability in (8.34) is computable with a reasonable amount of work on a computer. Some of these examples will be explored in the sequel and in a homework problem. For a general modulation and large M the problem is much tougher. One special case where a general result is available is the case of an orthogonal modulation, i.e., $\rho_{ij} = 0,\ \forall i \neq j$. This is explored in the homework.

8.2.3 Union Bound

A common upper bound on the probability in (8.34) has found great utility in the analysis of communication systems. This so called union bound is simply stated in general form as

$$P\left(\bigcup_{i=1}^{N} A_i\right) \leq \sum_{i=1}^{N} P(A_i) \tag{8.35}$$

where A_i are arbitrary events. The idea of the union bound is simply understood if one considers the Venn diagram of Fig 8.3. Consider the probability of each event, A_i to be proportional to the area in the Venn diagram. The union bound given in (8.35) counts the area in $A_1 \bigcap A_2$, $A_1 \bigcap A_3$, and $A_2 \bigcap A_3$ twice and the area of $A_1 \bigcap A_2 \bigcap A_3$ three times. The union bound is satisfied with equality if the events are disjoint.

The union bound can now be used to upperbound the word error probability in an M-ary optimum word demodulator. Using the union bound in (8.34) gives

$$\begin{aligned} P\left(\hat{\vec{I}} \neq j \left| \vec{I} = j \right.\right) &= P\left(\bigcup_{\substack{i=0 \\ j \neq i}}^{M-1} \left\{T_{i|j} > T_{j|j}\right\}\right) \\ &\leq \sum_{\substack{i=0 \\ j \neq i}}^{M-1} P\left(T_{i|j} > T_{j|j}\right). \end{aligned} \tag{8.36}$$

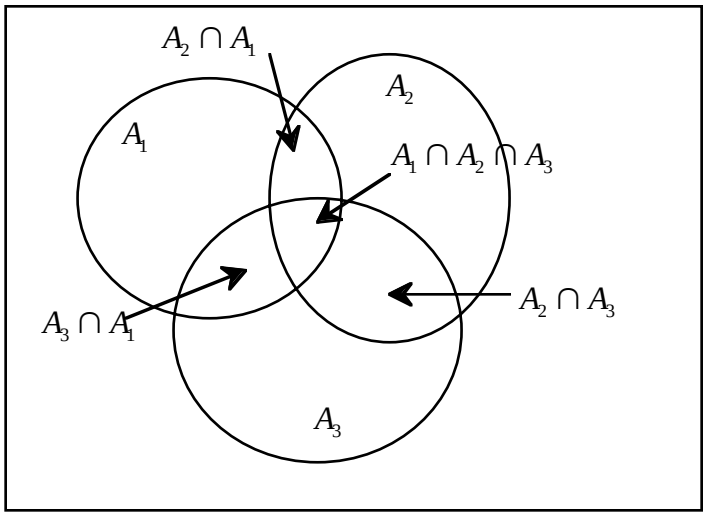


Figure 8.3: A Venn diagram to explain the union bound.

$P\left(T_{i|j} > T_{j|j}\right)$ is the probability that the i^{th} maximum likelihood metric for the transmitted signal is greater than the maximum likelihood metric corresponding to the actual transmitted signal. This probability is often denoted the pair–wise error probability (PWEP).

The PWEP has a form readily analyzed. Recall the form for the decision metrics are

$$\begin{aligned} T_{i|j} &= \Re\left[\rho_{ji}\sqrt{E_i E_j}\right] - \frac{E_i}{2} + N_I^{(i)} \\ T_{j|j} &= \frac{E_j}{2} + N_I^{(j)}. \end{aligned} \tag{8.37}$$

Consequently the PWEP is given as

$$\begin{aligned} P\left(T_{i|j} > T_{j|j}\right) &= P\left(\Re\left[\rho_{ji}\sqrt{E_i E_j}\right] - \frac{E_i}{2} + N_I^{(i)} > \frac{E_j}{2} + N_I^{(j)}\right) \\ &= P\left(N_{ij} > \frac{E_i + E_j}{2} - \Re\left[\rho_{ji}\sqrt{E_i E_j}\right]\right) \qquad (8.38) \\ &= P\left(N_{ij} > \frac{\Delta_E(i,j)}{2}\right) \qquad (8.39) \end{aligned}$$

where $N_{ij} = N_I^{(i)} - N_I^{(j)}$. Since N_{ij} is a zero mean Gaussian random variable the PWEP is simply expressed as the tail probability of a Gaussian random variable. Since

$$N_{ij} = \Re\left[\int_{-\infty}^{\infty} W_z(t)[x_i^*(t) - x_j^*(t)]dt\right] \tag{8.40}$$

the variance is given as

$$\text{var}\left(N_{ij}\right) = \frac{N_0}{2}\left[E_i + E_j - 2\Re\left[\rho_{ji}\sqrt{E_i E_j}\right]\right] = \frac{N_0 \Delta_E(i,j)}{2}. \tag{8.41}$$

Using this variance in (8.38) gives

$$\begin{aligned} P\left(T_{i|j} > T_{j|j}\right) &= \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_i + E_j - 2\Re\left[\rho_{ji}\sqrt{E_i E_j}\right]}{4N_0}}\right) \\ &= \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\Delta_E(i,j)}{4N_0}}\right) \end{aligned} \tag{8.42}$$

where again

$$\Delta_E(i,j) = \int_{-\infty}^{\infty} |x_j(t) - x_i(t)|^2 \, dt \tag{8.43}$$

is the squared Euclidean distance between $x_i(t)$ and $x_j(t)$. It should be noted that the PWEP is exactly the same as the binary error probability given in the previous chapter.

The PWEP can be combined with (8.36) and (8.33) to give a union bound to the overall average word error probability. The conditional probability of error given that $x_j(t)$ was transmitted is now upperbounded by

$$P\left(\hat{\vec{I}} \neq j \,\middle|\, \vec{I} = j\right) \leq \sum_{i \neq j} \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\Delta_E(i,j)}{4N_0}}\right). \tag{8.44}$$

Consequently an upperbound to the word error probability is given as

$$P_W(E) \leq P_{WUB}(E) = \sum_{j=0}^{M-1} \sum_{i \neq j} \frac{1}{2M}\text{erfc}\left(\sqrt{\frac{\Delta_E(i,j)}{4N_0}}\right) \tag{8.45}$$

and this upperbound is often referred to as the union bound in the literature. It is important to realize that this sum is over all possible transmitted signal and all possible decoding errors for each of the possible transmitted signals.

Example 8.5: We can enumerate all the Euclidean distances for the pulse width modulation example in a table

$\Delta_E(i,j)$

	$\vec{I} = i$			
$\vec{I} = j$	0	1	2	3
0	0	$\frac{8E_b}{10}$	$\frac{16E_b}{10}$	$\frac{24E_b}{10}$
1	$\frac{8E_b}{10}$	0	$\frac{8E_b}{10}$	$\frac{16E_b}{10}$
2	$\frac{16E_b}{10}$	$\frac{8E_b}{10}$	0	$\frac{8E_b}{10}$
3	$\frac{24E_b}{10}$	$\frac{16E_b}{10}$	$\frac{8E_b}{10}$	0

The important points resulting from the union bound are

1. Squared Euclidean distance is again an important performance criterion.
2. The union bound typically has the form

$$P_W(E) \leq P_{WUB}(E) = \sum_{k=1}^{N} \frac{A_d(k)}{2M}\text{erfc}\left(\sqrt{\frac{\Delta_E(k)}{4N_0}}\right) \tag{8.46}$$

where the sum enumerates the $N < M(M-1)/2$ possible different squared Euclidean distances, $\Delta_E(k),\ k = 1, N$. Note that $A_d(k)$ is the number of signal pairs having a squared Euclidean distance between them of $\Delta_E(k)$ and that $\sum_{k=1}^{N} A_d(k) = M(M-1)$. The combination of the enumeration of the squared Euclidean distance, $\Delta_E(k)$, and the number of signal pairs $A_d(k)$ for each distance is often denoted the squared Euclidean distance spectrum weight. All sets of pairs $\{\Delta_E(k), A_d(k)\},\ k = 1, N$ is often denoted the squared Euclidean distance spectrum.

Example 8.6: For pulse width modulation we have $N = 3$ with $\Delta_E(1) = \frac{8E_b}{10}$, $\Delta_E(2) = \frac{16E_b}{10}$, and $\Delta_E(3) = \frac{24E_b}{10}$ and $A_d(1) = 6$, $A_d(2) = 4$, and $A_d(3) = 2$. Consequently the union bound is expressed as

$$P_W(E) \leq P_{WUB}(E) = \frac{3}{4}\text{erfc}\left(\sqrt{\frac{8E_b}{40N_0}}\right) + \frac{1}{2}\text{erfc}\left(\sqrt{\frac{16E_b}{40N_0}}\right) + \frac{1}{4}\text{erfc}\left(\sqrt{\frac{24E_b}{40N_0}}\right). \quad (8.47)$$

3. The bound is often dominated by the minimum squared Euclidean distance of the signal set, $\Delta_E(\min)$, where

$$\Delta_E(\min) = \min_{\substack{j\in\{0,\ldots,M-1\}\\ i\in\{0,\ldots,M-1\}\\ i\neq j}} \Delta_E(i,j). \quad (8.48)$$

Since $\text{erfc}\,(x) \sim \exp[-x^2]$ it is usually the case that

$$\text{erfc}\left(\sqrt{\frac{\Delta_E(\min)}{4N_0}}\right) \gg \text{erfc}\left(\sqrt{\frac{\Delta_E(i,j)}{4N_0}}\right) \quad \forall \Delta_E(i,j) \neq \Delta_E(\min). \quad (8.49)$$

4. The union bound is usually tight at high SNR. The bound over counts the cases when more than one maximum likelihood metric for a nontransmitted word is greater than the maximum likelihood metric for the transmitted word. At high SNR the probability of two or more metrics being greater than the true metric is very small. Consequently the overbounding probability is very small.

5. Using items 3. and 4. allows one to deduce that

$$P_W(E) \approx \frac{A_d(\min)}{2M}\text{erfc}\left(\sqrt{\frac{\Delta_E(\min)}{4N_0}}\right) \quad (8.50)$$

is a good approximation to the true error performance at high SNR.

6. A concept that arises often in performance analysis is the concept of geometric uniformity [For91]. A geometrically uniform signal set is one in which the conditional distance spectrum of each of the possible transmitted signals is the same. The important characteristic of a geometrically uniform signal set is the number of terms in the union bound reduces from $M(M-1)$ to $M-1$. Note that pulse width modulation considered in Example 8.4 is not a geometrically uniform signal set, but $x_0(t)$ and $x_3(t)$ have the same conditional distance spectrum as do $x_1(t)$ and $x_2(t)$.

8.2.4 Signal Design

Signal design for M-ary modulation is typically best optimized with a max-min approach. Since the minimum squared Euclidean distance will dominate performance of an optimum demodulator the best signal design will happen when the minimum distance between all pairs of the M signals is maximized.

The spectral characteristics of an M-ary signal set can be characterized with a straightforward extension of the techniques used for binary signals. Recall from (6.5) that the average energy spectrum per bit for a transmitted signal, $X_z(t)$, where K_b bits are transmitted is

$$D_{x_z}(f) = \frac{E\left[G_{x_z}(f)\right]}{K_b}. \tag{8.51}$$

The average here again is over random data bits that are being transmitted.

8.3 Optimum Bit Demodulation

The demodulation structure that minimizes the bit error probability is different than the structure which minimizes word error probability. In the MAPWD or the MLWD the optimality criterion was word error probability. It seems obvious that often minimizing bit error probability will be a more important design goal than minimizing word error probability. A word error regardless of the number of bit errors is treated the same in a MLWD or a MAPWD. Consequently if the bit error rate is more important than the word error rate then a new demodulation structure is necessary. The minimum bit error rate demodulator has some similarities to the minimum word error rate but will be more complex to implement.

8.3.1 Optimum Bit Demodulation Receivers

First some notation will be introduced to aid with the discussion of the optimum bit detectors. Again assume that K_b bits, $\vec{I} = [I(1)\ I(2)\cdots I(K_b)]$, are being transmitted and that M signals, $x_i(t),\ i \in \{0,\ldots,M-1\}$, are being used to represent the possible word values that are to be transmitted. Without loss of generality assume that the interest is in minimizing the bit error rate for the k^{th} bit, $I(k)$. We know from the previous chapter that the optimum demodulator, the MAPBD, is

$$P\left(I(k)=1\,|y_z(t)\right) \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} P\left(I(k)=0\,|y_z(t)\right). \tag{8.52}$$

In words, the optimum demodulator computes the a posteriori probability (APP) of the bit values and makes the bit decision based on which APP is larger. Recall we denoted this architecture the MAPBD. Unfortunately there is no simple way to express the APP, $P\left(I(k)=m\,|y_z(t)\right)$, since the transmitted signal is not conditionally deterministic with knowledge of $I(k)$. The transmitted signal *is* conditionally deterministic if all the other bits are fixed and this idea can be combined with total probability to give a method to compute the minimum bit error rate demodulator. Define $\vec{I}^{(-)}(k) = [I(1)\cdots I(k-1)\ I(k+1)\cdots I(K_b)]$. $\vec{I}^{(-)}(k)$ represents all the other bits besides the one for which an optimum bit decision is to be made. For simplicity in notation let the values of $\vec{I}^{(-)}(k)$ be enumerated by the integers $n = 0, M/2-1$ and using total probability then the APP can be expressed as

$$\begin{aligned} P\left(I(k)=m\,|y_z(t)\right) &= \sum_{n=0}^{M/2-1} P\left(I(k)=m, \vec{I}^{(-)}(k)=n\,|y_z(t)\right) \\ &= \sum_{n=0}^{M/2-1} P\left(\vec{I}=\{m,n\}\,|y_z(t)\right) \end{aligned} \tag{8.53}$$

where $\vec{I}=\{m,n\}$ denotes the data word corresponding to $I(k)=m$ and $\vec{I}^{(-)}(k)=n$. The essential idea in MAPBD for M-ary signaling is that each APP of data words containing $I(k)=m$ is computed and

then total probability is used to obtain the marginal posterior probability of $I(k) = m$. This average is over the 2^{K_b-1} possible values of the $\vec{I}^{(-)}(k)$.

Example 8.7: Consider the case of $K_b = 3$ and $k = 2$ where $I^{(-)}(2) = [I(1)\ I(3)]$. The bits and the words are enumerated as

$I(1)$	$I(2)$	$I(3)$	i	m	n
0	0	0	0	0	0
0	0	1	1	0	1
0	1	0	2	1	0
0	1	1	3	1	1
1	0	0	4	0	2
1	0	1	5	0	3
1	1	0	6	1	2
1	1	1	7	1	3

Our earlier discussion demonstrated that $P\left(\vec{I} = i\,|y_z(t)\right)$ has a computable form. Recall from (8.18) that the APP for $\vec{I} = i$ is

$$CP\left(\vec{I} = i\,|y_z(t)\right) = \exp\left[\frac{2}{N_0}\Re\left[V_i(T_p)\right] - \frac{E_i}{N_0}\right]\pi_i \tag{8.54}$$

where C is a constant. This is exactly the decision metric that is computed in the MAPWD. Consequently the MAPBD simply computes the decision metrics in the MAPWD, the word APPs, and performs an average over the data words as

$$\begin{aligned} P\left(I(k) = m\,|y_z(t)\right) &= \sum_{n=0}^{M/2-1} \exp\left[\frac{2}{N_0}\Re\left[V_{\{m,n\}}(T_p)\right] - \frac{E_{\{m,n\}}}{N_0}\right]\pi_{\{m,n\}} \\ &= \sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{m,n\}}}{N_0}\right]\pi_{\{m,n\}}. \end{aligned} \tag{8.55}$$

Note the optimum bit decision is

$$\sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{1,n\}}}{N_0}\right]\pi_{\{1,n\}} \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} \sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{0,n\}}}{N_0}\right]\pi_{\{0,n\}}. \tag{8.56}$$

In a similar fashion the MLBD for M-ary modulations has the form

$$\sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{1,n\}}}{N_0}\right] \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} \sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{0,n\}}}{N_0}\right]. \tag{8.57}$$

Denoting the average metric for $I(k) = i$ as

$$L_i(k) = \sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{i,n\}}}{N_0}\right] \tag{8.58}$$

then the MLBD has the form

$$L_1(k) \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} L_0(k). \tag{8.59}$$

It is worth noting that optimum bit demodulation algorithms must know the value of N_0. This value of the noise power is one additional level of knowledge needed for MLBD detection that is not needed for MLWD detection.

Example 8.8: Consider the case of $K_b = 3$ and $k = 2$ where $I^{(-)}(2) = [I(1)\ I(3)]$. The average likelihood metrics for $I(2)$ are

$$L_0(2) = \exp\left[\frac{2T_0}{N_0}\right] + \exp\left[\frac{2T_1}{N_0}\right] + \exp\left[\frac{2T_4}{N_0}\right] + \exp\left[\frac{2T_5}{N_0}\right] \quad (8.60)$$

$$L_1(2) = \exp\left[\frac{2T_2}{N_0}\right] + \exp\left[\frac{2T_3}{N_0}\right] + \exp\left[\frac{2T_6}{N_0}\right] + \exp\left[\frac{2T_7}{N_0}\right] \quad (8.61)$$

8.3.2 Comparison with Optimum Word Demodulation

A pseudo-code comparison of MLBD and MLWD will illustrate the similarities and differences between the algorithms.

MLBD

- For $i \in \{0, \ldots, M-1\}$
  1. Compute T_i
  2. Compute $\exp(2T_i/N_0)$
- For $k \in \{1, \ldots, K_b\}$
  1. Identify $I^{(-)}(k)$
  2. Compute
$$L_0(k) = \sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{0,n\}}}{N_0}\right]$$
  3. Compute
$$L_1(k) = \sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{1,n\}}}{N_0}\right]$$
  4. Compute
$$\hat{I}(k) = \arg\max_{i=0,1} L_i(k)$$

MLWD

- For $i \in \{0, \ldots, M-1\}$
  1. Compute T_i
- $\hat{\vec{I}} = \arg\max_{i\in\{0,\ldots,M-1\}} T_i$

Recall the complexity of the MLWD is $O(M)$=$O(2^{K_b})$. The MLBD computes all the same metrics as the MLWD and then computes an average for each bit decision. Consequently the complexity of MLBD is $O(K_bM)$=$O(K_b2^{K_b})$. Since $K_b = \ln(M)$, the increase in the complexity for MLBD is not significant when M is large.

The performance of bit demodulation algorithms is in general hard to quantify due to the averaging of the word metrics. I am unaware of rigorous performance analysis for bit detection algorithms. Performance is usually assessed via a computer simulation except in special cases. One characteristic

that still holds is that the minimum Euclidean distance will still dominate the performance at high SNR. Several points are worth noting in conclusion:

1. The performance difference between MLWD and MLBD is small especially at high SNR.
2. MLWD is less complex than MLBD.

These two characteristics led to the MLWD dominating in practical implementations up until the late 1990s. Using the APPs that the MLBD (or MAPBD) computes can offer significant improvements in performance in suboptimum iterative demodulation schemes that were popularized by the appearance of Turbo codes [BG96, BGT93] and by the rediscovery of low density parity check codes [Gal62]. Because of the utility of APPs in iterative demodulation schemes we will continue to look at both ideas in parallel in the remainder of the text.

8.4 Examples

This chapter is concluded by considering two obvious and important examples of M-ary carrier modulated digital communication: M-ary frequency shift keying (MFSK) and M-ary phase shift keying (MPSK). In these examples it is assumed that $\pi_i = 1/M,\ i \in \{0, \ldots, M-1\}$.

8.4.1 M-ary FSK

MFSK modulation sends the word of information by transmitting a carrier pulse of one of M frequencies. This is an obvious simple signalling scheme and one used in many early modems. The signal set is given as

$$\begin{aligned} x_i(t) &= \sqrt{\frac{K_b E_b}{T_p}} \exp\left[j2\pi f_d(2i - M + 1)t\right] \qquad 0 \le t \le T_p \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{8.62}$$

where f_d is known as the frequency deviation. The frequency difference between adjacent frequency pulses in the signal set is $2f_d$. It is apparent that each waveform in a MFSK signal set has equal energy that has here been set to $E_s = K_b E_b$.

The matched filter impulse response is given as

$$h_i(t) = x_i^*(T_p - t). \tag{8.63}$$

Since MFSK is an equal energy signal set the energy correction term is not needed in the demodulator and the ML decision metrics are given as

$$T_i = \Re\left[\int_0^{T_p} Y_z(t) x_i^*(t) dt\right] = \sqrt{\frac{K_b E_b}{T_p}} \Re\left[\int_0^{T_p} Y_z(t) \exp\left[-j2\pi f_d(2i - M + 1)t\right] dt\right] \qquad i \in \{0, \ldots, M-1\}. \tag{8.64}$$

The MLWD demodulator is then given as

$$\hat{\vec{I}} = \arg \max_{i \in \{0, \ldots, M-1\}} T_i. \tag{8.65}$$

This optimum demodulator has M filters outputs to compute and then must select the signal corresponding to the largest.

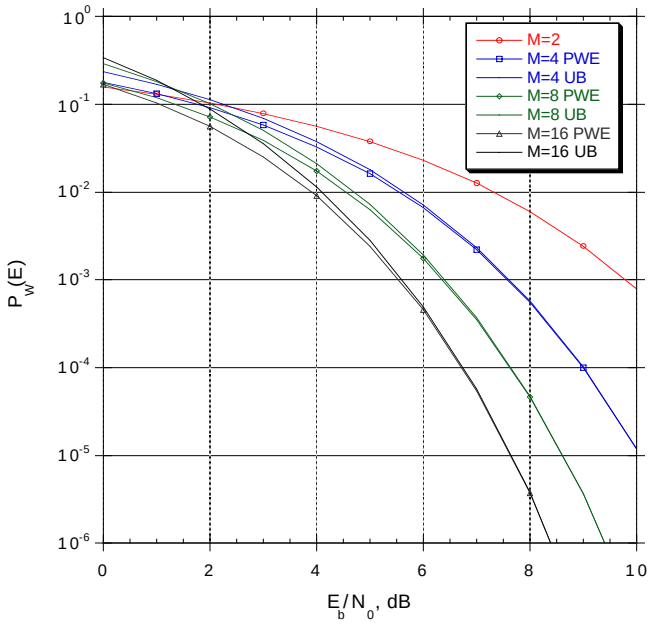


Figure 8.4: The true $P_W(E)$ and the union bound for orthogonal MFSK.

Performance of MFSK is a function of the frequency spacing, $2f_d$. Using the results from BFSK, i.e., (7.80), the pair-wise squared Euclidean distance is given as $\Delta_E(i,j) = 2K_bE_b\left(1 - \frac{\sin(4\pi f_d(i-j)T_p)}{4\pi f_d(i-j)T_p}\right)$. Consequently for MFSK the placement of the tones is a bit more delicate problem. The union bound is given as

$$P_{WUB}(E) = \sum_{n=1}^{M-1} \frac{M-n}{M} \text{erfc}\left(\sqrt{\frac{K_bE_b}{2N_0}\left(1 - \frac{\sin(4\pi f_d nT_p)}{4\pi f_d nT_p}\right)}\right). \tag{8.66}$$

The selection of the frequency spacing, f_d, becomes a tricky problem for $M > 2$ as one needs to balance all the terms in the union bound to optimize the performance. The MFSK signal set is not in general a geometrically uniform signal set.

For the special case of $\Re[\rho_{ij}] = 0$ both the $P_W(E)$ can be easily computed and a simple form for the union bound results. This is known as orthogonal MFSK and is achieved if $f_d = \frac{n}{4T_p}$ where n is a positive integer. The closed form probability of error expression for orthogonal MFSK is explored in the homework and plotted in Fig. 8.4 for $M = 2, 4, 8, 16$. The important thing to realize for MFSK is that for large enough E_b/N_0, the $P_W(E)$ is monotonically decreasing with M. In fact one can show that the error rate can be made arbitrarily small when $E_b/N_0 > \ln(2)$ [Sha48]. Consequently the performance of a digital communication system can be improved by transmitting more bits or having more waveforms to choose from. This result is also counter normal engineering intuition. Shortly we will show that this performance improvement is achieved only at the cost of an increase in the bandwidth of the signal. The union bound for orthogonal MFSK is

$$P_W(E) \leq \frac{M-1}{2} \text{erfc}\left(\sqrt{\frac{K_bE_b}{2N_0}}\right). \tag{8.67}$$

For $\Re[\rho_{ij}] = 0$ MFSK is a geometrically uniform signal set. Fig. 8.4 also plots the union bound for MFSK. It is obvious that the union bound converges to the true error probability at high SNR. For

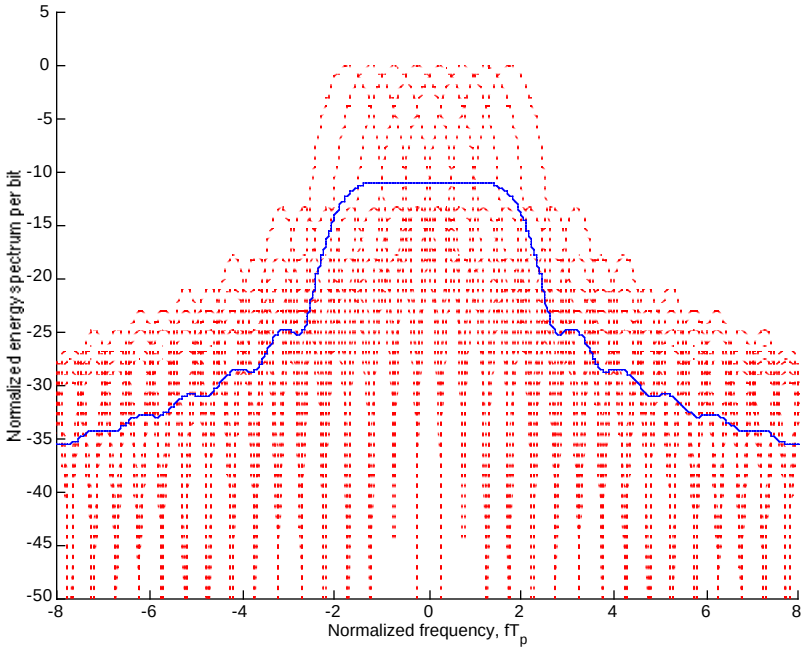


Figure 8.5: For 8FSK the $G_{x_i}(f)$ for $i \in \{0, \ldots, 7\}$ and $D_{x_z}(f)$.

the particular case of MFSK the bound is tight enough to be indiscernibly different on the graph for $E_b/N_0 > 6$dB. This characteristic for the union bound holds for most signal sets optimally demodulated in the presence of an AWGN.

The average energy spectrum per bit is again used to characterize the spectral efficiency. Recall the average energy spectral density per bit is given for M-ary modulations as

$$D_{x_z}(f) = \frac{1}{K_b} \sum_{i=0}^{M-1} \pi_i G_{x_i}(f). \tag{8.68}$$

Recall that the energy spectrum of the individual waveforms is given as

$$G_{x_i}(f) = K_b E_b T_p \left(\frac{\sin\left(\pi(f - f_d(2i - M + 1))T_p\right)}{\pi(f - f_d(2i - M + 1))T_p} \right)^2. \tag{8.69}$$

Recall that the minimum frequency separation needed to achieve an orthogonal modulation is $f_d T_p = 0.25$ and that by considering (8.68) and (8.69) it is obvious that the spectral content is growing proportional to $B_T = f_d(2^{K_b+1})$. An example of each of the individual energy spectrums and the average energy spectrum for 8FSK is plotted in Fig. 8.5. The transmission rate of MFSK is $W_b = K_b/T_p$. The spectral efficiency then is approximately $\eta_B = K_b/2^{K_b-1}$ and decreases with the number of bits transmitted or equivalently decreases with M. Conversely MFSK provides monotonically increasing performance with M. Consequently MFSK has found use in practice when lots of bandwidth is available and good performance is required.

The MLBD can be formed for the MFSK modulation. Recall the likelihood metrics must be processed to give

$$L_0(k) = \sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{0,n\}}}{N_0}\right] \qquad L_1(k) = \sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{1,n\}}}{N_0}\right]$$

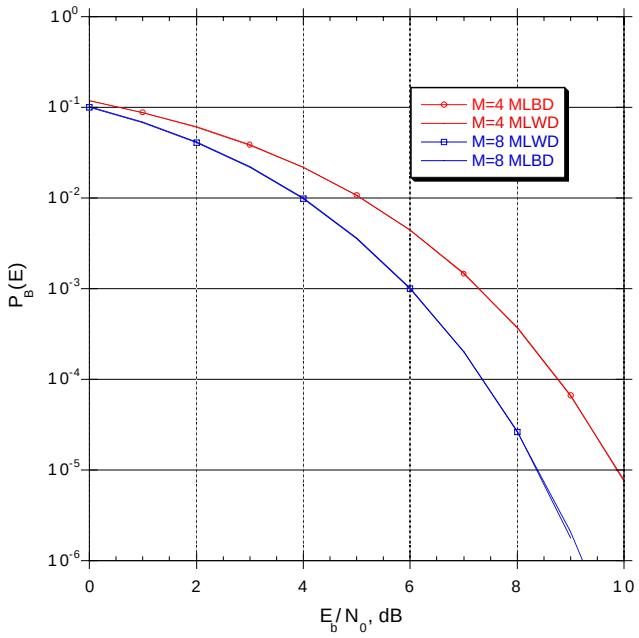


Figure 8.6: The $P_B(E)$ for MLWD and MLBD.

and the decision is made via

$$\hat{I}(k) = \arg \max_{i=0,1} L_i(k).$$

The MLBD does not simplify further for MFSK. Fig. 8.6 also contains the performance for the MLBD for $M = 4$. The performance is slightly better than the MLWD but not significantly. The majority of the improvement is due to the fact that each word error does not result in a bit error at every position. Consequently for orthogonal MFSK the bit error rate improvement for MLBD does not seem to justify the additional complexity.

The advantages of MFSK are summarized as

- Very simple to generate. Simply gate one of M oscillators on depending on the word to be sent.
- Performance (error rate) improves monotonically with K_b. This is counter intuition as normal intuition would expect performance degrading with increasing K_b.

The disadvantages of MFSK are summarized as

- The bandwidth increases exponentially with K_b, hence the spectral efficiency of MFSK decreases with K_b.
- Complexity increases exponentially with K_b.

Additional important points demonstrated by this example

- Union bound is simple to compute and is a tight bound at moderate to high SNR.
- MLBD has only slightly better bit error rate performance than the MLWD.

8.4.2 M-ary PSK

MPSK modulation sends the word of information by transmitting a carrier pulse of one of M phases. This is an obvious simple signalling scheme and one used in many modems. The first form of the signal set one might consider is given as

$$x_i(t) = \begin{cases} \sqrt{\frac{K_b E_b}{T_p}} \exp\left[j\frac{\pi(2i+1)}{M}\right] & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{8.70}$$

The phases have been chosen uniformly spaced around the unit circle. This methodology maximizes the minimum Euclidean distance. It is apparent that each waveform in an MPSK signal set has equal energy that has here been set to $E_s = K_b E_b$. Unfortunately if the phases are chosen corresponding to the word value, adjacent phases which are close in Euclidean distance will have many bits different in the two words. A technique mitigating the effects of word errors on the $P_B(E)$ by mapping of binary words into phases in a way to where adjacent phases are only different in one bit is known as Gray coding [Gra53]. An example of Gray coding for 4-ary PSK[1] is given as

$$\begin{aligned} \vec{I} = 0 = [0\ 0]^T \quad \theta_0 = \pi/4 \qquad \vec{I} = 1 = [1\ 0]^T \quad \theta_1 = 3\pi/4 \\ \vec{I} = 2 = [0\ 1]^T \quad \theta_2 = -\pi/4 \qquad \vec{I} = 3 = [1\ 1]^T \quad \theta_3 = -3\pi/4. \end{aligned} \tag{8.71}$$

This example will consider MLWD for the general MPSK signal set and MLBD for the Gray coded 4PSK given in (8.71).

Since MPSK is an equal energy signal set the energy correction term is not needed in the demodulator and the ML decision metrics are given as

$$\begin{aligned} T_i &= \Re\left[\int_0^{T_p} Y_z(t) x_i^*(t) dt\right] = \sqrt{\frac{K_b E_b}{T_p}} \Re\left[\int_0^{T_p} Y_z(t) \exp\left[-j\frac{\pi(2i+1)}{M}\right] dt\right] \\ &= \Re\left[\exp\left[-j\frac{\pi(2i+1)}{M}\right] \sqrt{\frac{K_b E_b}{T_p}} \int_0^{T_p} Y_z(t) dt\right] \\ &= \Re\left[\exp\left[-j\frac{\pi(2i+1)}{M}\right] Q\right] \qquad i \in \{0, \ldots, M-1\} \end{aligned} \tag{8.72}$$

where

$$Q = \sqrt{\frac{K_b E_b}{T_p}} \int_0^{T_p} Y_z(t) dt \tag{8.73}$$

is denoted the pulse shape matched filter output. The MLWD computes the output of one filter, Q, derotates this value by each of the possible transmitted phasors and picks the signal that gives the largest real value. This decision rule is equivalent to picking the transmitted phase that is closest to the phase of the pulse shape matched filter so that

$$\hat{\vec{I}} = \arg \max_{i \in \{0,\ldots,M-1\}} T_i = \arg \min_{i \in \{0,\ldots,M-1\}} \left| Q_p - \frac{\pi(2i+1)}{M} \right| \tag{8.74}$$

where $Q_p = \arg\{Q\}$.

Surprisingly this optimum demodulator has only one filter output to compute. This filter output is then processed to produce the decision metric. The word decision is then made on the basis of which

[1] 4-ary PSK modulation is often denoted quaternary PSK or QPSK.

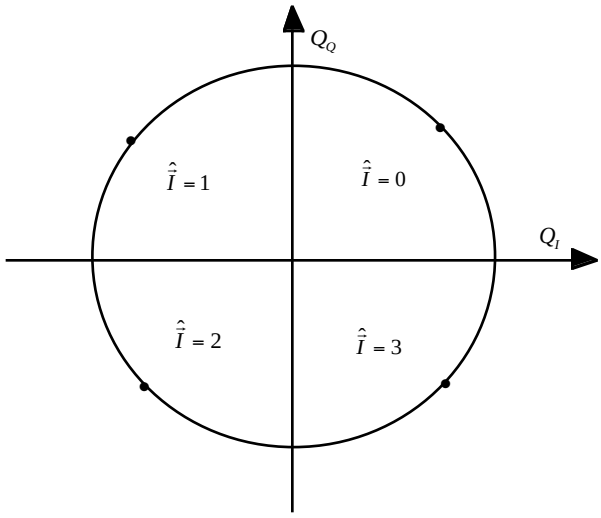


Figure 8.7: The decision regions for 4PSK.

of these metric is the largest. The decision rule reduces down to the establishing of decision regions in the complex plane for the pulse matched filter output Q. The decision regions for 4PSK are shown in Fig. 8.7. The fact that the M-ary demodulator only has one filter is since the transmitted signal set has the form

$$x_i(t) = D_i u(t) \tag{8.75}$$

where

$$u(t) = \begin{cases} \sqrt{\dfrac{K_b E_b}{T_p}} & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{8.76}$$

is the pulse shape. Modulations which have the form of (8.75) are often denoted linear modulations by communications engineers. Since linear modulations greatly reduce the demodulator complexity they will be explored in detail in the sequel.

Performance of MPSK is a function of the phase spacing, $\frac{2\pi}{M}$. The squared Euclidean distance is given as $\Delta_E(i,j) = 2K_bE_b\left(1-\cos\left(\frac{2\pi(i-j)}{M}\right)\right)$. Note since cosine is a periodic function, MPSK is a geometrically uniform signal set. The union bound for MPSK is given as

$$P_W(E) \le \frac{1}{2}\text{erfc}\left(\sqrt{\frac{K_bE_b}{N_0}}\right) + \sum_{i=1}^{M/2-1}\text{erfc}\left(\sqrt{\frac{K_bE_b}{2N_0}\left[1-\cos\left(\frac{2\pi i}{M}\right)\right]}\right). \tag{8.77}$$

The MFSK signal set was not in general a geometrically uniform signal set but here we see that the MPSK signal set is geometrically uniform when the phases of the modulation symbols are uniformly distributed around the unit circle.

It is interesting to note that for MPSK the union bound can actually be tightened compared to the general result given in (8.77). This tightening of the union bound is due to the following

Property 8.1 *$P(A \cup B \cup C) = P(A \cup B)$ if $C \subset A \cup B$.*

Since the MPSK signal set is geometrically uniform we can consider $x_0(t)$ to be the transmitted signal without loss of generality. The pair-wise error probability is then given as

$$\begin{aligned}\left\{T_{i|0} > T_{0|0}\right\} &= \left\{\Re\left[\exp\left[-j\frac{\pi(2i+1)}{M}\right]Q\right] > \Re\left[\exp\left[-j\frac{\pi}{M}\right]Q\right]\right\} \\ &= \left\{\left|Q_p - \frac{\pi(2i+1)}{M}\right| < \left|Q_p - \frac{\pi}{M}\right|\right\}. \end{aligned} \tag{8.78}$$

This reduces to

$$\left\{T_{i|0} > T_{0|0}\right\} = \left\{\frac{\pi(i+1)}{M} \leq Q_p \leq \frac{\pi(M+i+1)}{M}\right\}. \tag{8.79}$$

Given this form of the pair-wise error probability it can be seen that

$$\left\{T_{1|0} > T_{0|0}\right\} \cup \left\{T_{M-1|0} > T_{0|0}\right\} = \left\{\frac{2\pi}{M} \leq Q_p \leq 2\pi\right\}. \tag{8.80}$$

Consequently for any $i \in \{2, \ldots, M-2\}$

$$\left\{T_{i|0} > T_{0|0}\right\} \subset \left\{T_{1|0} > T_{0|0}\right\} \cup \left\{T_{M-1|0} > T_{0|0}\right\}. \tag{8.81}$$

This implies that a tighter union bound for MPSK than (8.77) is given as

$$P_W(E) \leq \operatorname{erfc}\left(\sqrt{\frac{K_b E_b}{2N_0}\left[1 - \cos\left(\frac{2\pi}{M}\right)\right]}\right). \tag{8.82}$$

This tightened union bound only includes the minimum squared Euclidean distance terms. For MPSK these two pair-wise error probability terms include all the possible ways an error can occur.

MPSK word error probability performance can be computed exactly. If we assume that $x_k(t)$ was transmitted, the conditional matched filter output is given as

$$Q = \exp\left[j\frac{\pi(2k+1)}{M}\right]K_b E_b + N_z \tag{8.83}$$

where N_z is a complex Gaussian noise with a variance of

$$\operatorname{var}(N_z) = N_0 K_b E_b. \tag{8.84}$$

Considering the decision regions and the form of the conditional matched filter output, the probability of word error can be computed. This idea is explored in the homework problems. The resulting probability of word error is plotted in Fig. 8.8 for M=2,4,8,16. In contrast to MFSK the probability of error increases with M for MPSK. The performance degrades about 4dB for each doubling of M. The union bound is also plotted in Fig. 8.8 for M=2,4,8,16. The union bound is again asymptotically tight with moderate SNR. This again demonstrates the utility of a union bound performance analysis.

The average energy spectrum per bit is again used to characterize the spectral efficiency of MPSK. Recall the average energy spectral density per bit is given for M-ary modulations as

$$D_{x_z}(f) = \frac{1}{K_b}\sum_{i=0}^{M-1}\pi_i G_{x_i}(f). \tag{8.85}$$

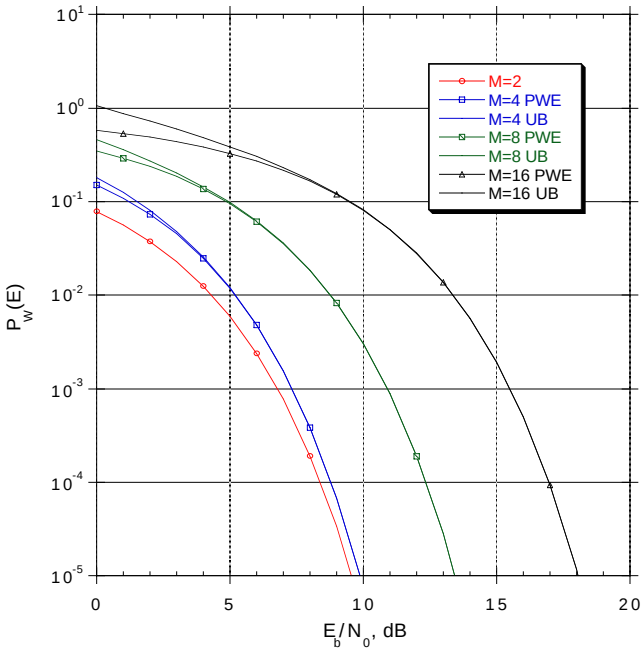


Figure 8.8: The true $P_W(E)$ and the union bound for MPSK.

Recall that the energy spectrum of the individual waveforms is given as

$$G_{x_i}(f) = E_b T_p \left| \exp\left[j\frac{\pi(2i+1)}{M} \right] \right|^2 \left(\frac{\sin(\pi f T_p)}{\pi f T_p} \right)^2 = E_b T_p \left(\frac{\sin(\pi f T_p)}{\pi f T_p} \right)^2 \tag{8.86}$$

The occupied bandwidth of MPSK, $B_T = 1/T_p$, does not increase with M while the bit rate does increase with M, $W_b = K_b/T_p$ to provide a spectral efficiency of $\eta_B = K_b$. Unfortunately the performance of MPSK modulation decreases with increasing M. Consequently MPSK modulations are of interest in practice when the available bandwidth is small and the SNR is large.

The MLBD for the Gray coded 4PSK yields some interesting results. Considering the decoder for $I(1)$ we have

$$L_1(1) = \exp\left[\frac{2T_1}{N_0} \right] + \exp\left[\frac{2T_3}{N_0} \right] \tag{8.87}$$

where due to (8.72) the ML decision metrics are given as

$$T_1 = \Re\left[\exp\left(-j\frac{3\pi}{4} \right) Q \right] = \frac{-1}{\sqrt{2}} Q_I + \frac{1}{\sqrt{2}} Q_Q \tag{8.88}$$

$$T_3 = \Re\left[\exp\left(j\frac{3\pi}{4} \right) Q \right] = \frac{-1}{\sqrt{2}} Q_I - \frac{1}{\sqrt{2}} Q_Q. \tag{8.89}$$

Combining these results give

$$L_1(1) = \exp\left[\frac{-\sqrt{2}Q_I}{N_0} \right] \left(\exp\left[\frac{\sqrt{2}Q_Q}{N_0} \right] + \exp\left[\frac{-\sqrt{2}Q_Q}{N_0} \right] \right) \tag{8.90}$$

Likewise

$$L_0(1) = \exp\left[\frac{\sqrt{2}Q_I}{N_0}\right]\left(\exp\left[\frac{\sqrt{2}Q_Q}{N_0}\right] + \exp\left[\frac{-\sqrt{2}Q_Q}{N_0}\right]\right) \tag{8.91}$$

The resulting binary decision rule for $I(1)$ is

$$\begin{aligned} L_1(1) &\underset{\hat{I}(1)=0}{\overset{\hat{I}(1)=1}{\gtrless}} L_0(1) \\ \exp\left[\frac{-\sqrt{2}Q_I}{N_0}\right] &\underset{\hat{I}(1)=0}{\overset{\hat{I}(1)=1}{\gtrless}} \exp\left[\frac{\sqrt{2}Q_I}{N_0}\right] \\ Q_I &\underset{\hat{I}(1)=1}{\overset{\hat{I}(1)=0}{\gtrless}} 0. \end{aligned} \tag{8.92}$$

Note this MLBD simplifies down to a very simple threshold test. In fact this threshold test is exactly the same as the threshold test for BPSK given in (7.92). Similarly the MLBD for $I(2)$ is

$$Q_Q \underset{\hat{I}(2)=1}{\overset{\hat{I}(2)=0}{\gtrless}} 0. \tag{8.93}$$

Now it can be seen that the MLBD for Gray coded QPSK is exactly two BPSK demodulators operating in parallel. This simplification of a general M-ary demodulator down to a set of binary demodulators is due to the bits being orthogonally modulated into the data symbols. This can be seen in that Gray coded QPSK is essentially two orthogonal BPSK signals (one on the I component and one on the Q component of the carrier). This orthogonality idea will come up in the sequel as methods are investigated for reducing the complexity of demodulators for large number of transmitted bits. The advantages of MPSK are summarized as

- Very simple to generate. Simply change the phase of an oscillator to one M values depending on the word to be sent.
- No increase in the bandwidth occupancy with increasing K_b. Consequently spectral efficiency increases with K_b.
- Demodulation complexity does not increase exponentially with K_b

The disadvantages of MPSK are summarized as

- Performance decreases monotonically with K_b.

Additional important points demonstrated by this example

- The most general union bound is simple to compute and is a tight bound at moderate to high SNR.
- The demodulator structure of MPSK allows a tighter union bound to be computed by considering the overlap in the decision regions.
- MLBD can have a greatly simplified form in some special cases.

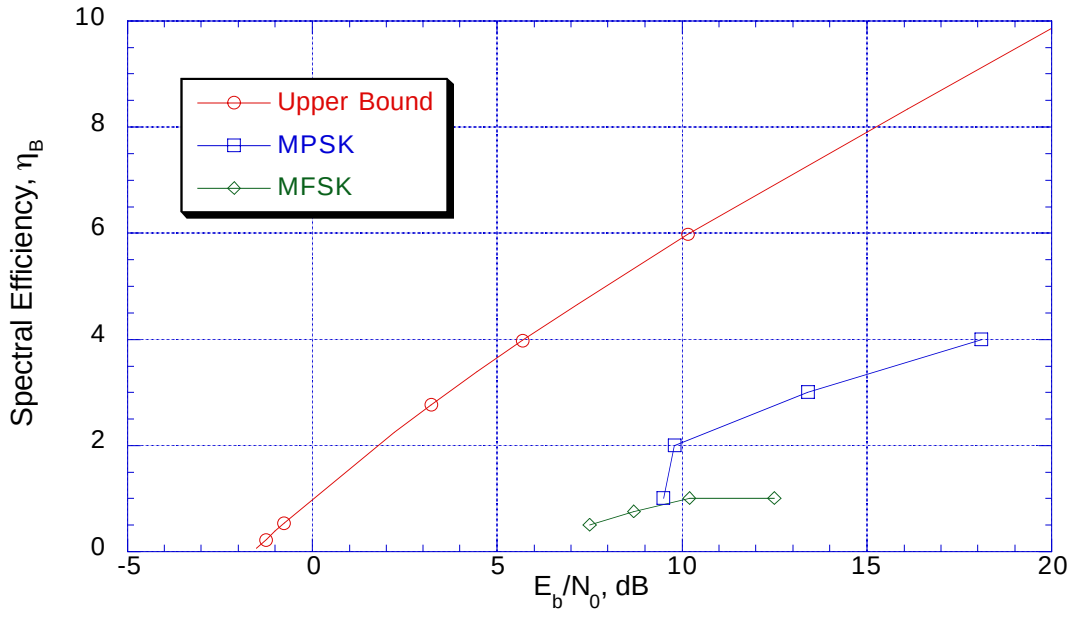


Figure 8.9: A comparison of the spectral efficiency of MPSK and MFSK with the upper bound.

8.4.3 Discussion

This section introduced two example modulations to transmit K_b bits of information: MFSK and MPSK. MPSK has the advantage in being able to supply an increasing spectral efficiency with K_b at the cost of requiring more E_b/N_0 to achieve the same performance. MFSK can provide improved performance with K_b but at a cost of a loss of spectral efficiency. Also the decoding complexity of MPSK is significantly less than the decoding complexity of MFSK. As a final point it is worth comparing the spectral efficiency performance of these two modulations with the upper bounds provided by information theory (see Section 6.3). As before we will denote reliable communication as being an error rate of 10^{-5}. The operating points of MFSK and MPSK and the upper bound on the possible performance are plotted in Fig. 8.9. It is clear from this graph that different modulations give us a different set of points in a performance versus spectral efficiency tradeoff. Also the two examples we considered in this Chapter have a performance much lower than the upperbound provided by information theory. This is still not too disturbing as we have lots of digital communication theory to explore.

8.5 Homework Problems

Problem 8.1. Find the probability of word error in MLWD for an arbitrary M for the case of equal energy orthogonal signaling, $\Re[\rho_{ij}] = 0,\ \forall i \neq j$. *Hint*: Condition on a value of the correct likelihood metric and find the conditional probability of word error (a distribution function of a max(•) random variable). Next average over the density function of correct likelihood metric. Compute numeric values and plot for M=4,8,16 for values of E_b/N_0=0–10dB. The first derivation of this result is available in [Kot60] but it is worth it to try to obtain the solution yourself before checking the reference. What value of f_d achieves orthogonality for the MFSK example considered in this chapter for a given T_p.

Problem 8.2. Show that if you design an equicorrelated ($\Re[\rho_{ij}] = \rho \leq 0,\ \forall i \neq j$), equal energy M-ary

signaling scheme then

$$\rho \geq -\frac{1}{M-1}.$$

Interpret what this means as M gets large. *Hint*: Consider a composite signal that is a sum of all M signals.

Problem 8.3. Recall from Chapter 5 that the phase error between the signal phase and the measured phase, $E_P = Y_P - x_P$, of the complex envelope of a carrier modulated signal in bandpass Gaussian noise, $Y_z = x_z + N_z$, is given as

$$f_{E_P}(e_p) = \begin{cases} \dfrac{\exp[-P]}{2\pi} + \sqrt{\dfrac{P}{4\pi}}\cos(e_p)\exp\left[-P\sin^2(e_p)\right]\left(1+\operatorname{erf}\left(\sqrt{P}\cos(e_p)\right)\right) & -\pi \leq e_p \leq \pi \\ 0 & \text{elsewhere} \end{cases} \tag{8.94}$$

where $P = \frac{x_A^2}{\mathrm{var}(N_z)}$.

a) Plot this PDF for $P = 10$dB.

b) Use this result to compute the word error probability for M-ary PSK. Compute numeric values and plot for M=4,8,16 for values of E_b/N_0=0–20dB.

Problem 8.4. A simple 4-ary modulation is given as

$$x_i(t) = d_i(1)u(t) + d_i(2)u(t-T)$$

where $u(t)$ is a pulse shape with energy E_u having support on $[0, T_u]$ and $T_u > T$, $i = 2 * I(2) + I(1)$, and $d_i(l) = (-1)^{I(l)}$. This signal is to be detected in the presence of an AWGN with one-sided spectral density of N_0 and the words are a priori equally likely, $\pi_i = 0.25$ $i \in \{0, \ldots, 3\}$.

a) What is the length of the transmission, T_p?

b) Find and detail out the optimum word error demodulator as a function of $V_u(T)$.

c) Find the union bound to the optimum word error demodulator probability of word decision error.

d) Define the following matched filter outputs

$$Q(1) = \int_0^{T_p} y_z(t)u^*(t)dt \qquad Q(2) = \int_0^{T_p} y_z(t+T)u^*(t)dt.$$

Find a simple form for the optimum demodulators given above as a function of $Q(1)$ and $Q(2)$. $Q(1)$ and $Q(2)$ are sufficient statistics for optimal detection.

e) Find and detail out the optimum bit error probability demodulator.

f) If $Q(1) = 0.5$, $Q(2) = -0.1$, $N_0 = 0.3$, $V_u(T) = 0.1$, and $E_u = 1$ what is the optimum word and bit decisions?

f) What conditions must exist such that simple decisions can be made on the value of $I(k)$ based solely on $Q(k)$ without loss in optimality?

Problem 8.5. Show via the union bound that M-ary FSK signaling can have a $P_W(E)$ arbitrarily small by a proper selection of K_b as long as $E_b/N_0 > 2\ln(2)$. Actually this statement is true for $E_b/N_0 > \ln(2)$ [Sha48] but the result is harder to prove (but worth an attempt by the interested student!)

Problem 8.6. 8-ary phase shift keying (8PSK) has an received signal of the form

$$Y_z(t) = x_i(t) + W_z(t) = d_i u(t) + W_z(t) \tag{8.95}$$

where $W_z(t)$ is a complex AWGN and if $\vec{I} = i$ then $d_i = \exp\left(\frac{j\pi i}{4}\right) \quad i \in \{0, \ldots, 7\}$. Assume $\pi_i = \frac{1}{8} \quad i \in \{0, \ldots, 7\}$.

a) Find the minimum probability of word error demodulator.

b) Find and plot the tightest union bound to the probability of word error for the demodulator found in a) for E_b/N_0=0-13dB.

c) Gray coding assigns bit patterns to M-ary signal points in such a way that adjacent signals (in Euclidean space) only have one bit different. Find a Gray code mapping for this 8PSK signal set.

d) Show that with the demodulator in a) and the bit to symbol mapping derived in c) that the resulting BEP is

$$P_B(E) = \frac{1}{3}\left[\text{erfc}\left(\sqrt{\frac{3E_b}{N_0}}\sin\left(\frac{\pi}{8}\right)\right) + \text{erfc}\left(\sqrt{\frac{3E_b}{N_0}}\cos\left(\frac{\pi}{8}\right)\right)\left[\frac{1}{2} + \frac{1}{2}\text{erf}\left(\sqrt{\frac{3E_b}{N_0}}\sin\left(\frac{\pi}{8}\right)\right)\right]\right] \tag{8.96}$$

e) Given the mapping in c) find the minimum probability of bit error demodulator. (Note any bit of the three will do.)

f) Run a computer simulation of the MLBD performance and plot the bit error probability for E_b/N_0=0-13dB along with performance of the MLWD given in (8.96).

Problem 8.7. Consider a 4-ary communication waveform that achieves $W_b = 2/T_p$ having the form shown in Fig. 8.10. Assume all words have equal priors and the corrupting noise is an AWGN with a one-sided spectral density of N_0.

a) Identify the MLWD structure.

b) Give the union bound for the performance.

c) Identify the 3dB bandwidth of the transmitted waveform and the spectral efficiency.

A proposed approach to increase the throughput of digital communications waveforms like that shown in Fig. 8.10 is variable phase shift keying (VPSK) [Wal97]. In variable phase shift keying the transition time between different voltage levels can be modulated to provide more possible waveforms. For example Fig. 8.11 shows a waveform with three possible transition times.

d) Design a VPSK signal set based on Fig. 8.10 that has $W_b = 3/T_p$. *Hint: You need eight waveforms and only two of four waveforms in Fig. 8.10 have transitions.*

e) Optimize performance of the signal set proposed in d) as a function of the shift parameter, τ_s and compare it to the original waveform.

f) Compute the spectral efficiency (using 3dB bandwidth again) for this case of optimal performance.

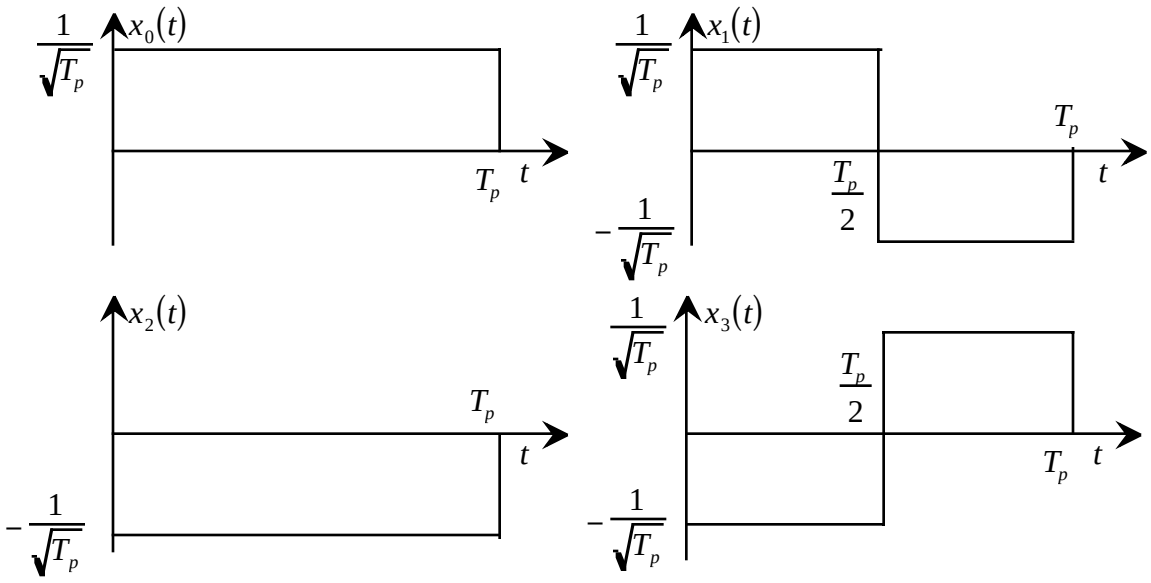


Figure 8.10: A 4-ary modulation.

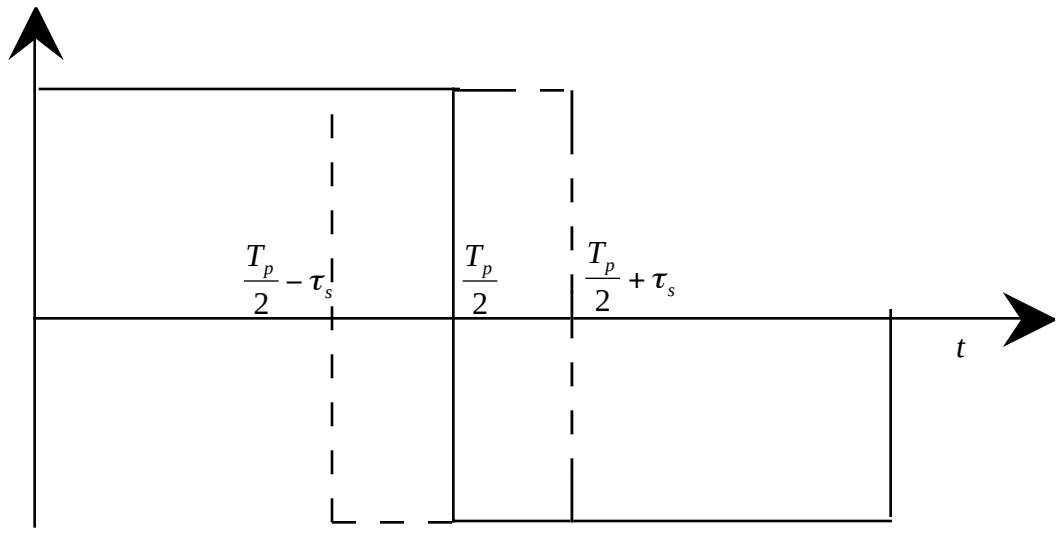


Figure 8.11: A waveform from Fig. 8.10 where the transition time is modulated.

Problem 8.8. In wireless communications it is often useful to observe a received waveform on more than one receiver antenna to improve the reliability of the decision. A situation like this is depicted in Fig. 8.12 for $L_r = 2$ antennas where the received signal at the j^{th} antenna is given as

$$Y_j(t) = c_j X_z(t) + W_j(t) \tag{8.97}$$

where C_j is a complex constant that represents j^{th} channel distortion and $W_j(t)$ is a white Gaussian noise with $R_{W_j}(\tau) = N_0\delta(\tau)$. Assume that $W_1(t)$ is independent of $W_2(t)$. Assume that c_1 and c_2 are known at the demodulator and that $\pi_0 = 0.5$.

a) Optimal decisions are based on APPs, i.e., $P(I = m|y(t))$ Bayes rule tells us that

$$P(I = m|y) = \frac{f_Y(y|I = m)\,\pi_m}{f_Y(y)}. \tag{8.98}$$

The results from detection theory tells us that

$$\frac{f_Y(y|I = m)}{f_Y(y)} = C\exp[2T_m/N_0] \tag{8.99}$$

How would the computation of the APP change if two observations with independent noises are obtained.

b) Using the results from a) for a given $x_0(t)$ and $x_1(t)$ find the minimum probability of error receiver obtained by observing $Y_j(t) = y_j(t) \quad j = 1, 2$. Simplify as much as possible. The resulting receiver structure is known as maximal ratio combining [Bre59] and essentially gives a matched filter in space and time.

c) Compute the error rate performance of the optimum receiver as a function of c_1, c_2, and $\Delta_E(0, 1)$

d) Knowing the values of c_1 and c_2 what is the signaling scheme that would minimize the bit error rate for a fixed transmitted energy per bit.

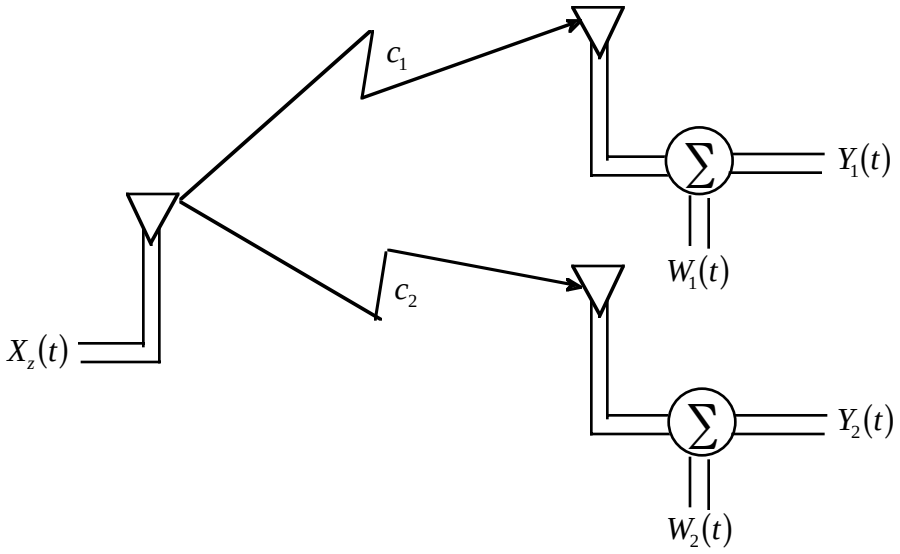


Figure 8.12: Observing a signal on multiple antennas.

Problem 8.9. In this problem we examine a binary communication system where three decisions are possible, $\hat{I}=1, \hat{I}=0, \hat{I}=E$ where E represents a no decision (erasure). The two possible baseband transmitted signals are $x_0(t)$ and $x_1(t)=-x_0(t)$ with $\pi_0=\pi_1=0.5$. Recall the optimum binary demodulator is of the form

$$P\left(I=1\,|v_I\right) \underset{\hat{I}=0}{\overset{\hat{I}=1}{\gtrless}} P\left(I=0\,|v_I\right) \tag{8.100}$$

where v_I is the output of the matched filter. The generalization considered in this problem is

$$\begin{array}{ll} \hat{I}=1 & P\left(I=1\,|v_I\right)>0.75 \\ \hat{I}=E & 0.25<P\left(I=1\,|v_I\right)\leq 0.75 \\ \hat{I}=0 & P\left(I=1\,|v_I\right)\leq 0.25 \end{array} \tag{8.101}$$

a) Assume $y_z(t)=x_z(t)+W_z(t)$ where $W_z(t)$ is a complex additive white Gaussian with $R_{W_z}(\tau)=N_0\delta(\tau)$, then the demodulator in (8.101) simplifies to a threshold test as in the binary case. Identify the decision statistic and the two decision thresholds.

b) Give an expression for the bit error probability $P\left(\hat{I}\neq I, \hat{I}\neq E\right)$ in terms of E_b, and N_0 for this signal set and demodulator. Plot this performance in comparison to the traditional binary demodulator.

Problem 8.10. Consider the MFSK example given in the text. Use the union bound to find the optimum frequency spacing for

a) $M=4$ at $E_b/N_0=8$dB.

b) $M=8$ at $E_b/N_0=9$dB.

For each case plot the union bound and compare it to the union bound for $\Re[\rho_{ij}]=0$. A computer might be a friend in this problem.

Problem 8.11. The touch-tone dialing in a telephone is a form of M-ary communication. When a key is pressed on a telephone two tones are generated, i.e.,

$$x_i(t)=A\cos\left(2\pi f_i(1)t\right)+A\cos\left(2\pi f_i(2)t\right) \tag{8.102}$$

This type of modulation is referred to as dual tone multiple frequency (DTMF) modulation. The DTMF tones can send $K_b=4$ bits even though there are only twelve keys on the phone. The modulation mappings are shown in Table 8.1. To simplify the problem assume $T_p=1$ seconds. Note also that the modulation is given as a real signal.

a) Plot the transmitted signal when person tries to call the operator ($i=13$) for $A=1$ over the interval $[0,0.1]$.

b) Transmission on telephone lines are often thought of as having $f_c=1200$ Hz. Give the simplest form of the complex envelope when a person tries to call the operator ($i=13$).

c) What is E_i $i\in\{0,\ldots,15\}$.

d) Detail out the MLWD. Can you get away with only 8 filters instead of 16? If so show the structure.

e) Compute and plot the union bound to the probability of word error for E_b/N_0=0-10dB. Note with a 16-ary modulation a computer will be your friend in this problem.

$f_i(1)$	$f_i(2)$			
	1209 Hz	1336 Hz	1477 Hz	1633 Hz
697 Hz	i=0 1	i=1 ABC 2	i=2 DEF 3	i=3 A
770 Hz	i=4 GHI 4	i=5 JKL 5	i=6 MNO 6	i=7 B
852 Hz	i=8 PRS 7	i=9 TUV 8	i=10 WXY 9	i=11 C
941 Hz	i=12 *	i=13 oper 0	i=14 #	i=15 D

Table 8.1: The dual tone multiple frequency (DTMF) modulation mappings

Problem 8.12. Consider a 4-ary communication waveform set that achieves $W_b = 2/T_p$ having the form shown in Fig.8.13. assume all words are equally likely and that the corrupting noise is an AWGN with a one-side spectral density of N_0.

a) Compute the value of A and B such that $E_i = 2E_b$ $i = 0, 3$

b) Identify the MLWD structure.

c) Show that only two filters are needed to implement this structure

d) Compute the union bound to the probability of word error.

Problem 8.13. Three bits of information are transmitted with an 8-ary modulation that has been called pulse position modulation (PPM) and is characterized with

$$\begin{aligned} x_0(t) &= u(t) \qquad x_1(t) = u(t - T_p/8) \qquad x_2(t) = u(t - T_p/4) \qquad x_3(t) = u(t - 3T_p/8) \qquad (8.103) \\ x_4(t) &= u(t - T_p/2) \qquad x_5(t) = u(t - 5T_p/8) \qquad x_6(t) = u(t - 3T_p/4) \qquad x_7(t) = u(t - 7T_p/8) \end{aligned}$$

where

$$u(t) = \begin{cases} \sqrt{\dfrac{24E_b}{T_p}} & 0 \le t \le T_p/8 \\ 0 & \text{elsewhere.} \end{cases} \qquad (8.104)$$

Recall for comparison that 8-ary orthogonal frequency shift keying (8FSK) has a minimum $f_d = \frac{1}{4T_p}$ and a transmitted signal given as

$$x_i(t) = \begin{cases} \sqrt{\dfrac{3E_b}{T_p}} \exp\left[\dfrac{j\pi t(i - 3.5)}{T_p}\right] & 0 \le t \le T_p \\ 0 & \text{elsewhere.} \end{cases} \qquad (8.105)$$

a) Detail out the maximum likelihood word demodulator for 8PPM. Are there any simplifications that are possible due to using PPM?

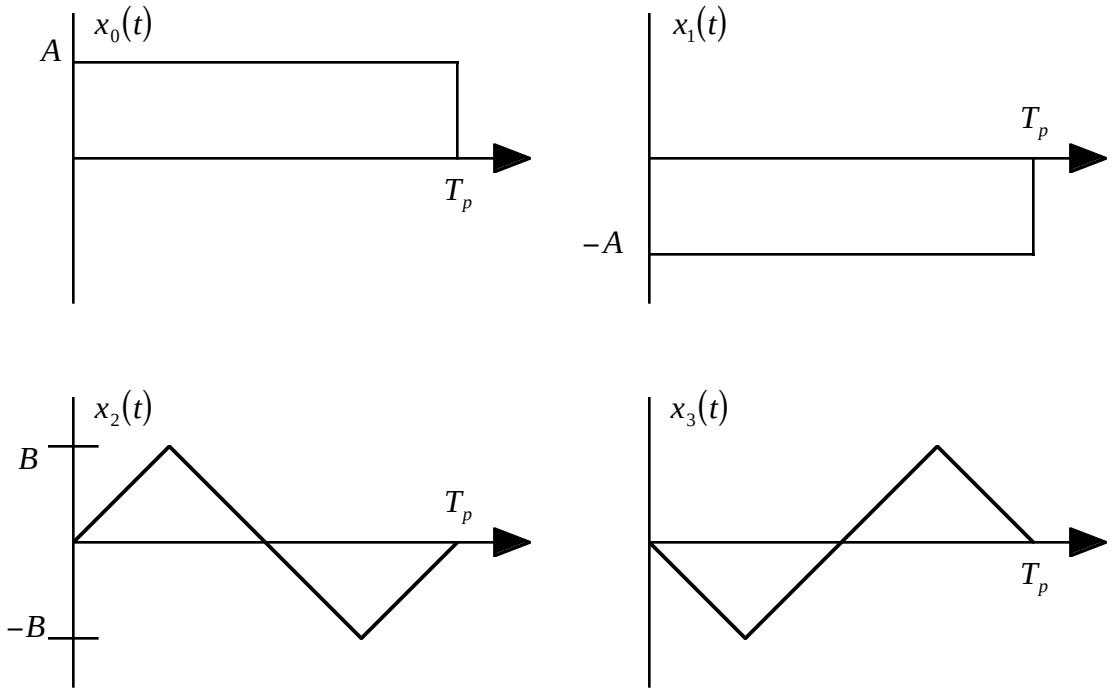


Figure 8.13: The waveforms for a 4-ary communication system.

b) Find the union bound to the word error performance for 8PPM. How will this performance compare to 8FSK?

c) Find the average energy spectrum per bit for 8PPM. How does this spectrum compare to 8FSK?

d) As a communication engineer assess the advantages and disadvantages of 8PPM versus 8FSK.

e) Assume a normal binary mapping into the word indices and find the maximum likelihood bit demodulator (MLBD) for the first bit.

Problem 8.14. In the wireless local network protocol denoted IEEE 802.11b the lowest rate modulation is a binary modulation using the two waveforms given in Fig. 7.29. A modified system chooses to send $K_b = 2$ equally likely bits by adding two more waveforms as shown in Fig. 8.14

a) Select A such that the energy per bit is E_b.

b) Show how to compute T_2.

c) Compute the union bound for this 4-ary modulation.

Problem 8.15. Recall that

$$\Re\left[W_z(t)x_i^*(t)\right] = W_I(t)x_{i,I}(t) + W_Q(t)x_{i,Q}(t) \tag{8.106}$$

and use it to show that the correlation coefficient between the noise in different maximum likelihood metrics is

$$\rho = \frac{E\left[N_I^{(i)}N_I^{(k)}\right]}{\frac{N_0}{2}\sqrt{E_iE_k}} = \Re\left[\rho_{ik}\right]. \tag{8.107}$$

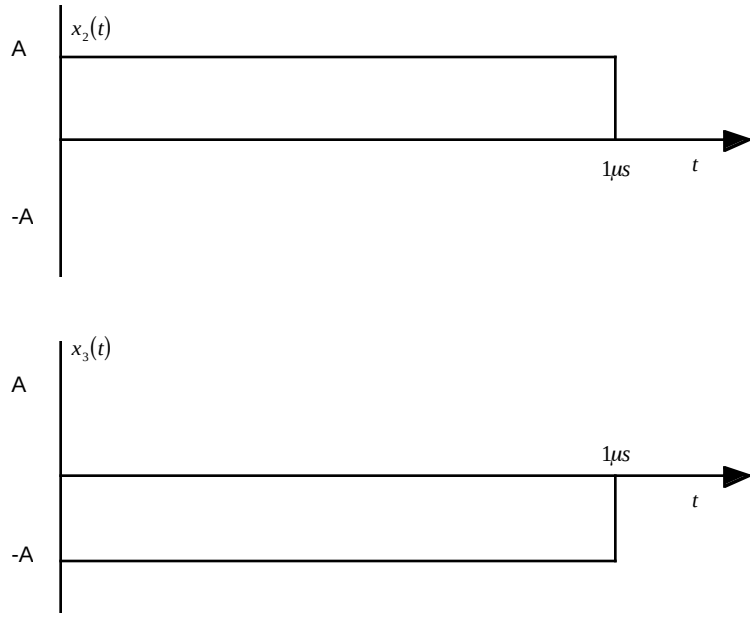


Figure 8.14: Two more waveforms for a 4-ary modulation.

Problem 8.16. Assume $K_b = 3$, equally likely transmitted signals, that the maximum likelihood metrics are given as

$$\begin{aligned} T_0 &= 1.375 \quad T_1 = 1 \quad T_2 = 1.25 \quad T_3 = 1 \quad T_4 = 0.75 \quad T_5 = 0.875 \\ T_6 &= 1.25 \quad T_7 = 1.2, \end{aligned} \tag{8.108}$$

and $N_0 = 0.5$.

a) Find the maximum likelihood word demodulation output word and bits.

b) Compute $L_0(k)$ and $L_1(k)$ for $k \in \{1, 2, 3\}$.

c) Find each of the maximum likelihood bit decisions. Are any of the maximum likelihood bit decisions different than the bit decisions in the maximum likelihood word demodulator output? If so why?

Problem 8.17. Assume $K_b = 2$ and that the maximum likelihood metrics and the a priori probabilities are all available, i.e., T_i and $\pi_i \quad i = 0, 1, 2, 3$.

a) Find the MAPBD for $I(1)$ and $I(2)$.

b) Taking the exponential function of the maximum likelihood metric greatly expands the range of the possible values needed in the processing of the signals in a MAPBD. Consequently the processing of the signals in MAPBD is often done in the "log" domain. An operation that has proved useful for operating in the "log" domain is known as the max* operation [**?**], specifically

$$\max^*(A, B) = \log\left(\exp[A] + \exp[B]\right). \tag{8.109}$$

Show that this can be evaluated as

$$\max^*(A, B) = \max(A, B) + \log\left(1 + \exp\left[-|B - A|\right]\right) \tag{8.110}$$

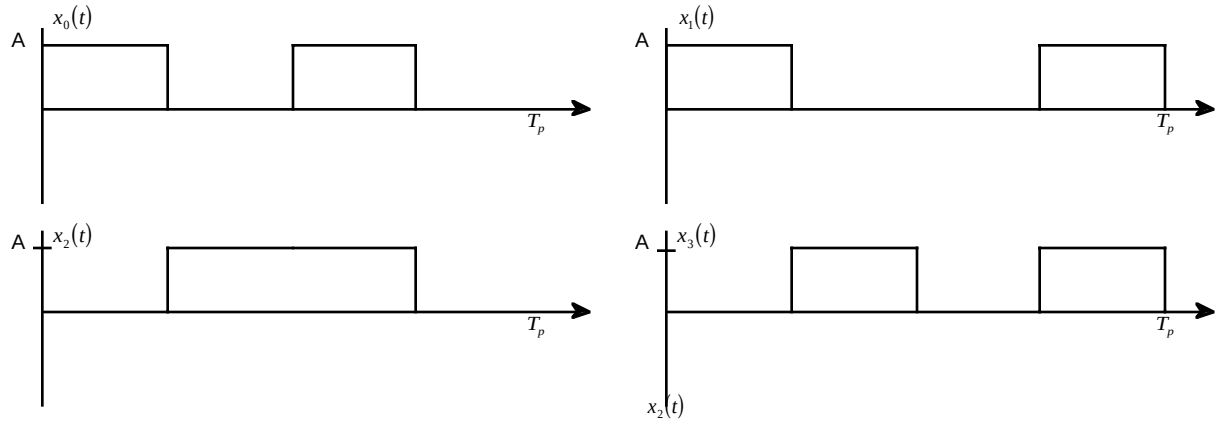


Figure 8.15: Waveforms for a 4-ary modulation.

c) Plot the funtion $\log(1+\exp[X])$ for $X<0$. How might you approximate the computation of this function to reduce complexity.

d) Show how to use max* to compute the MAPBD for $I(1)$ and $I(2)$. Use the notation $p_i = \log(\pi_i)$.

e) Show that

$$\max^*(A,B,C) = \max^*(\max^*(A,B),C). \tag{8.111}$$

Discuss how this impacts the computation of the MAPBD for $K_b > 2$.

Problem 8.18. The four waveforms in Fig. 8.15 are to be used to send $K_b = 2$ bits. Assume the received signal is distorted by an AWGN with a one–sided spectral density of N_0 and that all the signals are a priori equally likely.

a) Select A such that the energy per bit is E_b.

b) What is the average energy spectrum per bit?

c) Detail out the MLWD. It is possible to have the MLWD be implemented with one filter that is sampled at four different time instances. Give the form for this filter.

d) Compute the union bound for this 4-ary modulation.

e) Would this modulation perform better or worse than 4-ary orthogonal FSK in terms of word error probability?

f) Would this modulation have better or worse spectral efficiency than 4-ary orthogonal FSK?

g) Assume that $\vec{I} = i = [I(1)I(2)]$, find the MLBD for both $I(1)$ and $I(2)$. There turns out to be a simple threshold test for this problem. What characteristic enables this simple threshold test.

8.6 Example Solutions

Problem 8.18.

a) $2E_b = \frac{A^2 T_p}{2}$ so that $A = 2\sqrt{\frac{E_b}{T_p}}$.

b) Defining

$$u_4(t) = \begin{cases} 1 & 0 \le t \le T_p/4 \\ 0 & \text{elsewhere} \end{cases} \tag{8.112}$$

then it is apparent that

$$x_i(t) = u_4\left(t - \tau_{1,i}\right) + u_4\left(t - \tau_{2,i}\right). \tag{8.113}$$

Consequently the Fourier transform is given as

$$\begin{aligned} X_i(f) &= \frac{AT_p}{4}\text{sinc}\left(\frac{ft_p}{4}\right)\exp\left[-j2\pi f\tau_{1,i}\right] + \frac{AT_p}{4}\text{sinc}\left(\frac{ft_p}{4}\right)\exp\left[-j2\pi f\tau_{2,i}\right] && (8.114) \\ &= \frac{AT_p}{4}\text{sinc}\left(\frac{ft_p}{4}\right)\left(\exp\left[-j2\pi f\tau_{1,i}\right] + \exp\left[-j2\pi f\tau_{2,i}\right]\right) && (8.115) \end{aligned}$$

The energy spectrum of the individual waveforms is then given as

$$G_{x_i}(f) = \frac{A^2T_p^2}{16}\left(\text{sinc}\left(\frac{ft_p}{4}\right)\right)^2\left[2 + 2\cos\left(2\pi f\left(\tau_{1,i} - \tau_{2,i}\right)\right)\right]. \tag{8.116}$$

The average energy spectrum is then given as

$$D_{x_z}(f) = \frac{A^2T_p^2}{32}\left(\text{sinc}\left(\frac{ft_p}{4}\right)\right)^2\left[4 + 2\cos\left(\pi fT_p\right) + \cos\left(\frac{\pi fT_p}{2}\right) + \cos\left(\frac{3\pi fT_p}{2}\right)\right]. \tag{8.117}$$

c) The ML decision metric is

$$T_i = \Re\left[\int_0^{T_p} Y_z(t)x_i(t)dt\right] \qquad i = 0, \cdots, 3 \tag{8.118}$$

and the decoded bit is

$$\hat{\vec{I}} = \arg\max_{i=0,\cdots,3} T_i \tag{8.119}$$

It should be noted that this modulation is an equal energy modulation scheme and no energy correction scheme is required. The one filter has an impulse response of $u_4(t)$ where

$$Q(t) = \int Y_z(\lambda)u_4(t-\lambda)d\lambda. \tag{8.120}$$

The four sample points are

$$\begin{aligned} Q_1 = Q(T_p/4) &= \int_0^{T_p/4} Y_z(t)dt \qquad Q_2 = Q(T_p/2) = \int_{T_p/4}^{T_p/2} Y_z(t)dt && (8.121) \\ Q_3 = Q(3T_p/4) &= \int_{T_p/2}^{3T_p/4} Y_z(t)dt \qquad Q_4 = Q(T_p) = \int_{3T_p/4}^{T_p} Y_z(t)dt. && (8.122) \end{aligned}$$

The ML metrics are given as

$$T_0 = \Re\left[Q_1 + Q_3\right] \quad T_1 = \Re\left[Q_1 + Q_4\right] \quad T_2 = \Re\left[Q_2 + Q_3\right] \quad T_3 = \Re\left[Q_2 + Q_4\right] \tag{8.123}$$

d) A quick examination of the waveforms shows the signal set is geometrically uniform so that

$$P_{WUB}(E) = \text{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right) + \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{8.124}$$

e) Recall the union bound for orthogonal 4FSK is

$$P_{WUB}(E) = \frac{3}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{8.125}$$

minimum squared Euclidean distance between signals for 4FSK is 3dB larger than the minimum squared Euclidean distance for the signal set considered in this problem.

f) The spectral efficiency of 4FSK is approximately $\eta_B = 1$ while it is apparent from (8.117) the spectral efficiency of the considered modulation is approximately $\eta_B = 0.5$

g) Note that

$$T_{00} = \Re\left[Q_1 + Q_3\right] \quad T_{01} = \Re\left[Q_1 + Q_4\right] \quad T_{10} = \Re\left[Q_2 + Q_3\right] \quad T_{11} = \Re\left[Q_2 + Q_4\right] \tag{8.126}$$

so it is easy to see that $T_{m_1 m_2} = T_{m_1}^{(1)} + T_{m_2}^{(2)}$ where

$$T_{m_k}^{(k)} = \Re\left[Q_{2(k-1)+m_k+1}\right] \tag{8.127}$$

and the maximum likelihood metric has an additive form as needed for a simple MLBD. The threshold tests are

$$\Re\left[Q_1\right] \underset{\hat{I}(1)=1}{\overset{\hat{I}(1)=0}{\gtrless}} \Re\left[Q_2\right] \qquad \Re\left[Q_3\right] \underset{\hat{I}(1)=1}{\overset{\hat{I}(1)=0}{\gtrless}} \Re\left[Q_4\right] \tag{8.128}$$

Chapter 9

Managing the Complexity of Optimum Demodulation

The results of the previous chapter have shown that the idea of transmitting and demodulating many bits of information is a direct extension of the ideas for transmitting and demodulating one bit of information. The biggest impediment to using these concepts in practical communication systems is the fact that the optimal demodulator complexity grows exponentially with the number of bits transmitted ($O(2^{K_b})$). The goal for a practical system has to be a demodulator with complexity that grows linearly with the number of bits ($O(K_b)$). Fortunately the example of M-ary phase shift keyed (MPSK) modulation considered in Chapter 8 gives us two insights into how complexity of optimal demodulation can be made practical. These two characteristics of MPSK modulation are

1. MPSK has a form $x_i(t) = d_i u(t)$ where a modulation symbol, d_i, is linearly modulated on a pulse shape, $u(t)$. When a modulation takes this form only one matched filter needs to be computed for the optimum demodulator. The complexity of the demodulator is then greatly reduced.

2. Gray coded 4PSK has an optimum bit demodulator that has exactly the same form as a demodulator for a single bit being transmitted in isolation. If this decoupling of each bit decision can be generalized it would allow the complexity of the optimal demodulator to have a complexity that is linear in the number of bits transmitted.

9.1 Linear Modulations

Linear modulation has the form

$$x_i(t) = d_i \sqrt{E_b} u(t) \; i \in \{0, \ldots, M-1\}. \tag{9.1}$$

The function of time $u(t)$ is known as the pulse shape. In this text it will be assumed that $u(t)$ has unit energy, $E_u = 1$. The modulation symbol, d_i, is, in general, complex valued function of the transmitted information word, $D_z = a\left(\vec{I}\right)$, and $d_i = a(i) \in \Omega_d$. The function $a(i)$ is known as the constellation mapping. A commonly used graphic for interpreting linear modulations is the constellation plot. A constellation plot for 4PSK modulation, $\Omega_d = \left\{\sqrt{2}e^{\frac{j\pi}{4}}, \sqrt{2}e^{\frac{j3\pi}{4}}, \sqrt{2}e^{\frac{-j\pi}{4}}, \sqrt{2}e^{\frac{-j3\pi}{4}}\right\}$, is shown in Fig. 9.1. For a consistent energy normalization the constellation will always be chosen such that

$$\sum_{i=0}^{M-1} |d_i|^2 \pi_i = K_b. \tag{9.2}$$

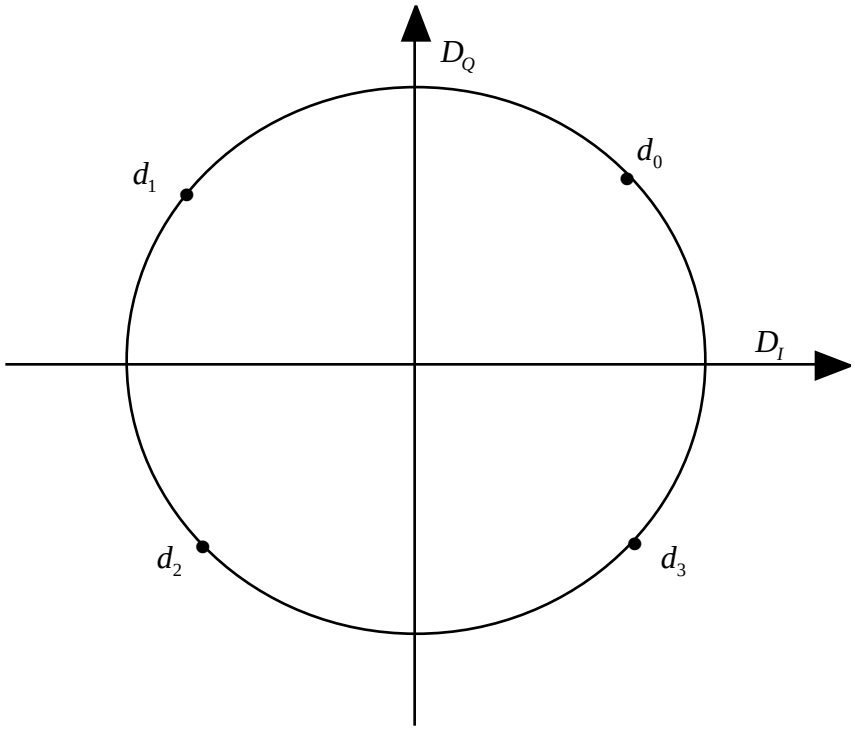


Figure 9.1: The constellation plot of 4PSK modulation.

9.1.1 MLWD for Linear Modulation

The MLWD for a linear modulation has a greatly simplified form. Recall the optimum demodulator has the form

$$\hat{\vec{I}} = \arg \max_{i \in \{0,\ldots,M-1\}} \Re\left[V_i(T_p)\right] - \frac{E_i}{2} \tag{9.3}$$

where the i^{th} matched filter output has the form

$$\begin{aligned} V_i(T_p) &= \int_0^{T_p} y_z(t) x_i^*(t) dt \\ &= d_i^* \sqrt{E_b} \int_0^{T_p} y_z(t) u^*(t) dt \\ &= d_i^* \sqrt{E_b} Q \end{aligned} \tag{9.4}$$

where Q is denoted that pulse shape matched filter output and is given as

$$Q = \int_0^{T_p} y_z(t) u^*(t) dt \tag{9.5}$$

Note the matched filter can be viewed as a fixed filter with an impulse response $u^*(T_p - t)$ that is sampled at time T_p. A block diagram of the demodulator for linear modulations is shown in Fig. 9.2. Several observations can be made about the this structure:

1. The number of filtering operations in the demodulation has been reduced from $M = 2^{K_b}$ to one when compared to the demodulator for a general M-ary modulation.

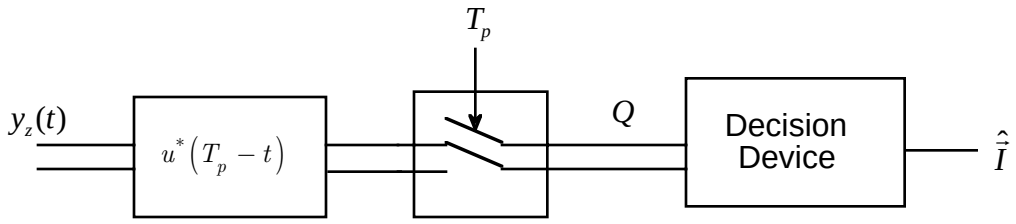


Figure 9.2: The block diagram of a demodulator for linear modulation.

2. The decision device can still potentially have a complexity that is $O\left(2^{K_b}\right)$.

3. The energy of the i^{th} transmitted signal is $E_i = |d_i|^2 E_b$.

The optimum demodulator for linear modulation has an intuitively pleasing geometric interpretation. The optimum demodulator can be expressed as

$$\hat{\vec{I}} = \arg \max_{i \in \{0,\ldots,M-1\}} \sqrt{E_b} \Re \left[d_i^* Q \right] - \frac{|d_i|^2 E_b}{2}. \tag{9.6}$$

Multiplying by 2 and adding a constant with respect to the possible transmitted symbols, $-|Q|^2$, gives

$$\begin{aligned} \hat{\vec{I}} &= \arg \max_{i \in \{0,\ldots,M-1\}} -|Q|^2 + 2\Re \left[d_i^* Q \sqrt{E_b} \right] - |d_i|^2 E_b \\ &= \arg \max_{i \in \{0,\ldots,M-1\}} -\left| Q - d_i \sqrt{E_b} \right|^2 \\ &= \arg \min_{i \in \{0,\ldots,M-1\}} \left| Q - d_i \sqrt{E_b} \right|^2 \end{aligned} \tag{9.7}$$

Consequently the optimum demodulator computes Q and the finds the possible transmitted constellation point $d_i\sqrt{E_b}$ which has the minimum square Euclidean distance to Q. This is known by communications engineers as a minimum distance decoder. This minimum distance decoder induces decision regions in the complex plane. These decision regions represent the matched filter output values that correspond to a most likely transmitted symbol. The decision regions for 4PSK MLWD is shown in Fig. 9.3. In the sequel the decision regions will be denoted A_i where if $Q \in A_i$ then $\hat{I} = i$ represents the MLWD output.

Recall the matched filter output when the i^{th} word was transmitted has the form

$$\begin{aligned} Q &= \int_0^{T_p} y_z(t) u^*(t) dt = \int_0^{T_p} x_i(t) u^*(t) dt + \int_0^{T_p} W_z(t) u^*(t) dt \\ &= d_i \sqrt{E_b} + N_z(T_p). \end{aligned} \tag{9.8}$$

With this view the signal portion of the matched filter output corresponds to the transmitted symbol scaled by the bit energy, $\sqrt{E_b}$, and the noise causes a random translation of the signal in the complex plane. The idea of a minimum distance decoder is simply trying to find which signal set would correspond to the smallest noise magnitude. For example if d_0 from the 4PSK constellation in Fig.9.1 is the transmitted symbol then the matched filter output would have a representation as given in Fig. 9.4.

9.1.2 Performance Evaluation for Linear Modulation

The symbol error probability calculation also has a nice geometric interpretation. Comparing the signal form in Fig. 9.4 to decision regions in Fig. 9.3 an error will occur when the noise, $N_z(T_p)$, pushes the

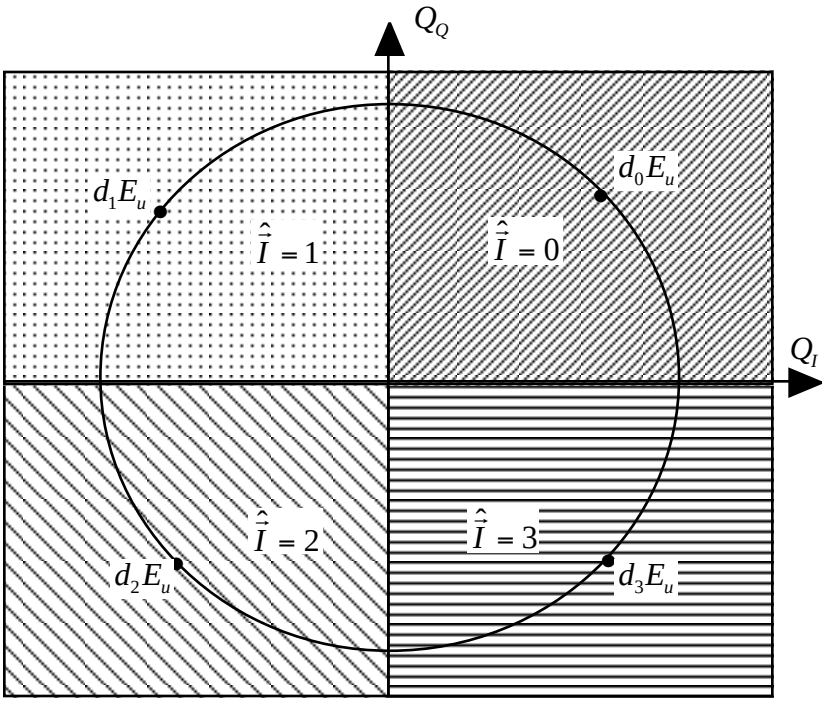


Figure 9.3: The minimum distance decision regions for 4PSK.

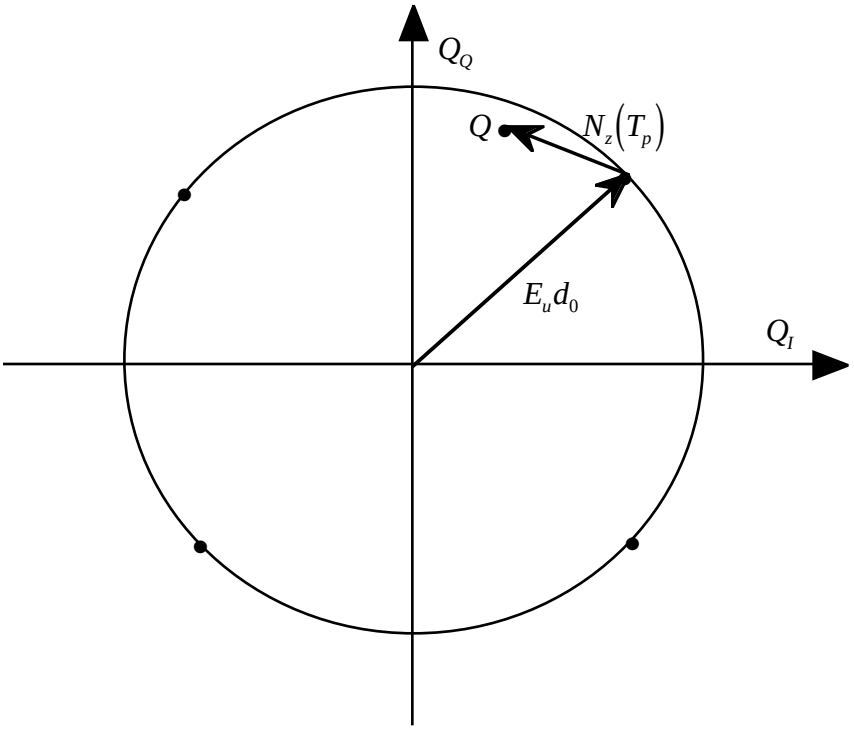


Figure 9.4: An example geometric interpretation of the signal and noise in linear modulation. $D_z = d_0 = \sqrt{2}e^{j\pi/4} = 1 + j$.

matched filter into a decision region that does not correspond to the transmitted symbol. Since $N_z(T_p)$ is a complex Gaussian random variable with

$$\mathrm{var}\left(N_z(T_p)\right) = E_u N_0 = N_0 \qquad \mathrm{var}\left(N_I(T_p)\right) = \frac{N_0}{2} = \mathrm{var}\left(N_Q(T_p)\right) \tag{9.9}$$

the PDF of the noise is easily computed using (4.26). An example noise PDF is plotted in Fig. 9.6-a). The probability of word error conditioned on the i^{th} symbol being sent is

$$P_W(E\,|i\,) = P\left(Q \notin A_i\right) = P\left(d_i\sqrt{E_b} + N_z(T_p) \notin A_i\right). \tag{9.10}$$

Consequently this probability is simply an integral of the noise PDF, i.e.,

$$P_W(E\,|i\,) = \int_{R_i} f_{N_z}(n_z)dn_z = \int 1_{R_i}(n_z) f_{N_z}(n_z)dn_z \tag{9.11}$$

where the indicator function is used to mask the PDF only in the regions where noise will cause an error. Fig. 9.6-b) shows the masked PDF for a 4PSK decision regions given in Fig. 9.3 for $d_0 = \sqrt{2}\exp\left(j\frac{\pi}{4}\right)$.

The $P_W(E)$ for a given K_b can be a significant function of how the constellation points are placed in the complex plane. For example consider two common linear modulations, pulse amplitude modulation (PAM) and phase shift keyed (PSK) modulation that have been implemented often in engineering practice. PSK uses only the phase to transmit information while PAM uses only the amplitude to transmit information. Examples of the constellation plots for these two modulations are shown in Fig. 9.5 for $K_b = 2$. Note 4-ary pulse amplitude modulation (4PAM) is characterized with $\Omega_d = \left\{\pm\sqrt{2}/\sqrt{5}, \pm 3\sqrt{2}/\sqrt{5}\right\}$. The $P_W(E)$ for each of these linear modulations are given for $K_b = 1, 2$ as (see problems)

$$P_W(E) = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \qquad \text{BPSK/BPAM} \tag{9.12}$$

$$P_W(E) = \mathrm{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) - \left(\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)\right)^2 \qquad \text{4PSK} \tag{9.13}$$

$$P_W(E) = \frac{3}{4}\mathrm{erfc}\left(\sqrt{\frac{2E_b}{5N_0}}\right) \qquad \text{4PAM.} \tag{9.14}$$

These curves are plotted in Fig 9.7. It is clear from Fig 9.7 that how constellation points are chosen in sending K_b bits with a linear modulation can have a significant impact on the performance.

The union bound for the word error probability of a linear modulation is straightforward to compute. The union bound can be formed by noting that the pair-wise Euclidean squared distance is given as

$$\Delta_E(i,j) = E_b\left|d_i - d_j\right|^2. \tag{9.15}$$

Consequently the union bound can be form only by looking at the constellations and does not need to worry about the pulse shaping function.

So an additional benefit of linear modulation beside the simple MLWD is that the $P_W(E)$ can always be calculated. This calculation might require a little computer work (two dimensional numerical integration) but this effort is not unreasonable for a modern communication engineer. While often a general M-ary modulation must resort to a union bound computation to characterize the $P_W(E)$ this is not necessary for linear modulation.

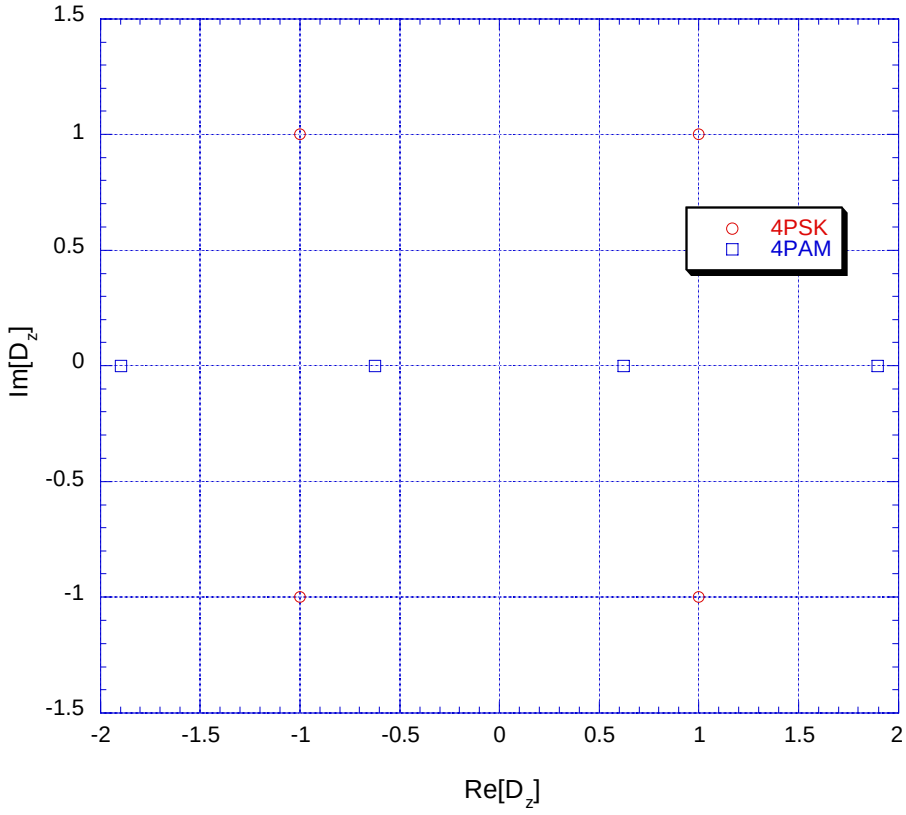
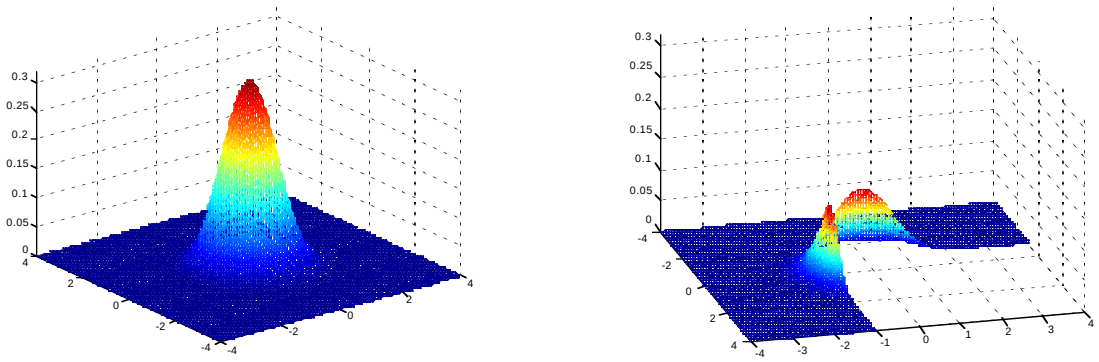


Figure 9.5: Constellation diagrams for 4PAM and 4PSK.



a) Noise PDF

b) Noise PDF after masking, $d_0 = 1 + j1,\ E_b = 1$

Figure 9.6: A noise PDF corresponding to var $(N_z(T_p)) = 1$.

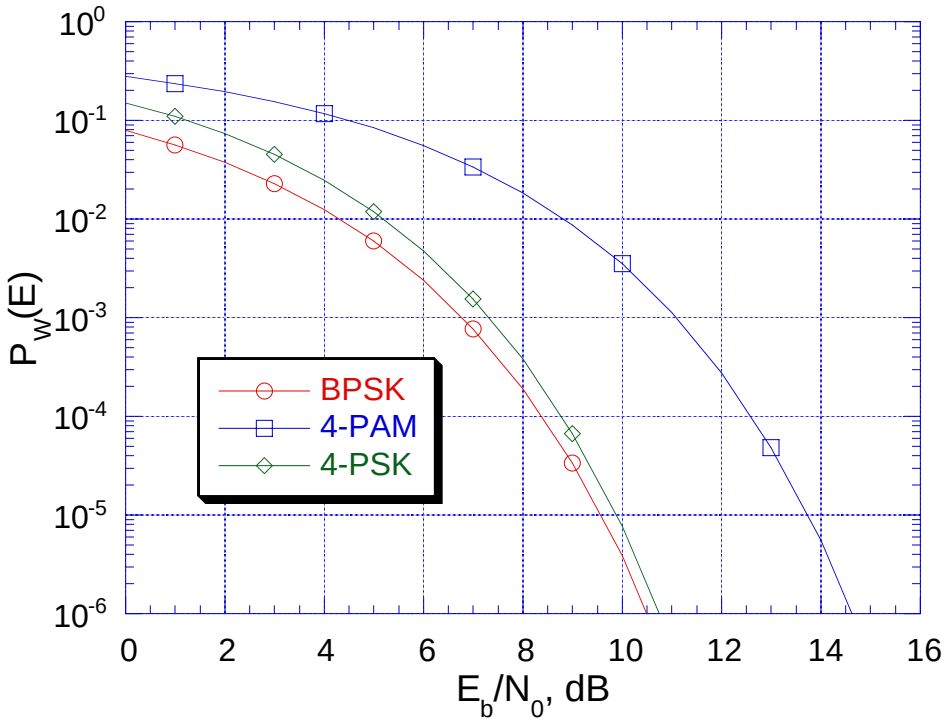


Figure 9.7: Word error probability curves for several linear modulations.

9.1.3 Spectral Characteristics of Linear Modulation

For linear modulations the average energy spectrum per bit is straightforward to compute.

$$D_{x_z}(f) = \frac{E\left[G_{x_z}(f)\right]}{K_b} = G_u(f)\frac{E\left[|D_z|^2\right]}{K_b} = G_u(f)\frac{\sum_{i=0}^{M-1}|d_i|^2\pi_i}{K_b} = E_b G_u(f). \tag{9.16}$$

The important point to note about (9.16) is that the average energy spectrum per bit is entirely (up to a constant multiple) a function of the pulse shape spectrum, $G_u(f)$. Transmitting more bits (increasing K_b) has no effect on the occupied spectrum. Consequently with a bandwidth efficient choice of the pulse shape, $u(t)$, linear modulation is a very effective modulation for bandlimited channels. A design of a pulse shape that will meet a specified spectral mask is explored in the projects at the end of the Chapter 7.

9.1.4 Summary of Linear Modulation

Important points about linear modulation are

- Demodulation of linear modulation does not increase exponentially with the number of bits transmitted since only one matched filter output needs to be computed and simple decision regions can be identified. Linear modulation is often used because it has a low demodulation complexity.

- The energy spectrum per bit of a linear modulation is entirely a function of the pulse shape that is used. Bandwidth efficient pulse shapes are often used in conjunction with M-ary linear modulation when bandwidth efficiency is of paramount importance. This will be examined in more detail in Chapter 10.

- Performance gets significantly worse as M gets larger. This can be seen by examining the performance of linear modulations in both Fig 9.7 and Fig 8.8. Consequently linear modulation provides a classic tradeoff in bandwidth efficiency (highly bandwidth efficient) versus performance (not highly energy efficient).

- The selection of the constellation points in an M-ary linear modulation can significantly affect performance. This can be seen by examining the performance of 4PAM and 4PSK (both K_b=2) in Fig 9.7. 4PAM has a 3dB worse performance than 4PSK since the 4PAM constellation points have not been placed as effectively in the complex plane as the 4PSK constellation points. An interesting paper that investigates optimal and suboptimal constellations is [FGL$^+$84].

9.2 Low Complexity Optimal Bit Decisions

Recall from Chap 8 that Gray coded 4PSK has an optimum bit demodulator that has exactly the same form as a demodulator for a single bit being transmitted in isolation. If this decoupling of each bit decision can be generalized it would allow the complexity of the optimal demodulator to have a complexity that is linear in the number of bits transmitted. Here we explore the most general conditions that must exist for optimum signal bit demodulation to have the same complexity as the case when the bit is translated in isolation as discussed in Chap. 7.

9.2.1 Optimum Bit Demodulation

We know from the previous chapter that the optimum bit demodulator, the MAPBD, is

$$P\left(I(k)=1\,|y_z(t)\right) \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} P\left(I(k)=0\,|y_z(t)\right). \tag{9.17}$$

Simplifications presented in Chap. 8 show the optimum bit decision is given as

$$\sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{1,n\}}}{N_0}\right]\pi_{\{1,n\}} \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} \sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{0,n\}}}{N_0}\right]\pi_{\{0,n\}}. \tag{9.18}$$

From this point forward we will assume each of the bits is generated independently, i.e.,

$$\pi_i = \prod_{l=1}^{K_b} \pi_{m_l}(l) \tag{9.19}$$

where $\vec{I} = i = [m_1\ m_2\ \ldots\ m_{K_b}]$. Without this assumption there is little hope in finding simplified optimum single bit decisions as all the individual bits are related before they are even transmitted. In a similiar fashion the MLBD for M-ary modulations has the form

$$\sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{1,n\}}}{N_0}\right] \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} \sum_{n=0}^{M/2-1} \exp\left[\frac{2T_{\{0,n\}}}{N_0}\right]. \tag{9.20}$$

Note that the summation in (9.20) can only have a simple form if the terms corresponding to the bit of interest, $I(k)$, can be factored out of the summation. It is easy to see that this can only be if

$$T_{\{m,n\}} = T_n^{(-)} + T_m^{(k)} \tag{9.21}$$

where

$$T_n^{(-)} = f_1\left(I^{(-)}(k)=n\right) \qquad T_m^{(k)} = f_2\left(I(k)=m\right). \tag{9.22}$$

Since the maximum likelihood metric has the sum form for every possible transmitted word, the APP for each word can be formed into a product of two terms.

Since the APP of each word can be factored into a term related to the bit of interest and all other bits, the MAPBD simplifies to

$$\exp\left[\frac{2T_1^{(k)}}{N_0}\right]\pi_1(k)\sum_{n=0}^{M/2-1}\exp\left[\frac{2T_n^{(-)}}{N_0}\right]\pi_n \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} \exp\left[\frac{2T_0^{(k)}}{N_0}\right]\pi_0(k)\sum_{n=0}^{M/2-1}\exp\left[\frac{2T_n^{(-)}}{N_0}\right]\pi_n \tag{9.23}$$

Cancelling out the common terms in (9.23) gives

$$\exp\left[\frac{2T_1^{(k)}}{N_0}\right]P\left(I(k)=1\right) \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} \exp\left[\frac{2T_0^{(k)}}{N_0}\right]P\left(I(k)=0\right). \tag{9.24}$$

In is very important to note that the form in (9.24) is exactly the same form as the MAP detector for one bit transmitted in isolation (see Chapter 7 and 8). In other words, if the maximum likelihood metric

can be put into an additive form for the bit of interest like in (9.21) then each bit can be optimally detected independently. Likewise for MLBD if $T_{\{m,n\}} = T_n^{(-)} + T_m^{(k)}$ then

$$\exp\left[\frac{2T_1^{(k)}}{N_0}\right] \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} \exp\left[\frac{2T_0^{(k)}}{N_0}\right]. \tag{9.25}$$

The monotonicity of the exponential function results in

$$T_1^{(k)} \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} T_0^{(k)}. \tag{9.26}$$

The outstanding question is how do we get

$$T_{\{m,n\}} = \Re\left[V_{\{m,n\}}(T_p)\right] - \frac{E_{\{m,n\}}}{2} = T_n^{(-)} + T_m^{(k)}? \tag{9.27}$$

First it should be noted that if the additive form holds for each of the K_b bits then the maximum likelihood metric needs to be in an additive form as in

$$T_i = \sum_{k=1}^{K_b} T_{m_k}^{(k)} \tag{9.28}$$

where $\left\{\vec{I}\right\} = \{I(1) = m_1, \cdots, I(K_b) = m_{K_b}\}$. In other words each T_i can be broken up into a sum of terms which are individually only a function of a single bit position, k, and bit value for that position, $I(k) = m_k$.

In general there are a wide variety of ways that this sum form can be achieved for the ML decision metric. At this point we return to the example of Gray coded 4PSK modulation that was introduced in Chap. 8 to give one example. Denoting $i = \{m_1, m_2\}$, it is apparent that Gray coded 4PSK is a constant energy linear modulation (M=4) with

$$d_i = d_{m_1} + jd_{m_2}, \tag{9.29}$$

where the in-phase and quadrature modulation is given as $d_m = (-1)^m$. It is clear that the in-phase part of the message symbol is a function of only $I(1)$ and the quadrature part of the message symbol is a function of only $I(2)$. Consequently the ML decision metric is given as

$$\begin{aligned} T_i &= \sqrt{E_b}\Re\left[d_i^* Q\right] - E_b = \sqrt{E_b}\Re\left[\left(d_{m_1} + jd_{m_2}\right)^* Q\right] - E_b \\ &= \sqrt{E_b}\Re\left[d_{m_1} Q\right] - \frac{E_b}{2} + \sqrt{E_b}\Re\left[-jd_{m_2} Q\right] - \frac{E_b}{2} = T^{(1)} + T^{(2)}. \end{aligned} \tag{9.30}$$

Gray coded 4PSK has a simple optimum bit demodulator because each of the two bits is modulated on an orthogonal set of signals (the I and Q components of the data symbol). The most commonly used techniques to achieve this simple demodulation structure is considered in the next section.

9.2.2 Orthogonal Modulations

Many of the modulation schemes that are used in practice for their reduced complexity optimal demodulation are due to the following property

Property 9.1 (Orthogonal Modulations) *If*

$$X_i(t) = \sum_{l=1}^{K_b} x\left(m_l, l, t\right) \tag{9.31}$$

where again $\vec{I} = i = [m_1\ m_2\ \ldots\ m_{K_b}]$ *then the MLBD metric can be put in the additive form required for simplified bit detection if the waveforms used to transmit each of the bits are orthogonal, i.e.,*

$$\Re\left[\int_{-\infty}^{\infty} x\left(m_l, l, t\right) x^*\left(m_k, k, t\right) dt\right] = 0 \qquad \forall k \neq l, m_l = 0, 1, m_k = 0, 1 \tag{9.32}$$

Proof: The ML metric consists of two components: the matched filter output and the energy correction term. The matched filter term for the transmitted signal that has the form in (9.31) is given as

$$\begin{aligned}\Re\left[\int_{-\infty}^{\infty} y_z(t) x_i^*(t) dt\right] &= \Re\left[\int_{-\infty}^{\infty} y_z(t) \left(\sum_{l=1}^{K_b} x\left(m_l, l, t\right)\right)^* dt\right] \\ &= \sum_{l=1}^{K_b} \Re\left[\int_{-\infty}^{\infty} y_z(t) x^*\left(m_l, l, t\right) dt\right]. \end{aligned} \tag{9.33}$$

Consequently just due to the form in (9.31) the matched filter output has the additive form required for simplified detection. The remaining term is the energy correction term given for (9.31) as

$$\begin{aligned}\frac{E_i}{2} &= \frac{1}{2}\int_{-\infty}^{\infty} |x_i(t)|^2 dt = \frac{1}{2}\int_{-\infty}^{\infty} \sum_{l=1}^{K_b} x\left(m_l, l, t\right) \left(\sum_{k=1}^{K_b} x\left(m_k, k, t\right)\right)^* dt \\ &= \frac{1}{2}\sum_{l=1}^{K_b}\int_{-\infty}^{\infty} \left|x\left(m_l, l, t\right)\right|^2 dt + \frac{1}{2}\sum_{l=1}^{K_b}\sum_{\substack{k=1 \\ k\neq l}}^{K_b}\int_{-\infty}^{\infty} x\left(m_l, l, t\right)\left(x\left(m_k, k, t\right)\right)^* dt \qquad (9.34) \\ &= \frac{1}{2}\sum_{l=1}^{K_b} E_{x(m_l,l)} + \sum_{l=2}^{K_b}\sum_{k=1}^{l} \Re\left[\int_{-\infty}^{\infty} x\left(m_l, l, t\right)\left(x\left(m_k, k, t\right)\right)^* dt\right]. \qquad (9.35)\end{aligned}$$

Clearly the second term in (9.35) will go to zero if the orthogonality condition holds and then the energy correction term will also have an additive form. □

An interesting characteristic of orthogonal modulations is that the decision metric has exactly the same form as the single bit detector presented in Chapter 7. Using the above notation for orthogonal modulation we see that the optimal decision rule (MAPBD) is given as

$$\exp\left[\frac{2T_1^{(k)}}{N_0}\right]\pi_1 \underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} \exp\left[\frac{2T_0^{(k)}}{N_0}\right]\pi_0. \tag{9.36}$$

The MLBD has the form

$$\begin{aligned} T_1^{(k)} &\underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} T_0^{(k)} \\ \Re\left[V_1(k)\right] - \frac{1}{2}E_{x(1,k)} &\underset{\hat{I}(k)=0}{\overset{\hat{I}(k)=1}{\gtrless}} \Re\left[V_0(k)\right] - \frac{1}{2}E_{x(0,k)} \end{aligned} \tag{9.37}$$

where

$$V_{m_k}(k) = \int_{-\infty}^{\infty} y_z(t) x^* (m_k, k, t)\, dt \tag{9.38}$$

is the matched filter to the waveform used to transmit $I(k) = m_k$. Since the orthogonality reduces the demodulation to exactly that of a single bit developed in Chapter 7 we can deduce the performance is exactly that of the single bit demodulator. Consequently the bit error probability performance is lower bounded by

$$P_B(E) \geq \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right). \tag{9.39}$$

The orthogonality implies each decision is independent so that the word error probability for orthogonal modulation becomes (see Problem 9.12)

$$P_W(E) = 1 - (1 - P_B(E))^{K_b} \geq 1 - \left(1 - \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)\right)^{K_b}. \tag{9.40}$$

9.3 Orthogonal Modulation Examples

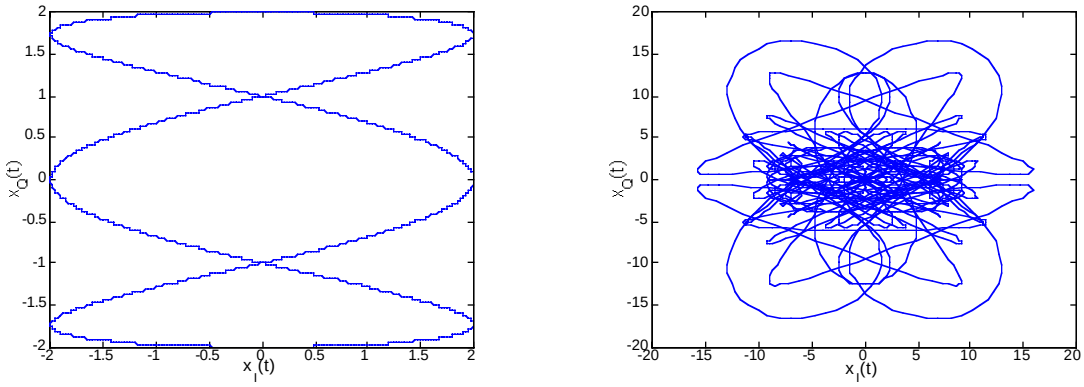
9.3.1 Orthogonal Frequency Division Multiplexing

A commonly used modulation that admits a simple optimal bit demodulation is orthogonal frequency division multiplexing (OFDM) [Cha66]. OFDM has found utility in telephone, cable, and wireless modems. With OFDM each of the K_b bits is independently modulated on a separate subcarrier frequency and the subcarrier frequencies are chosen to ensure the orthogonality. The format for an OFDM signal is

$$X_z(t) = \begin{cases} \displaystyle\sum_{l=1}^{K_b} D_z(l)\sqrt{\frac{E_b}{T_p}} \exp\left[j2\pi f_d(2l - K_b - 1)t\right] & 0 \leq t \leq T_p \\ 0 & \text{elsewhere} \end{cases} \tag{9.41}$$

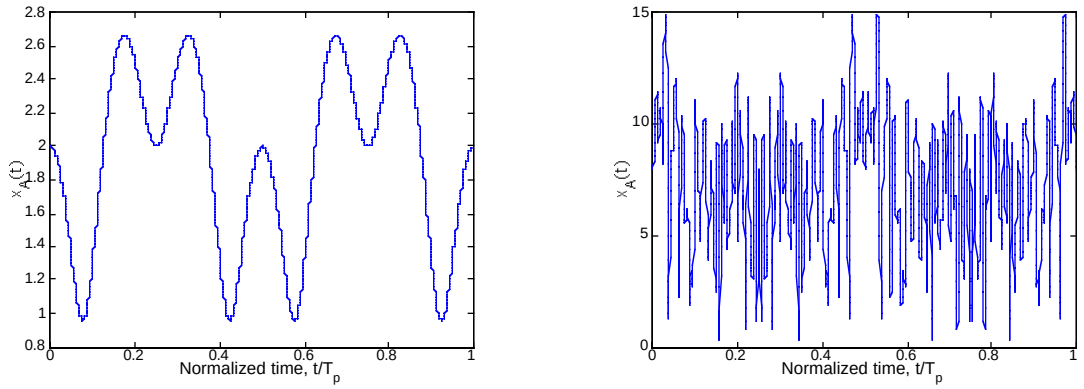
where $D(l) = a\left(I(l)\right)$ and $2f_d$ is the separation between adjacent freqeuncies that are used to transmit the information. The transmission rate of this form of OFDM is $W_b = K_b/T_p$ bits per second. In examining (9.41) it is clear that in this form of OFDM a binary linear modulation is used on each of the K_b different frequencies. For clarity of discussion the remainder of the section will assume the linear modulation is BPSK (i.e., $a(0) = 1$ and $a(1) = -1$). A more general form of OFDM could use any type of linear modulation and any number of bits per subcarrier. For example 4 bits could be transmitted per frequency using a 16QAM modulation (see Problem 9.1). Some of the generalizations of OFDM will be explored in the homework.

The OFDM transmitted waveform is a sum of K_b complex sinusoids. This transmitted waveform will have a complex envelope that changes significantly over the transmission time as the K_b complex sinusoids change in phase relative to each other. The larger the value of K_b the larger this variation over the transmission time will be. Fig. 9.8 shows the vector diagrams of some example OFDM transmitted waveforms. The vector diagram clearly becomes more complex as more bits are transmitted. In addition the peak value of the amplitude of the complex envelope increases with K_b. For example Fig. 9.9 shows $x_A(t)$ of some example OFDM transmitted waveforms. An OFDM waveform has a significant difference between the peaks in amplitude and the average value of the amplitude. This high peak to average



a) K_b=4 b) K_b=64

Figure 9.8: Example vector diagrams of OFDM transmissions.



a) K_b=4 b) K_b=64

Figure 9.9: Example $x_A(t)$ for OFDM transmissions.

ratio requires the radios in an OFDM system to have the dynamic range necessary to process the signal without distortion.

In the general framework of orthogonal modulations OFDM has set

$$x\left(I(l),l,t)\right) = \begin{cases} D_z(l)\sqrt{\dfrac{E_b}{T_p}}\exp\left[j2\pi f_d(2l-K_b-1)t\right] & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{9.42}$$

The orthogonality condition then corresponds to

$$\Re\left[\int_{-\infty}^{\infty} x\left(m_l,l,t\right)x^*\left(m_k,k,t\right)dt\right] = 0 \tag{9.43}$$

$$\Re\left[d_{m_l}d^*_{m_k}\frac{E_b}{T_p}\int_0^{T_p}\exp\left[j4\pi f_d(l-k)t\right]dt\right] = 0. \tag{9.44}$$

Using the result of (7.79) shows that the minimum spacing with an arbitrary complex constellation for orthogonality is achieved for $f_d = 1/(2T_p)$. If the constellation is restricted to have real values (e.g., BPSK) then the minimum frequency spacing is achieved with $f_d = 1/(4T_p)$.

Example 9.1: The wireless local area network standard denoted IEEE 802.11a uses OFDM. The 802.11a standard uses OFDM that can be represented with $K_b = 53$ with $T_p = 3.2\mu$s. The frequency spacing of the subcarrier tones is set with $f_d = 156.25$kHz because complex constellations are used in some modes of transmission. The potential transmission rate of the 802.11a waveform with BPSK modulation would be 16.25 MHz. The top transmission speed of 54Mbps is achieved by using a QAM modulation on each subcarrier. QAM constellations are investigated in the problems at the end of the chapter.

The optimal demodulator has the form of K_b parallel single bit optimal demodulators. This demodulator is shown in Fig. 9.10. Restricting ourselves to BPSK modulation on each subcarrier frequency the MLBD has the form

$$\Re\left[\int_0^{T_p} Y_z(t)\sqrt{\frac{1}{T_p}}\exp\left[-j2\pi f_d(2k-K_b-1)t\right]dt\right] = \Re\left[Q(k)\right] \underset{\hat{I}(k)=1}{\overset{\hat{I}(k)=0}{\gtrless}} 0 \tag{9.45}$$

Since OFDM is using a linear modulation on each subcarrier we will use $Q(k)$ to denote the matched filter output for the k^{th} bit or subcarrier. The demodulator computes a filter output for each bit, $k \in \{1,\ldots,K_b\}$, and hence the complexity of the OFDM optimum demodulator is $O(K_b)$ as opposed to the $O(2^{K_b})$ for an arbitrary modulation that transmits K_b bits of information. For future reference here we note that $Q(k)$ can be viewed at the Fourier transform of $Y_z(t)$ evaluated at $f = f_k = f_d((2k-K_b-1)$. The OFDM optimum demodulator evaluates the real part of the Fourier transform at K_b points symmetrically spaced around $f = 0$ and uses these Fourier transform values in a threshold test. Since $\Re\left[Q(k)\right] = D_z(k)\sqrt{E_b} + N_I(k)$ where $\text{var}\left(N_I(k)\right) = \frac{N_0}{2}$, the bit error probability performance of OFDM using BPSK on each subcarrier is clearly

$$P_B(E) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right). \tag{9.46}$$

The optimum demodulator for OFDM is essentially K_b single bit demodulators implemented in parallel (one for each subcarrier). Because of the orthogonality of the subcarriers the performance of an OFDM demodulator is the same a the demodulator for a single bit transmitted in isolation (see Chapter 7).

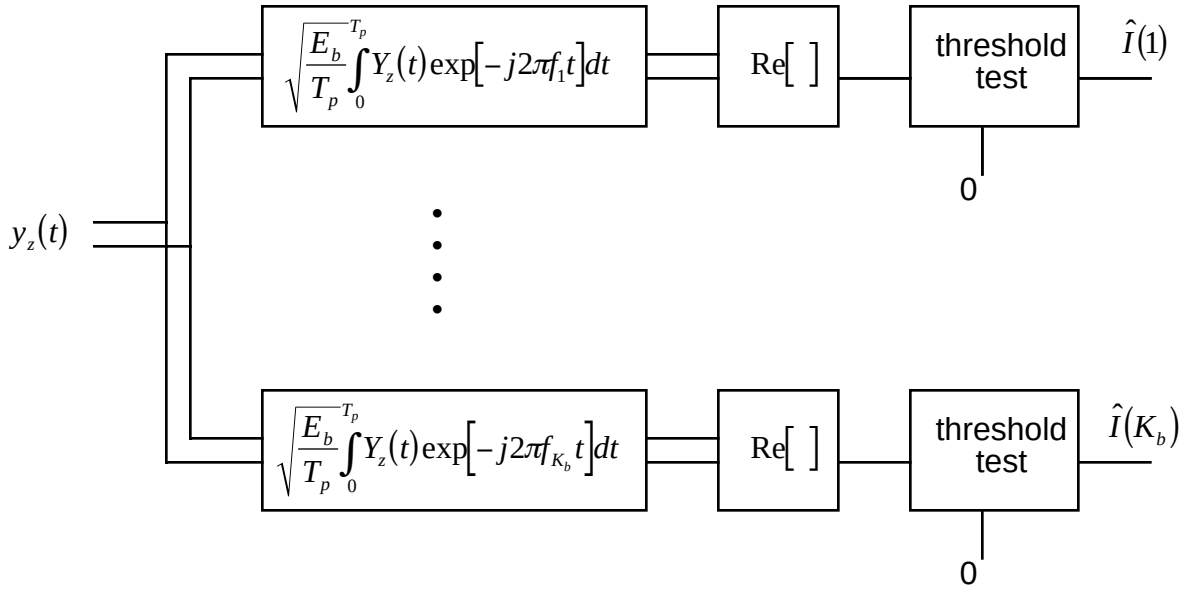


Figure 9.10: The optimal demodulator for OFDM using BPSK.

Spectral Characteristics of OFDM

Recall that the average energy spectrum per bit is

$$D_{x_z}(f) = \frac{E\left[G_{x_z}(f)\right]}{K_b}. \tag{9.47}$$

To simplify the notation needed in this discussion we will make the following definition.

Definition 9.1 *The Fourier transform of the unit energy rectangular pulse function*

$$u_r(t) = \begin{cases} \frac{1}{\sqrt{T_p}} & 0 \le t \le T_p \\ 0 & elsewhere \end{cases} \tag{9.48}$$

is

$$U_r(f) = \sqrt{T_p}\frac{\sin\left(\pi f T_p\right)}{\pi f T_p}\exp\left[-j\pi f T_p\right] \tag{9.49}$$

Taking the Fourier Transform of the OFDM signal (9.41) and using the frequency shift property of the Fourier transform gives

$$X_z(f) = \sum_{l=1}^{K_b} D_z(l)\sqrt{E_b}U_r\left(f - f_d(2l - K_b - 1)\right). \tag{9.50}$$

The energy spectrum is then given as

$$G_{X_z}(f) = E_b\sum_{l_1=1}^{K_b}\sum_{l_2=1}^{K_b} D_z(l_1)D_z^*(l_2)U_r\left(f - f_d(2l_1 - K_b - 1)\right)U_r^*\left(f - f_d(2l_2 - K_b - 1)\right). \tag{9.51}$$

The average energy spectrum per bit is then

$$D_{X_z}(f) = \frac{E_b}{K_b} E\left[\sum_{l_1=1}^{K_b}\sum_{l_2=1}^{K_b} D_z(l_1)D_z^*(l_2)U_r\left(f - f_d(2l_1 - K_b - 1)\right)U_r^*\left(f - f_d(2l_2 - K_b - 1)\right)\right]. \quad (9.52)$$

The average energy spectrum per bit can be greatly simplified when BPSK modulation is used on each frequency in OFDM. Recall each bit is assumed to be random and independently distributed. Examing the term corresponding to a fixed l_1 and l_2, $l_1 \neq l_2$, in (9.52) (one of K_b^2 terms in the summation) gives

$$\begin{aligned} & E\left[D_z(l_1)D_z^*(l_2)\right]U_r\left(f - f_d(2l_1 - K_b - 1)\right)U_r^*\left(f - f_d(2l_2 - K_b - 1)\right) \\ &= \left(\pi_0\pi_0 + \pi_1\pi_1 - \pi_0\pi_1 - \pi_1\pi_0\right)U_r\left(f - f_d(2l_1 - K_b - 1)\right)U_r\left(f - f_d(2l_2 - K_b - 1)\right) \\ &= \left(\pi_0 - \pi_1\right)^2 U_r\left(f - f_d(2l_1 - K_b - 1)\right)U_r\left(f - f_d(2l_2 - K_b - 1)\right) \end{aligned} \quad (9.53)$$

Examing the term corresponding to $l_1 = l_2$, in (9.52) gives

$$\begin{aligned} E\left[|D_z(l_1)|^2\right]\left|U_r\left(f - f_d(2l_1 - K_b - 1)\right)\right|^2 &= \left(\pi_0 + \pi_1\right)\left|U_r\left(f - f_d(2l_1 - K_b - 1)\right)\right|^2 \\ &= \left|U_r\left(f - f_d(2l_1 - K_b - 1)\right)\right|^2 \end{aligned} \quad (9.54)$$

The average energy spectrum per bit is then given as

$$\begin{aligned} D_{X_z}(f) &= \frac{E_b}{K_b}\sum_{l_1=1}^{K_b}\left|U_r\left(f - f_d(2l_1 - K_b - 1)\right)\right|^2 + \quad (9.55) \\ & \frac{E_b}{K_b}(\pi_1 - \pi_0)^2\sum_{l_2=1}^{K_b}\sum_{\substack{l_1=1 \\ l_1\neq l_2}}^{K_b} U_r\left(f - f_d(2l_1 - K_b - 1)\right)U_r^*\left(f - f_d(2l_2 - K_b - 1)\right). \end{aligned}$$

Further simplifications occur if each bit is equally likely. For equally likely bits the second term in (9.55) becomes zero. The spectrum in this case becomes

$$D_{X_z}(f) = \frac{E_b}{K_b}\sum_{l=1}^{K_b}\left|U_r\left(f - f_d(2l - K_b - 1)\right)\right|^2. \quad (9.56)$$

This spectrum is plotted in Fig. 9.11 for $K_b = 4$ and $f_d = 0.5/T_p$. The conclusions we can draw about the average energy spectrum of OFDM is that the bandwidth occupancy for this type of an OFDM is proportional to $1/T_p$ and to K_b. Recall that the transmission rate is $W_b = K_b/T_p$ bits/second. Consequently the transmission efficiency is in the neighborhood of 1 bit/s/Hz, though the exact number will be a function of how the engineering bandwidth is defined.

In conclusion OFDM provides a method to implement multiple bit transmission with good performance and reasonable complexity and spectral efficiency. The bit error probability performance of the OFDM scheme highlighted in this section gives the same bit error probability performance as BPSK used in isolation. The spectral efficiency is roughly 1 bit/s/Hz. The complexity of the optimum receiver is $O(K_b)$. Consequently OFDM has found significant utility in engineering practice.

9.3.2 Orthogonal Code Division Multiplexing

A second commonly used modulation that admits a simple optimal bit demodulation is orthogonal code division multiplexing (OCDM). Orthogonal code division multiplexing is often used in cellular radio

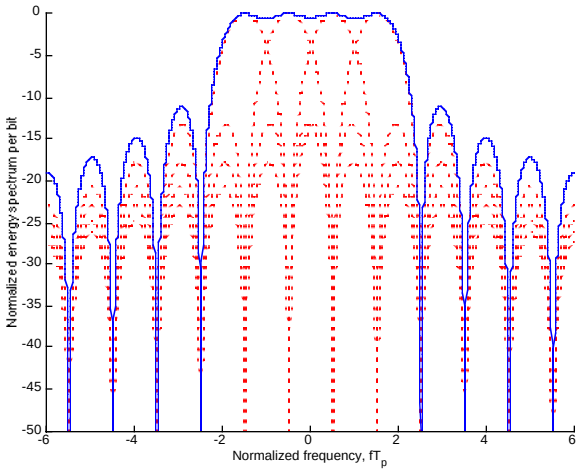


Figure 9.11: Average energy spectrum per bit for OFDM. $K_b = 4$, equally likely bits and BPSK modulation, with $f_d = 1/2T_p$.

communication and in satellite communication on the downlink. With OCDM each of the K_b bits is independently modulated on the same carrier frequency with an orthogonal waveform. This orthogonal waveform is often termed the spreading waveform. The typical format for an OCDM signal is

$$X_z(t) = \begin{cases} \sum_{l=1}^{K_b} D_z(l)\sqrt{E_b}s_l(t) & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{9.57}$$

where $D(l) = a\left(I(l)\right)$ and $s_l(t)$ is often denoted the spreading signal for the l^{th} bit. Here we assume that both $E\left[|D(l)|^2\right] = 1$ and that $E_{s_l} = 1$. The transmission rate of this form of OCDM is $W_b = K_b/T_p$ bits per second. In examining (9.57) it is clear that in this form of OCDM a binary linear modulation is used on each of the K_b different spreading waveforms. Again for clarity of discussion the remainder of the section will assume the linear modulation is BPSK (i.e., $a(0) = 1$ and $a(1) = -1$). It should be noted that OFDM is a special case of OCDM with $s_l(t) = \exp\left[j2\pi f_l t\right]/\sqrt{T_p}$.

In the general framework of orthogonal modulations OCDM has set

$$x\left(I(l), l, t\right)) = \begin{cases} D_z(l)\sqrt{E_b}s_l(t) & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{9.58}$$

The orthogonality condition then corresponds to

$$\Re\left[\int_{-\infty}^{\infty} x\left(m_l, l, t\right) x^*\left(m_k, k, t\right) dt\right] = E_b \Re\left[d_{m_l} d^*_{m_k} \int_0^{T_p} s_l(t) s^*_k(t) dt\right] = 0. \tag{9.59}$$

The remaining question is how to find a set of K_b orthogonal waveforms. There is a wide variety of ways to construct these waveforms and one example is given in Fig. 9.12 for K_b=4. Certainly this construction is not unique. The time waveform of OCDM can have significant variations across the transmission time. For instance Fig 9.13 shows the time plot of the output envelope[1] of two possible transmitted waveforms for the spreading waveforms introduced in Fig. 9.12 for K_b=4. In both cases

[1]Because the spreading waveforms are all real $x_A(t) = |x_I(t)|$.

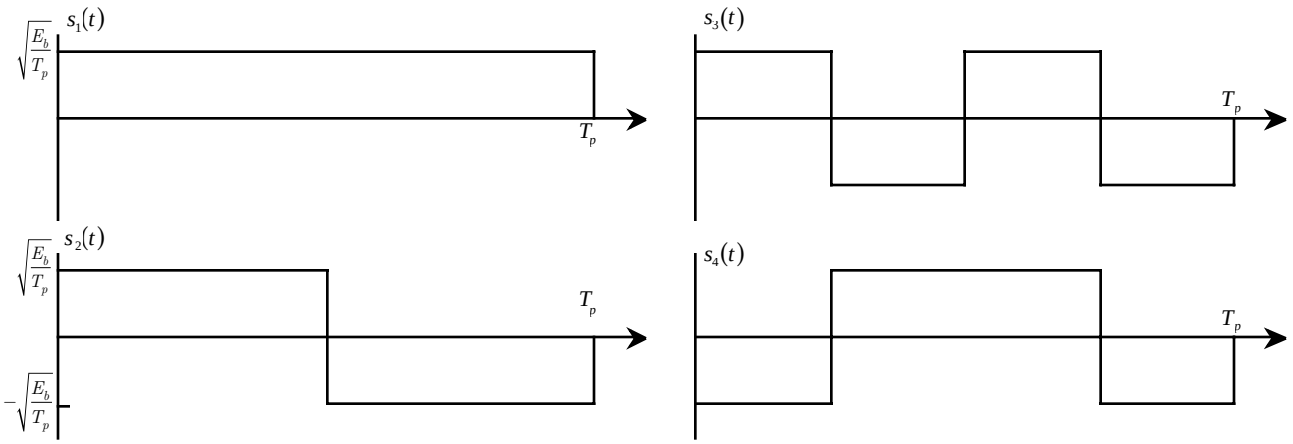


Figure 9.12: Example spreading waveforms for OCDM and $K_b = 4$.

the transmitted energy is equal but how that energy is spread across the transmit time is very much a function of the transmitted information. An OCDM waveform can have a significant difference between the peaks in amplitude and the average value of the amplitude. This difference will increase as K_b increases. This high peak to average power ratio requires the radios in an OCDM system to have the dynamic range necessary to process the signal without distortion.

Example 9.2: One of the standards for mobile telephones is denoted IS-95. IS-95 uses OCDM as the forward link modulation with $K_b = 64$. The spreading waveforms used to send each bit are Walsh functions which are discussed in Problem 9.22 (This problem shows how Walsh functions are derived from Hadamard matrices). In IS-95 $T_p = 52.08\bar{3}\mu$s so if BPSK modulation was used a bit rate of 1.2288Mbps could be achieved. In voice mobile telephony data rates only need to support voice band speech transmission. Hence IS-95 uses each of the spreading waveforms to transmit a bit to potentially 64 different users. This use of one physical layer radio link to support many information transactions is known as multiple access communications. Each spreading waveform is simultaneously capable of supporting a 19.2kbps transmission rate which is a data rate greater than necessary for modern speech encoders.

The optimal demodulator again has the form of K_b parallel single bit optimal demodulators. This demodulator is shown in Fig. 9.14. Restricting ourselves to BPSK modulation on each spreading waveform the MLBD has the form

$$\Re\left[\int_0^{T_p} Y_z(t)s_k^*(t)dt\right] = \Re\left[Q(k)\right] \underset{\hat{I}(k)=1}{\overset{\hat{I}(k)=0}{\gtrless}} 0 \tag{9.60}$$

The demodulator computes a filter output for each bit, $k \in \{1, \ldots, K_b\}$, and hence the complexity of the OCDM optimum demodulator is again $O(K_b)$ as opposed to the $O(2^{K_b})$ for an arbitrary modulation that transmits K_b bits of information. Since $\Re\left[Q(k)\right] = D_z(k)\sqrt{E_b} + N_I(k)$ where $\mathrm{var}\left(N_I(k)\right) = \frac{N_0}{2}$, the bit error probability performance of OCDM using BPSK on each spreading waveform is again

$$P_B(E) = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right). \tag{9.61}$$

The optimum demodulator for OCDM is again K_b single bit demodulators implemented in parallel (one for each orthogonal spreading waveform). Because of the orthogonality of the spreading waveforms the performance of an OCDM demodulator is the same a the demodulator for a single bit transmitted in isolation (see Chapter 7).

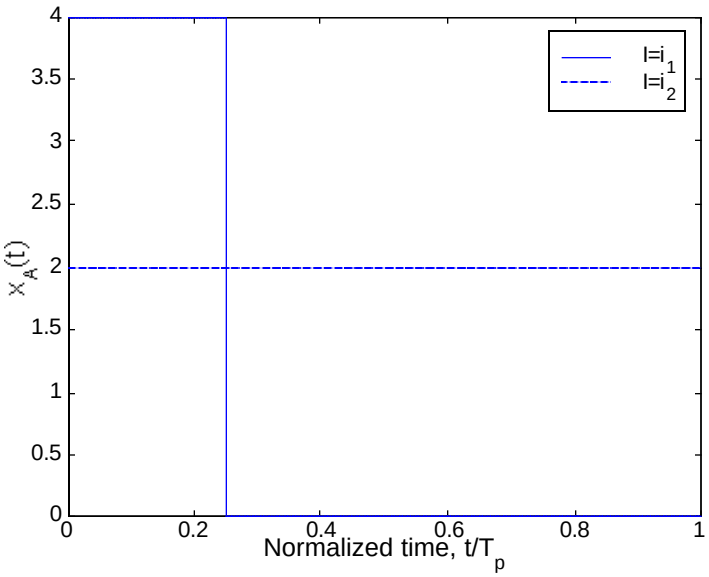


Figure 9.13: Output envelope time waveforms for the spreading waveforms in Fig. 9.12. $i_1 = [1110]$ and $i_2 = [1010]$

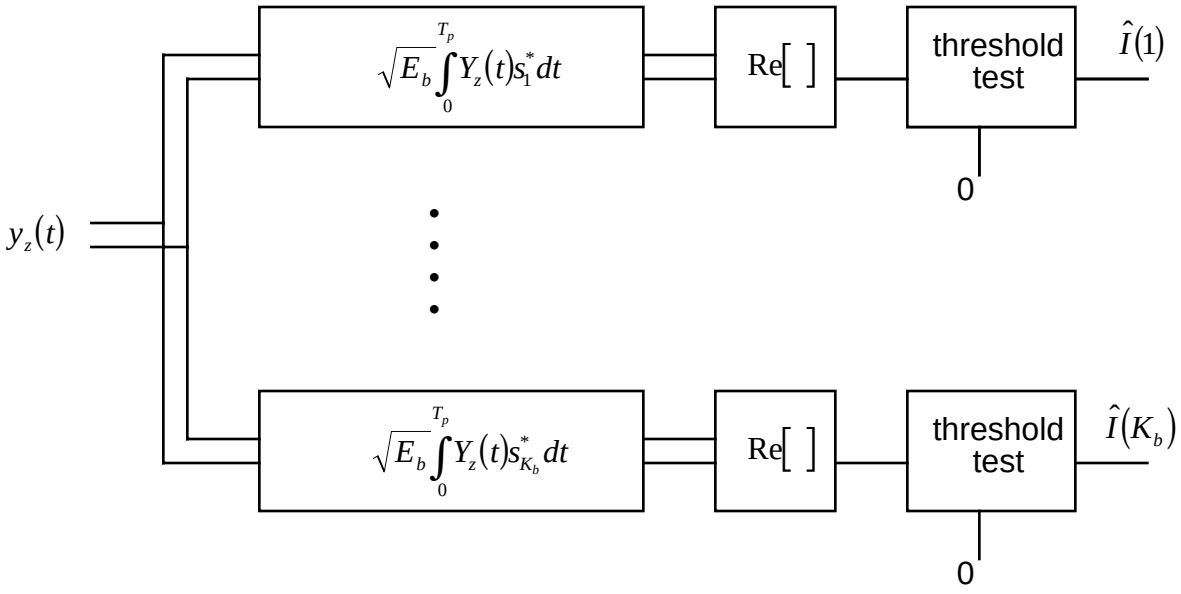


Figure 9.14: The optimal demodulator for OCDM using BPSK.

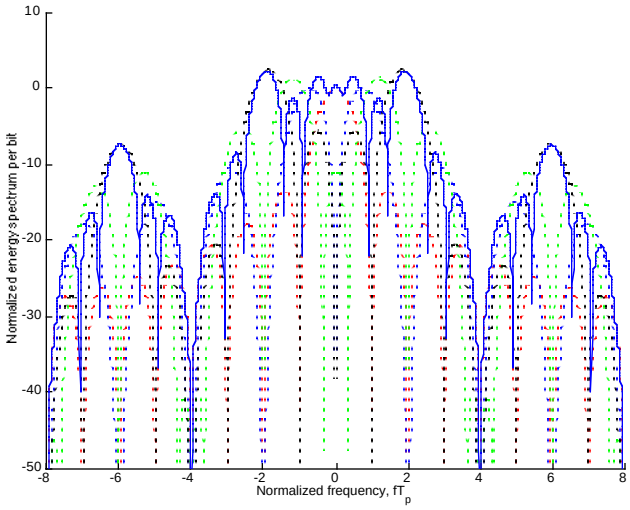


Figure 9.15: The average energy spectrum per bit for OCDM.

Spectral Characteristics of OCDM

Again the average energy spectrum per bit is

$$D_{x_z}(f) = \frac{E\left[G_{x_z}(f)\right]}{K_b} = \frac{E_b}{K_b} E\left[\sum_{l=1}^{K_b}\sum_{k=1}^{K_b} D_z(l) D_z^*(k) S_l(f) S_k^*(f)\right]. \tag{9.62}$$

Using the results of Section 9.3.1 and assuming for simplicity that each bit is equally likely and BPSK modulation is used, the spectrum in this case becomes

$$D_{X_z}(f) \quad = \quad \frac{E_b}{K_b}\sum_{l_1=1}^{K_b} \left|S_l\left(f\right)\right|^2 \tag{9.63}$$

where $S_l\left(f\right) = \mathcal{F}\left\{s_l(t)\right\}$. The important thing to note here is that the spectrum of OCDM waveforms is directly proportional to the spectrum of the chosen spreading waveforms. The spectrum of the set of spreading signals chosen in Fig. 9.12 is shown in Fig. 9.15. The conclusions we can draw about the average energy spectrum of OCDM is that the bandwidth occupancy for this type of an OCDM is proportional to $1/T_p$ and to K_b. The bandwidth is greater than if a single bit was sent in isolation since K_b orthogonal waveforms need to be constructed for each of the K_b transmitted bits.

9.3.3 Binary Stream Modulation

A third commonly used modulation that admits a simple optimal bit demodulation is orthogonal time division multiplexing. This is perhaps the most intuitive form of orthogonal modulation. Orthogonal time division multiplexing is used in a vast majority of digital communication systems in some form. The idea is simple, data is streamed in time (one bit following another) and for the remainder of this text this orthogonal time division multiplexing will be referred to as stream modulation. With a stream modulation each of the K_b bits is independently modulated on the same carrier frequency with a time

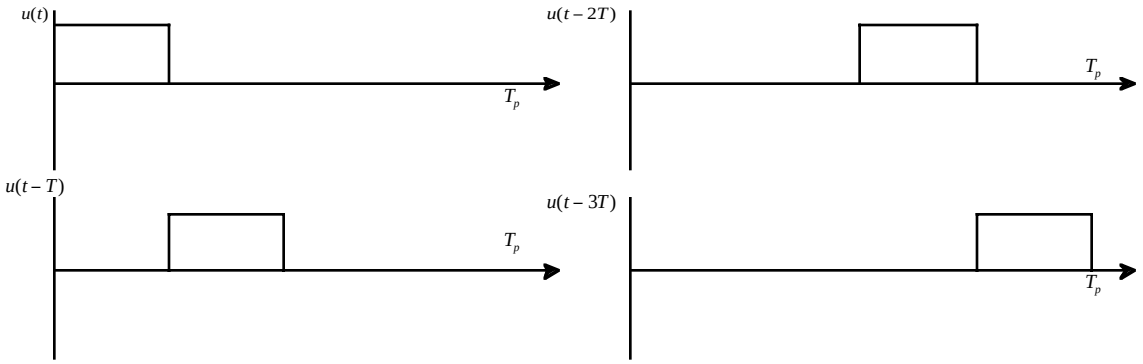


Figure 9.16: An example of time shifted pulses for a linear stream modulation and K_b=4.

shifted waveform. The typical format for a stream modulation using linear modulation is

$$X_z(t) = \sum_{l=1}^{K_b} D_z(l)\sqrt{E_b}u(t-(l-1)T) \tag{9.64}$$

where $D(l) = a\left(I(l)\right)$, T is known as the symbol or bit time in stream modulations, and $u(t)$ is the unit energy pulse of length T_u. The transmission rate of stream modulation is $W_b = K_b/T_p$ bits per second. It should be noted that $T_p = (K_b - 1)T + T_u$ and for large K_b that $W_b \approx 1/T$. In examining (9.64) it is clear that in stream modulation a binary linear modulation is used in each of the K_b different time intervals. Again for clarity of discussion the remainder of the section will assume the linear modulation is BPSK (i.e., $a(0) = 1$ and $a(1) = -1$).

In the general framework of orthogonal modulations stream modulation has set

$$x\left(I(l), l, t\right)) = \begin{cases} D_z(l)\sqrt{E_b}u(t-(l-1)T) & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{9.65}$$

where again $u(t)$ is denoted the pulse shape. The orthogonality condition then corresponds to

$$\begin{aligned} 0 &= \Re\left[\int_{-\infty}^{\infty} x\left(I(l), l, t\right) x^*\left(I(k), k, t\right) dt\right] \\ &= E_b\Re\left[D_z(l)D_z^*(k)\int_{-\infty}^{\infty} u(t-lT)u^*(t-kT)dt\right] = E_b\Re\left[D_z(l)D_z^*(k)V_u\left((k-l)T\right)\right]. \end{aligned} \tag{9.66}$$

This orthogonal time shift condition is often known as Nyquist's criterion for zero intersymbol interference (ISI) [Nyq28]. The remaining question is how to design orthogonal time shifted waveforms. There is a wide variety of ways to construct these waveforms but the simplest way is to limit $u(t)$ to only have support on $[0, T]$. For example if K_b=4 one can choose $T = T_p/4$ and have the set of time shifted waveforms as given in Fig. 9.16. Certainly this construction is not unique. It is interesting to note with the waveforms chosen in Fig. 9.16 the amplitude of the transmitted signal will be constant. The ability to more carefully control peak to average power ratio is one advantage of stream modulation.

The optimal demodulator has again has the form of K_b parallel single bit optimal demodulators. Since the sample modulation format is repeated in time the required filtering operation can also be serially with one filter sample serially. This demodulator is shown in Fig. 9.17. Restricting ourselves to

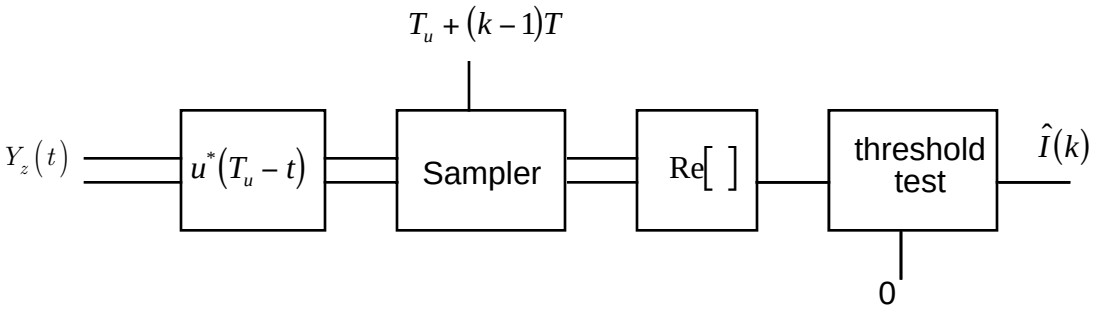


Figure 9.17: The optimal demodulator for stream modulations using BPSK.

BPSK modulation on each time shifted pulse the MLBD has the form

$$\Re\left[\int_0^{T_p} Y_z(t)u^*(t-(k-1)T)dt\right] = \Re\left[Q(k)\right] \underset{\hat{I}(k)=1}{\overset{\hat{I}(k)=0}{\gtrless}} 0 \tag{9.67}$$

The demodulator has one filter output whose output is sampled for each bit, $k \in \{1, \ldots, K_b\}$, and hence the complexity of the stream modulation optimum demodulator is again $O(K_b)$ as opposed to the $O(2^{K_b})$. The pulse shape matched filter output will be reference frequently in the sequel hence we formally define it.

Definition 9.2 *Recall the pulse shape matched filter has an impulse response $h_o(t) = u^*(T_u - t)$, so that the pulse shape matched filter output is*

$$Q_u(t) = \int_0^{T_p} Y_z(\lambda)u^*(\lambda + T_u - t)d\lambda. \tag{9.68}$$

where T_u is the pulse shape time support.

Using this definition it is easy to see that

$$Q(k) = Q_u(T_u + (k-1)T). \tag{9.69}$$

A characteristic of linear stream modulation that makes for an efficient implementation is that only one filter is needed to implement the demodulator. This pulse shape matched filter only needs to be sampled at different times to obtain the sufficient statistics for demodulation.

Finally, again the performance of stream modulation is upperbounded by performance on binary modulations transmitted in isolation. This is true since $\Re\left[Q(k)\right] = D_z(k)\sqrt{E_b} + N_I(k)$ where $\mathrm{var}\left(N_I(k)\right) = \frac{N_0}{2}$, the bit error probability performance of stream modulation using BPSK is

$$P_B(E) = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right). \tag{9.70}$$

Spectral Characteristics of Stream Modulations

Again the average energy spectrum per bit is

$$D_{x_z}(f) = \frac{E\left[G_{x_z}(f)\right]}{K_b} = \frac{E_b}{K_b}E\left[\sum_{l=1}^{K_b}\sum_{k=1}^{K_b} D_z(l)D_z^*(k)U(f)U^*(f)\exp\left[-j2\pi f(l-k)T\right]\right]. \tag{9.71}$$

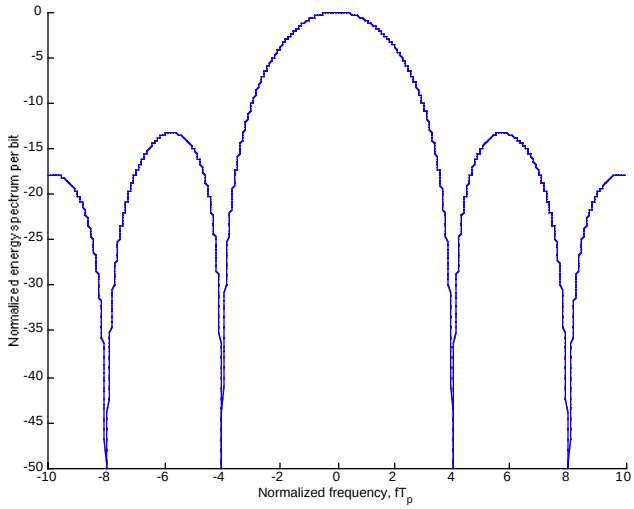


Figure 9.18: The average energy spectrum per bit for stream modulation.

Using the results of Section 9.3.1 and assuming for simplicity that each bit is equally likely and BPSK modulation is used, the spectrum in this case becomes

$$D_{X_z}(f) = \frac{E_b}{K_b} \sum_{l_1=1}^{K_b} |U(f)|^2 = E_b G_u(f) \tag{9.72}$$

where $U(f) = \mathcal{F}\{u(t)\}$ and $G_u(f) = |U(f)|^2$. The important thing to note here is that the spectrum of linear stream modulation is directly proportional to the spectrum of the chosen pulse shape. The spectrum of the linear stream modulation defined in Fig. 9.16 is shown in Fig. 9.18. The conclusions we can draw about the average energy spectrum of stream modulation is that the bandwidth occupancy is also proportional to $1/T_p$ and to K_b.

9.4 Conclusions

This chapter looked at two signal design techniques that simplify the optimal demodulator structures: linear modulation and orthogonal modulation. Linear modulation simplifies the demodulator by only requiring one matched filter. Linear modulation provides reduced performance but no bandwidth expansion as K_b grows larger. Consequently linear modulations often find utility when spectral efficiency is at a premium and the expected SNR is high. Orthogonal modulation is a technique where K_b bit are transmitted by using orthogonal waveforms for each bit. The orthogonality of the individual bit waveforms enables each bit to be detected optimally with a complexity identical to the situation where the bit was transmitted in isolation. Examples of orthogonal modulation included in this chapter were modulations where orthogonality was obtained by frequency spacing (OFDM), obtained by complex waveforms (OCDM), and obtained by time spacing (stream modulation). Both of these techniques address the exponential complexity of optimum demodulation by an appropriate signal design. Using signals designed as either a linear or an orthogonal modulation produces a performance upperbounded by the performance of BPSK. As a final note all deployed communication systems use stream modulation in some form even when using OFDM or OCDM. Also better than 90% of the digital communication systems use linear modulations in some form.

9.5 Homework Problems

Problem 9.1. 16QAM is a linear modulation with $D_i = D_{Ii} + jD_{Qi}$ where $D_{Ii}, D_{Qi} \in \{-3A, -A, A, 3A\}$ that can be used to transmit four bits per symbol (K_b=4).

a) Find the value of A such that E_b is the energy per bit.

b) Find the decision regions corresponding to the values of Q for MLWD of 16QAM.

c) Is 16QAM a geometrically uniform signal set?

d) Compute the probability of word error for MLWD, $P_W(E)$, for 16QAM and plot for E_b/N_0 =0, 15dB.

e) Compute the full union bound to the probability of word error for MLWD for 16QAM and plot for E_b/N_0 =0, 15dB. *Hint:* The union bound requires the computation of 15 distances for each of the 16 possible transmitted signals so it might be useful to use the computer.

f) It is possible to eliminate some terms in the union bound for 16QAM as was done for MPSK. Identify the union bound having the smallest number of terms for 16QAM.

g) Using the union bound result and assuming a pulse shape is chosen such that $B_T = 1/T_p$ plot the 16QAM spectral efficiency in Fig. 8.9.

Problem 9.2. Consider the two 8-ary (K_b=3) linear modulations whose constellations are given in Fig. 9.19. One of the 8-ary constellations has four points each placed on two concentric circles. The other 8-ary constellation has eight points selected from the 16QAM (see Problem 9.1) constellation in a checkerboard fashion.

a) Determine the value of each constellation point such that E_b is the energy per bit and $\Delta_1 = \Delta_2$.

b) Find the decision regions corresponding to the values of Q for MLWD of these two 8-ary modulations.

c) Which constellation has a better performance? Why?

Problem 9.3. This problem examines the probability of word error for MLWD of some common linear modulations.

a) Show that the probability of word error for 4PSK is given as

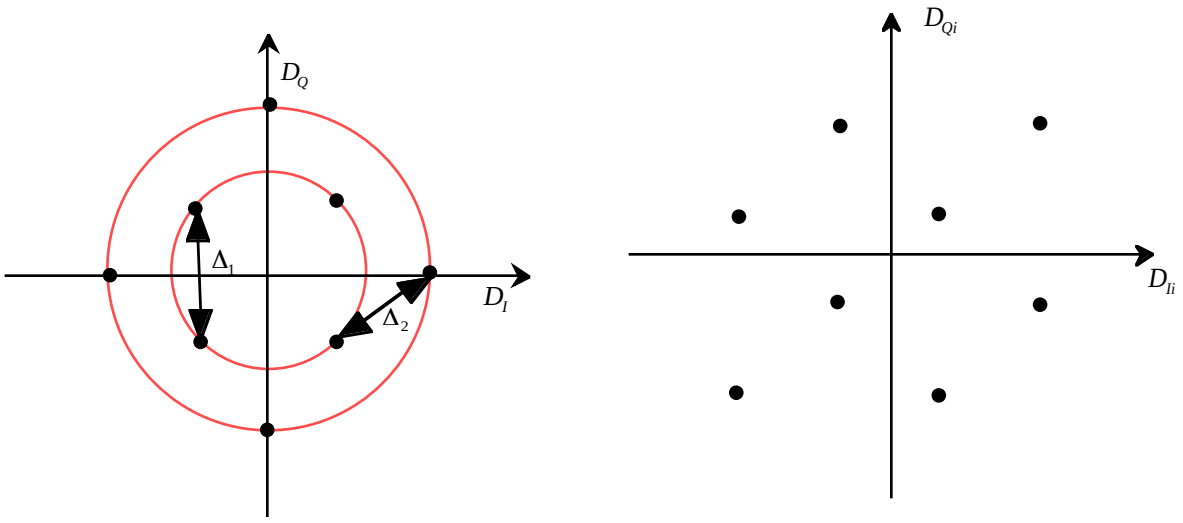
$$P_W(E) = \text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) - \left(\frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)\right)^2 \tag{9.73}$$

b) 4-ary equally spaced pulse amplitude modulation (PAM) has a constellation that is characterized with $\Omega_d = \left\{\pm\sqrt{2/5}, \pm 3\sqrt{2/5}\right\}$. Show that the probability of word error for 4PAM is given as

$$P_W(E) = \frac{3}{4}\text{erfc}\left(\sqrt{\frac{2E_b}{5N_0}}\right) \tag{9.74}$$

c) Plot the spectral efficiency versus E_b/N_0 of 4PAM on Fig. 8.9 using a word error rate of 10^{-5}.

Problem 9.4. This problem examines the full union bound to the probability of word error for MLWD of some common linear modulations.



a) b)

Figure 9.19: Two 8-ary constellation plots for linear modulations.

a) Show that the full union bound for 4PSK is given as

$$P_W(E) = \text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) + \frac{1}{2}\text{erfc}\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{9.75}$$

b) Is there a way to tighten the union bound for 4PSK?

c) Show that the full union bound for 4PAM ($\Omega_d = \left\{\pm\sqrt{2/5}, \pm 3\sqrt{2/5}\right\}$) is given as

$$P_W(E) = \frac{3}{4}\text{erfc}\left(\sqrt{\frac{2E_b}{5N_0}}\right) + \frac{1}{2}\text{erfc}\left(\sqrt{\frac{8E_b}{5N_0}}\right) + \frac{1}{4}\text{erfc}\left(\sqrt{\frac{18E_b}{5N_0}}\right) \tag{9.76}$$

d) Is there a way to tighten the union bound for 4PAM?

Show that these results are tight to the true values given in Problem 9.3 at high SNR. In each case the full union bound can be tightened. Derive the tightest union bound to the probability of error for both 4PSK and 4PAM.

Problem 9.5. Imagine that digital computers were first developed with five logic levels to match the number of fingers on the typical human hand. Digital communications would have evolved to be the transmission of quinits (having five possible values) as opposed to the transmission of bits. In this problem you will consider what this course might have been like if this alternate evolution had occurred.

Consider a 5-ary modulation of the form $X_z(t) = D_z u(t)$ corrupted by an additive white Gaussian noise with one-sided spectral density of N_0. Consider two possible signal sets:

i	$d_i^{(1)}$	$d_i^{(2)}$
0	1	A
1	$\exp\left(\frac{j2\pi}{5}\right)$	jA
2	$\exp\left(\frac{j4\pi}{5}\right)$	$-A$
3	$\exp\left(\frac{j6\pi}{5}\right)$	$-jA$
4	$\exp\left(\frac{j8\pi}{5}\right)$	0

where A is a positive real constant.

a) Choose the value of A such that both signal sets have the same average E_s.

b) Give the demodulator that minimizes the word error probability and define the decision regions for each of the two possible signal sets.

c) Are either signal sets geometrically uniform?

d) Compute the full union bound to $P_W(E)$ for each of the two signal sets.

e) Simplify the union bound as much as possible for each signal set by eliminating terms from the full union bound if possible.

f) Use numerical integration to compute the actual $P_W(E)$.

Problem 9.6. Consider one of the 5-ary modulations from the previous problem of the form $X_z(t) = D_z u(t)$ with equally likely symbols and corrupted by an additive white Gaussian noise with one-sided spectral density of N_0.

i	d_i
0	1
1	$\exp\left(\frac{j2\pi}{5}\right)$
2	$\exp\left(\frac{j4\pi}{5}\right)$
3	$\exp\left(\frac{j6\pi}{5}\right)$
4	$\exp\left(\frac{j8\pi}{5}\right)$

a) What is the average energy spectral density per quinit, $D_{x_z}(f)$?

b) If $u(t) = \sqrt{\frac{E_u}{W}}\frac{\sin(\pi W t)}{\pi W t}$ give $D_{x_z}(f)$?

c) If this particular communication system's electronics was particularly susceptible to 60Hz interference from the power supply, could you come up with a pulse shape that produced a communication signal that did not have significant spectral content at 60Hz. Assume that $T_p < 0.05$

Problem 9.7. A problem examining Alamouti space–time signaling [Ala98]. Consider a 4-ary digital communication system with equally likely words where the received signal is corrupted by an additive white Gaussian noise of two-sided spectral density of $N_0/2$. The system uses two antennas as shown in Fig. 9.20 and the four possible transmitted signals are shown in Fig. 9.21. The received signal is of the form

$$y_z(t) = c_1 x_{i1}(t) + c_2 x_{i2}(t) + W_z(t) = x_i(t) + W_z(t)$$

where c_1 and c_2 are known constants at the demodulator. The signal set is defined with

$$\begin{array}{llll} x_{01}(t) = u_1(t) & x_{11}(t) = u_2(t) & x_{21}(t) = u_4(t) & x_{31}(t) = u_3(t) \\ x_{02}(t) = u_2(t) & x_{12}(t) = u_3(t) & x_{22}(t) = u_1(t) & x_{32}(t) = u_4(t). \end{array}$$

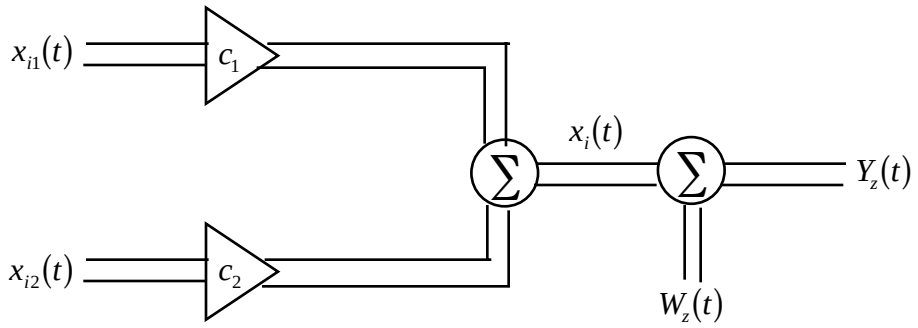


Figure 9.20: The two antenna communication system.

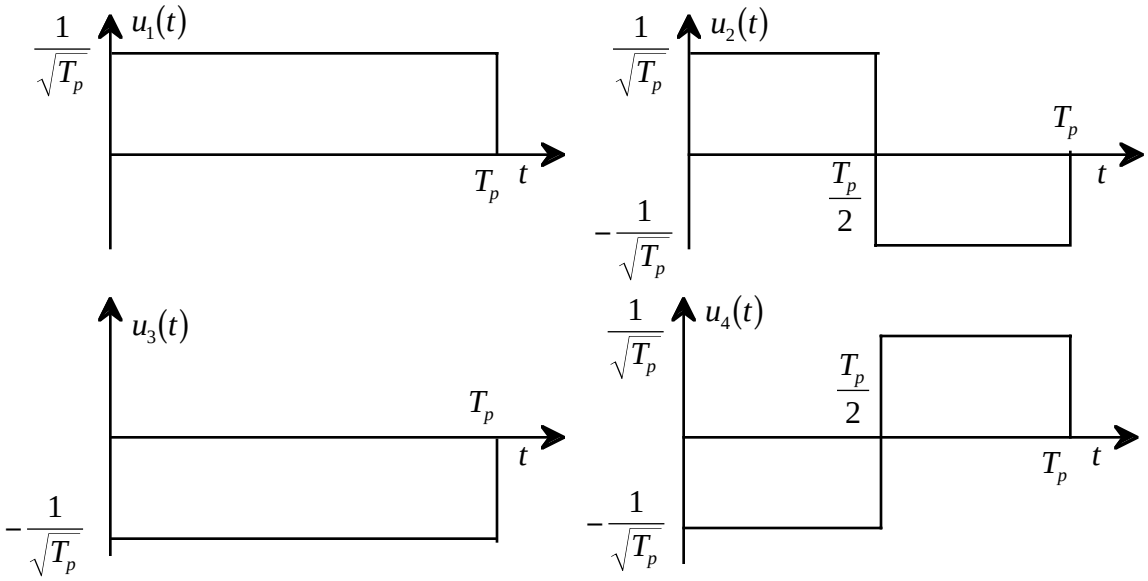


Figure 9.21: The transmitted waveforms.

a) Show that this system results in an equal received energy signal set for any values of c_1 and c_2. Compute E_s.

b) Detail out the optimum word demodulator (MLWD). Give impulse responses for any filters and simplify the structure as much as possible.

c) Compute the square Euclidean distance between each of the signal pairs.

d) Show that $x_i(t) = d_i(1)s_1(t) + d_i(2)s_2(t)$ where $s_1(t)$ is orthogonal to $s_2(t)$. This implies the maximum likelihood bit demodulator (MLBD) for each bit simplifies to a binary threshold test. Derive the MLBD for $I(1)$.

e) Compute the $P_B(E)$ for the MLBD for $I(1)$ and a given c_1 and c_2 conditioned on $\vec{I} = 0$.

Problem 9.8. Consider a demodulation scheme for a binary linear stream modulation where Nyquist's criterion for zero ISI holds. Prove that the real part of the noise sample taken from the matched filter

for bit $k_1,\ k_1 \in \{1,\ldots,K_b\}$, $N_I(k_1)$, is independent of the real part of the noise sample taken from the matched filter for bit $k_2,\ k_2 \in \{1,\ldots,K_b\}$, $N_I(k_2)$ where $k_1 \neq k_2$.

Problem 9.9. In Problem **??**.11 three pulse shapes are presented and this problem considers their use in a binary stream modulation of the form

$$X_z(t) = \sum_{l=1}^{K_b} x_{I(l)}(t-(l-1)T) \tag{9.77}$$

a) If $x_0(t) = u_1(t)$ and $x_1(t) = u_2(t)$ find the smallest value of T where Nyquist criterion is still satisfied.

b) If $x_0(t) = u_1(t)$ and $x_1(t) = u_2(t)$ plot the resulting transmitted waveform for $K_b = 10$ and $\vec{I} = [0\ 1\ 1\ 1\ 1\ 0\ 0\ 1\ 0\ 0]$ for the value of T selected in a).

c) If $x_0(t) = u_1(t)$ and $x_1(t) = u_2(t)$ plot the resulting matched filter output waveform both signals for $K_b = 10$ and $\vec{I} = [0\ 1\ 1\ 1\ 1\ 0\ 0\ 1\ 0\ 0]$ for the value of T selected in a).

d) Repeat a), b), and c) for the case when $x_0(t) = u_1(t)$ and $x_1(t) = u_3(t)$

Problem 9.10. In this problem we examine an M-ary linear wireless communication system that uses multiple transmit antennas. This problem assumes each of the L_t antennas simultaneously transmitts one bit of information. The complex envelope of the transmitted signal from the i^{th} antenna has the form $X_{zi}(t) = D_i u(t)$ where $D_i = a_i(I_i)$ is a complex modulation symbol with $|D_i| \leq 1$, I_i is the information bit transmitted from the i^{th} antenna, and $u(t)$ is the pulse shape. The constellation mapping is denoted with

$$a_i(0) = d_{0i} \qquad a_i(1) = d_{1i}. \tag{9.78}$$

The received signal has the form

$$Y_z(t) = \sum_{i=1}^{L_t} c_i X_{zi}(t) + W(t) \tag{9.79}$$

where c_i are the complex channel gains for the i^{th} antenna and $W(t)$ is a complex AWGN of one-sided spectral density of N_0, i.e. $R_W(\tau) = N_0\delta(\tau)$.

a) What is the bandpass signal corresponding to $X_{z1}(t)$ when a carrier frequency of f_c is used?

b) Show

$$Y_z(t) = D_z u(t) + W(t) \tag{9.80}$$

and find D_z. How many values can D_z take and what are they?

c) Assume $L_t = 2$ and that c_1 and c_2 are known at the receiver. Formulate the optimum demodulator and simplify as much as possible.

d) Assume BPSK modulation is used on each antenna $d_{0i} = 1$ and $d_{1i} = -1$ and compute the resulting union bound on the word error probability as a function of c_1 and c_2.

e) If it is known before transmission that $c_1 = 1$ and $c_2 = \frac{1}{\sqrt{2}}(1+j)$ postulate a binary modulation $D_i = a_i(I_i)$ for each antenna that would result in the best word error probability performance.

Problem 9.11. You have been asked to design a data communication system for the Zwertians. The Zwertians are a people from a distant solar system who have the unique characteristic of having 3 fingers. Consequently their number system is entirely base 3 and units of information are available in trinits (1 of 3 values). The Zwertians want a linear modulation, consequently a trinit is transmitted with the complex envelope

$$X_z(t) = D_z u(t) \tag{9.81}$$

where the pulse, $u(t)$, has energy $E_u = 1$. To complete your design you must specify an optimum demodulator and a 3-ary modulation scheme. Assume each value of the transmitted trinit is equally likely. The received signal is corrupted by an additive white Gaussian noise of one-sided spectral density N_0.

a) Assume the three possible transmitted symbols are denoted $d_i \quad i = 1, 2, 3$. Find the demodulator that minimizes the probability of symbol error.

b) Find the union bound to the symbol error in terms of N_0, and $d_i \quad i = 1, 2, 3$ for the optimum demodulator decision when $D_z = d_1$.

c) Assume that the average transmitted energy per symbol is constrained to be $E_s = 1$, find the optimum values of $d_i \quad i = 2, 3$ when $d_1 = 1$.

Problem 9.12. Show that if K_b bits are transmitted on orthogonal waveforms with BPSK modulation then the word error probability is given as

$$P_W(E) = 1 - \left(1 - \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)\right)^{K_b} \tag{9.82}$$

Using the characteristics of the binomial coefficients to arrange in (9.82) a sum like the union bound. From this sum can you identify the most probable error event for an orthogonal modulation and how bit errors many occur in the word. For example if $K_b = 4$ and $\vec{I} = [0\ 0\ 1\ 0]$ find all most probably error codewords.

Problem 9.13. Consider K_b=6 bits being transmitted with OFDM. Compute the output spectrum of the transmitted signal with the closest possible carrier spacing, sketch the optimum demodulator and estimate the resulting optimum word error probability performance for

a) 6 carrier frequencies each using BPSK modulation

b) 3 carrier frequencies each using Gray coded 4PSK modulation

c) 2 carrier frequencies each using 8PSK modulation.

Make a comparison between each of these transmission strategies in terms of performance and spectral efficiency

Problem 9.14. The OFDM transmitted signal is a sum of complex sinusoids and in general will be complex valued. In certain special cases the transmitted signal will be real valued only. When does this occur? Identify all the situations when this happens for K_b=4 and BPSK modulation on each carrier.

Problem 9.15. Identify two other spreading waveforms of equal energy that are mutually orthogonal for an OCDM system with K_b=3 besides

$$s_1(t) = \begin{cases} 1 & 0 \le t < \frac{T_p}{3} \\ \exp\left[\frac{j2\pi}{3}\right] & \frac{T_p}{3} \le t < \frac{2T_p}{3} \\ \exp\left[\frac{j4\pi}{3}\right] & \frac{2T_p}{3} \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{9.83}$$

A goal in finding these additional two waveforms is to make $D_{x_z}(f)$ as compact as possible.

Problem 9.16. Identify four spreading waveforms of length T_p that are equal energy and are mutually orthogonal for an OCDM system with K_b=4, and that use less transmission bandwidth than the example spreading waveforms shown in Fig. 9.12.

Problem 9.17. There is an alternate form for PWEP first introduced by Craig [Cra91] that proves to be useful in many situations and this problem will lead the student through the derivation. Consider a matched filter output Q for BPSK modulation when $D_z = a(1) = -1$ is transmitted. The probability of error in this case is given as (see Section 9.1.2)

$$P_B(E\,|I=1) = \int_{R_1} f_{N_z}(n_z)dn_z = \int 1_{R_1}(n_z) f_{N_z}(n_z)dn_z \tag{9.84}$$

where $R_1 = \{N_z : N_I \geq E_b\}$. Use the form $N_z(t) = N_A(t)\exp[jN_P(t)]$ to express the above integral in polar coordinates and derive

$$P_B(E\,|I=1) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \exp\left[\frac{-E_b}{N_0\cos^2(n_p)}\right] dn_p \tag{9.85}$$

Problem 9.18. The following is a pulse shape that is used in communication systems

$$u(t) = \begin{cases} \sqrt{\dfrac{2}{T_u}}\sin\left(\dfrac{\pi t}{T_u}\right) & 0 \leq t \leq T_u \\ 0 & \text{elsewhere.} \end{cases} \tag{9.86}$$

Consider that this pulse is used in a binary linear stream modulation of the form

$$X_z(t) = \sum_{l=1}^{K_b} D_z(l)\sqrt{E_b}u(t-(l-1)T) \tag{9.87}$$

where each bit value is equally likely and the received signal has the form $Y_z(t) = X_z(t) + W(t)$ with $W(t)$ is an AWGN with one-sided spectral density of N_0.

a) Compute $V_u(\tau)$.

b) How fast can the symbol or bit rate, $W_b = 1/T$, be and still have the MLBD simplify to a binary threshold test? Provide the form of the MLBD when BPSK modulation is used (i.e., $a(0) = d_0 = 1$ and $a(1) = d_1 = -1$).

c) What is the resulting performance of the optimal bit detector found in b).

d) Compute and plot $D_{x_z}(f)$.

At the maximum W_b for the waveform (9.87), a colleague of yours claims that you can double the transmission rate of the system in a) and still have optimal bit demodulators that are simple threshold tests. The claim is that this is achieved by sending a separate binary stream modulation in the Q-channel that is offset in time from the original one by $T_u/2$. The new transmitted signal (called minimum shift keying (MSK) by your colleague) has the form

$$\tilde{X}_z(t) = \sum_{l=1}^{K_b} D_{z1}(l)\sqrt{E_b}u(t-(l-1)T) + jD_{z2}(l)\sqrt{E_b}u\left(t-(l-1)T-\frac{T}{2}\right). \tag{9.88}$$

e) Show that $\tilde{X}_z(t)$ can be put in a form as given in (9.31). Check the orthogonality conditions to verify your colleague's assertion.

f) For the time interval $\frac{T}{2} \leq t \leq T$ find $\tilde{X}_A(t)$ and $\tilde{X}_P(t)$ as a function of $D_{z1}(1)$ and $D_{z2}(1)$.

g) Given the results in e) can the waveform in (9.88) be interpreted as an FSK waveform?

h) Compute $D_{x_z}(f)$ for MSK.

Problem 9.19. Consider a 4-ary modulation of the form $X_z(t) = D_z\sqrt{E_b}u(t)$ corrupted by an additive white Gaussian noise with one-sided spectral density of N_0. Consider two possible signal sets:

i	$d_i^{(1)}$	$d_i^{(2)}$
0	$\sqrt{2}$	$A\sqrt{3}$
1	$j\sqrt{2}$	jA
2	$-\sqrt{2}$	$-A\sqrt{3}$
3	$-j\sqrt{2}$	$-jA$

where A is a positive real constant.

a) Choose the value of A such that both signal sets have the same average E_s.

b) Give the demodulator that minimizes the word error probability and define the decision regions for each of the two possible signal sets.

c) Compute the union bound to $P_W(E)$ for each of the two signal sets.

d) Decide which signal set for a common E_s will give better performance with the optimum demodulator.

Problem 9.20. Show in OFDM that if $f_d = 1/T_p$ and K_b is odd that $X_z(0) = X_z(T_p)$. Show that this is not necessarily true for K_b even.

Problem 9.21. Show that if the orthogonality condition holds that the MLBD output and the MLWD output are exactly the same.

Problem 9.22. Define a 2×2 matrix

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{9.89}$$

and a recursion defined as

$$\mathbf{H}_{2^n} = \begin{bmatrix} \mathbf{H}_{2^{n-1}} & \mathbf{H}_{2^{n-1}} \\ \mathbf{H}_{2^{n-1}} & -\mathbf{H}_{2^{n-1}} \end{bmatrix}. \tag{9.90}$$

These sets of matrices are known as Hadamard matrices. Show that if $K_b = 2^n$ then the K_b spreading waveforms generated as

$$s_i(t) = \sum_{j=1}^{2^n} [\mathbf{H}_{2^n}]_{ij}\, u_{2^n}\left(t - (j-1)\frac{T_p}{2^n}\right) \qquad i = 1, 2^n \tag{9.91}$$

where

$$u_{2^n}(t) \;=\; \begin{cases} \frac{1}{\sqrt{T_p}} & 0 \leq t \leq \frac{T_p}{2^n} \\ 0 & \text{elsewhere} \end{cases} \tag{9.92}$$

constituted a set of spreading for an OCDM system. Specifically sketch the waveforms and show orthogonality for $K_b = 2$ and $K_b = 8$. For $K_b = 4$ this set of spreading waveforms is given in Fig. 9.12. A proof by induction for the general case would be most complete.

Problem 9.23. In the wireless local network protocol denoted IEEE 802.11b the lowest rate modulation is a binary modulation using the two waveforms given in Fig. 7.29.

a) If the two waveforms in Fig. 7.29 are to be used in a binary stream modulation, how fast can the symbol or bit rate, $W_b = 1/T$, be and still have the maximum likelihood bit demodulator simplify to a binary threshold test for each bit? *Hint: The answer is* $W_b > 1,000,000Hz$.

b) Assume the two waveforms in Fig. 7.29 were to be used in an orthogonal code division multiplexed system to send one bit. Find another spreading waveform, $s_2(t)$, for the case $K_b = 2$ defined on $[0, 1\mu s]$ with $E_{s_2} = E_b$ such that when BPSK modulation is used with $s_2(t)$ the maximum likelihood bit demodulator simplifies to a binary threshold test for each bit.

Problem 9.24. 8PAM is a linear modulation used to transmit $K_b = 3$ bits of information that is characterized with $\Omega_d \in \{\pm A, \pm 3A, \pm 5A, \pm 7A\}$. Assume the received signal is corrupted by an additive white Gaussian noise with one-sided spectral density of N_0.

a) For the case where $E_u = 1$, find the value of A such that $E_s = 3E_b$.

b) Define the decision regions for 8PAM that operate on the output of the pulse shape matched filter, Q.

c) Find the probability of word error conditioned on $D_z = A$.

d) A closed form expression exists for the probability of word error. Find it.

e) Define a mapping from bits $\vec{I} = [I(1)\ I(2)\ I(3)]$ to symbols that makes sense and use this mapping and the matched filter output, Q to find the MAPBD for $I(2)$.

Problem 9.25. 32QAM and 32Cross
Problem 9.26. Q^2PSK
Problem 9.27. I-Q Modulation with different rates on each channel.
Problem 9.28. OCDM with another spreading waveform.
Problem 9.29. Orthogonality in the frequency domain.
Problem 9.30. $K_b = 3$ compare and contrast 8FSK and 3 streamed symbols of BFSK.
Problem 9.31. QES98 with new pulse shape.

Chapter 10

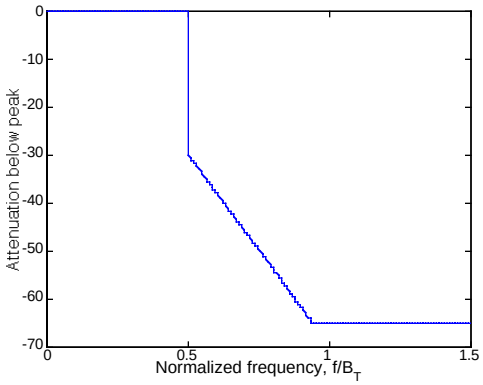
Spectrally Efficient Data Transmission

10.1 Spectral Containment

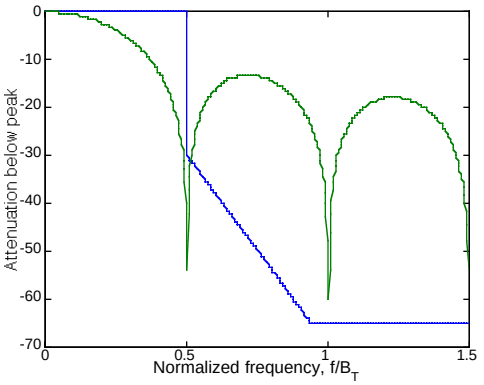
In digital communications it is often desirable to get many users in a band of frequencies. Examples include channels in broadcast or cable television and phone calls in mobile telephony. In this case it is often undesirable for emissions from one user to not interfere with the transmission from another user. In the modulations introduced in Chapter 9 the transmitted spectrum have significant out of band power due to the $\text{sinc}\,(fT_p)$ characteristic in the spectrum. This $\text{sinc}\,(fT_P)$ characteristic is due to the using of a rectangular pulse in each of these modulations. This out of band power can produce significant interference among users.

To prevent significant adjacent channel interference a spectral mask is often imposed on the transmission. Fig. 10.1-a) shows a typical spectral mask that must be met by transmitter electronics. Within the transmission bandwidth, B_T, the spectrum is not regulated. Outside the transmission bandwidth the output power is slowly tapered off until a large attenuation is achieved at the next adajacent channel. This mask is the spectral emissions mask for the narrowband radio services band as defined by the Federal Communications Commission [Com04]. Forcing a transmitter to obey this type of emissions mask allows more channels to operate with a better performance in a given frequency band. Meeting this emission mask will require a more sophisticated signal design than has been investigated up to this point in the text. For example Fig. 10.1-b) shows the mask in comparison to the spectrum of a digital modulation that uses a rectangular pulse shape as was used in Chapter 9. Clearly a rectangular pulse shape is not well motivated in this application due to the high sidelobes in frequency. These high sidelobes will cause significant interference to adjacent channel users. This is especially true when a near-far situation exists. Designing pulse shapes that meet a spectral mask (one for a mobile telephone) was investigated in Project 7-1. Meeting this emission mask and maintaining the orthogonality that enables low complexity demodulation as was introduced in Chapter 9 is the focus of this chapter.

Another situation that often arises is where the physical channel has significant spectral shaping characteristics. The best example of this is digital communications on analog telephone lines. The original analog telephone lines were designed to work with 4kHz baseband signals and to minimize the effects of aliasing there is a sharp cutoff filter in each phone line. Consequently it is important to find communication techniques that can accomodate this frequency selective characteristic of the channel. For example Fig. 10.2 shows what might be a typical communication system model. If the channel is known at the transmitter then this knowledge can be included as part of the signal design, i.e., the design problem then focuses on specifying the form of $X_z(t)$ or $X_c(t)$ to achieve a desired form for $R_z(t)$ or $R_c(t)$. Being able to design a communication system that can operate in that known environment is another goal of this chapter.



a) The mask.



b) The mask compared with an example $\text{sinc}\left(2f/B_T\right)^2$ spectrum.

Figure 10.1: A typical spectral emission mask.

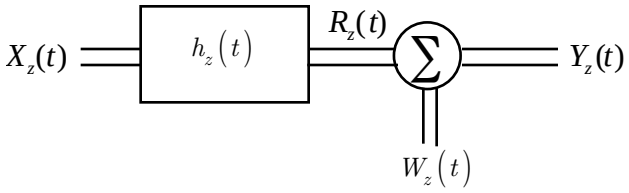


Figure 10.2: A frequency selective channel model.

10.2 Squared Cosine Pulse

The following Fourier transform pair is frequently used in bandwidth efficient communications

Definition 10.1 *The squared cosine pulse family is characterized as*

$$p_c(t) = \begin{cases} 1 & |t| \le \dfrac{(1-\alpha)T_z}{2} \\ \cos^2\left(\dfrac{\pi}{2\alpha T_z}\left(|t| - \dfrac{(1-\alpha)T_z}{2}\right)\right) & \dfrac{(1-\alpha)T_z}{2} \le |t| \le \dfrac{(1+\alpha)T_z}{2} \\ 0 & \textit{elsewhere} \end{cases} \tag{10.1}$$

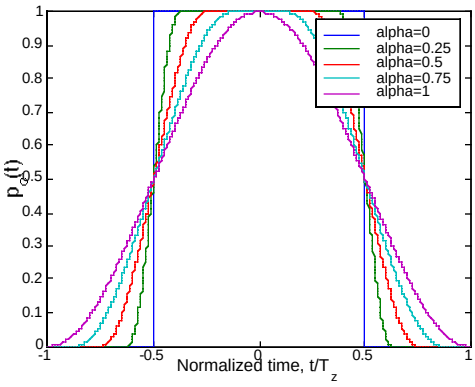
where $0 \le \alpha \le 1$. *The Fourier transform of the squared cosine pulse family is*

$$P_c(f) = \frac{\cos\left(\pi\alpha f T_z\right)}{1-\left(2\alpha f T_z\right)^2}\, T_z \text{sinc}\left(fT_z\right). \tag{10.2}$$

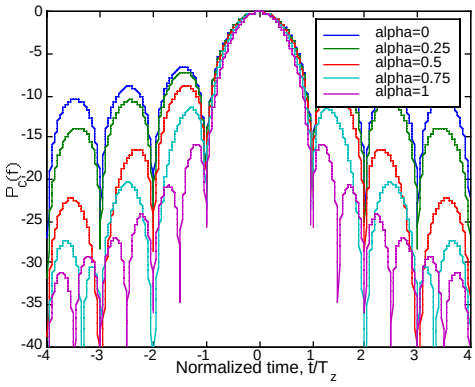
Fig.10.3 shows a plot of the time function and the Fourier transform. When $\alpha = 0$ then $p_c(t)$ becomes the rectangular pulse of width T_z. When $\alpha > 0$ then the Fourier transform dies off much quicker than the Fourier transform of the rectangular pulse. This characteristic is useful for making the transmissions more bandwidth efficient and meeting a spectral emissions mask. The squared cosine pulse shape has two additional important characteristics that make it useful for bandwidth efficient digital communications.

1. For $|t| \le T_z$ the squared cosine pulse satisfies

$$p_c(t) + p_c(t - T_z) + p_c(t + T_z) = 1. \tag{10.3}$$



a) $p_c(t)$



b) $P_c(f)$

Figure 10.3: The time and frequency characteristics of the squared cosine pulse.

2. $P_c\left(\frac{n}{T_z}\right) = 0$ for n an integer and $n \neq 0$.

The first characteristic indicates that the area under $p_c(t)$ is constant for any value of α even though the length of the pulse, $(1+\alpha)T_z$, is changing with α. This characteristic is due to the fact that $\cos(x - \pi/2) = \sin(x)$ and that $\cos^2(x) + \sin^2(x) = 1$. The second characteristic indicates the Fourier transform goes to zero at the same values of frequency as Fourier transform of the rectangular pulse. The zeros of the Fourier transform are due to the $T_z \text{sinc}\left(fT_z\right)$ term.

The squared cosine pulse family is often referred to as the raised cosine pulse in the literature. The reason for this can be seen in the case of $\alpha = 1$ where the pulse is given using a trigonometric identity as

$$p_c(t) = \begin{cases} \cos^2\left(\frac{\pi t}{2T_z}\right) & |t| \leq T_z \\ 0 & \text{elsewhere} \end{cases} = \begin{cases} \frac{1}{2} + \frac{1}{2}\cos\left(\frac{\pi t}{T_z}\right) & |t| \leq T_z \\ 0 & \text{elsewhere} \end{cases} \tag{10.4}$$

Why this text uses the notation squared cosine family should be apparent by the end of the chapter.

A second Fourier transform pair is also frequently used in bandwidth efficient communications

Definition 10.2 *The cosine pulse family is characterized as*

$$u_c(t) = \begin{cases} 1 & |t| \leq \frac{(1-\alpha)T_z}{2} \\ \cos\left(\frac{\pi}{2\alpha T_z}\left(|t| - \frac{(1-\alpha)T_z}{2}\right)\right) & \frac{(1-\alpha)T_z}{2} \leq |t| \leq \frac{(1+\alpha)T_z}{2} \\ 0 & \textit{elsewhere} \end{cases} \tag{10.5}$$

where $0 \leq \alpha \leq 1$. *The Fourier transform of the cosine pulse family is*

$$U_c(f) = \frac{\sqrt{T_z}(1-\alpha)\text{sinc}\left(fT_z(1-\alpha)\right)}{1-(4\alpha fT_z)^2} + \frac{4\sqrt{T_z}\alpha\cos\left(\pi fT_z(1+\alpha)\right)}{\pi\left(1-(4\alpha fT_z)^2\right)}. \tag{10.6}$$

This pulse is also often described in the literature as a square root raised cosine pulse. It should be noted that neither of the two important characteristics of the squared cosine pulse family hold true for the cosine pulse family.

10.3 Spectral Shaping in OFDM

OFDM is a modulation that independently modulates data bits on different carrier frequencies. A simple demodulator results if the modulation on each of the carrier frequencies are orthogonal. With the rectangular pulse shape assumed in Chapter 9 the orthogonality condition between the signal for carrier l and carrier k was shown to be

$$\Re\left[D_z(l)D_z^*(k)\frac{E_b}{T_p}\int_{-\infty}^{\infty}|u_r(t)|^2\exp\left[j4\pi f_d(l-k)t\right]dt\right]=0. \tag{10.7}$$

The spectrum is also a direct function of the shaping pulse. The spectrum in the case of equally likely bits independently modulated on each carrier becomes

$$D_{X_z}(f) \quad = \quad \frac{E_b}{K_b}\sum_{l=1}^{K_b}\left|U_r\left(f-f_d(2l-K_b-1)\right)\right|^2. \tag{10.8}$$

The desired goal is to identify a shaping pulse, $u_s(t)$, that both maintains the orthogonality condition of (10.7) and gives better out of band spectral emission characteristics than the rectangular pulse shape.

The results of Section 10.2 can be used to identify spectral efficient transmission strategies for OFDM. To generalize OFDM we can use an arbitrary pulse shape, $u_s(t)$. This implies that the transmitted signal has the form

$$X_z(t)=\sum_{l=1}^{K_b}D_z(l)\sqrt{\frac{E_b}{T_p}}u_s(t)\exp\left[j2\pi f_d(2l-K_b-1)t\right] \tag{10.9}$$

and the orthogonality condition becomes

$$\Re\left[D_z(l)D_z^*(k)\frac{E_b}{T_p}\int_{-\infty}^{\infty}|u(t)|^2\exp\left[j4\pi f_d(l-k)t\right]dt\right]=0. \tag{10.10}$$

Defining $p(t)=|u_s(t)|^2$ and restating the orthogonality condition as

$$\Re\left[D_z(l)D_z^*(k)P\left(2f_d(k-l)\right)\right]=0, \tag{10.11}$$

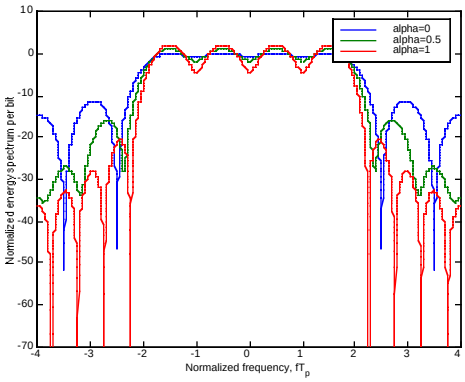
it is immediately obvious that orthogonality holds for an arbitrary pulse shape and an arbitrary constellation only if

$$P(2f_d(k-l))=0 \qquad \forall\ l\neq k. \tag{10.12}$$

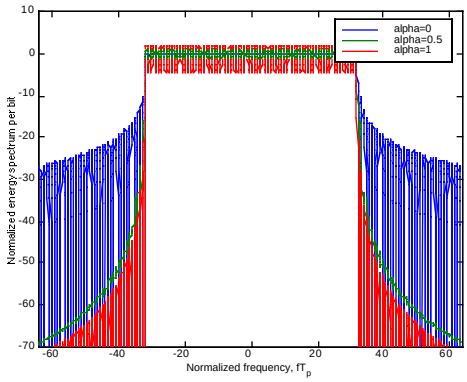
It is obvious considering the characteristics of the squared cosine pulse that a selection of $p(t)=p_c(t)$ with $T_z=1/(2f_d)$ will give the desired orthogonality. This implies that a bandwidth efficient implementation of OFDM has the form

$$X_{zs}(t)=\sum_{l=1}^{K_b}D_z(l)\sqrt{\frac{E_b}{T_p}}u_c(t)\exp\left[j2\pi f_d(2l-K_b-1)t\right]=u_c(t)X_{zu}(t). \tag{10.13}$$

where $X_{zu}(t)$ is the original OFDM modulation without shaping. Using this form of spectral shaping is convenient since it only has to be done on the transmitted signal after modulation. The smaller the value of α the smaller the region in time over which the the OFDM signal needs to be shaped. Fig. 10.4-a) show a plot of the output spectrum for OFDM for K_b=4 for various values of α. It should be noted that the cosine pulse shape does not allow a stringent spectral mask to be met with a small K_b, but with larger K_b the spectrum becomes quite compact. This can be observed in Fig. 10.4-b) which shows a plot of the output spectrum for OFDM for K_b=64 for various values of α.



a) $K_b = 4$.



b) $K_b = 64$.

Figure 10.4: The average energy spectrum per bit of OFDM for various raised cosine time pules.

10.4 Spectral Shaping in Linear Stream Modulations

Linear stream modulation is the most prevalent form of digital communications. Recall linear stream modulations have the form

$$X_z(t) = \sum_{l=1}^{K_b} D_z(l)\sqrt{E_b}u(t-(l-1)T) \tag{10.14}$$

The orthogonality condition that needs to be satisfied such that simple demodulation is possible is given as

$$\Re\left[D_z(l)D_z^*(k)V_u\left((k-l)T\right)\right] = 0. \tag{10.15}$$

The average energy spectrum per bit of linear stream modulation for the case of equally likely symmetric constellations was shown in Chapter 9 to be

$$D_{X_z}(f) = E_b\left|U\left(f\right)\right|^2 = E_bG_u(f) \tag{10.16}$$

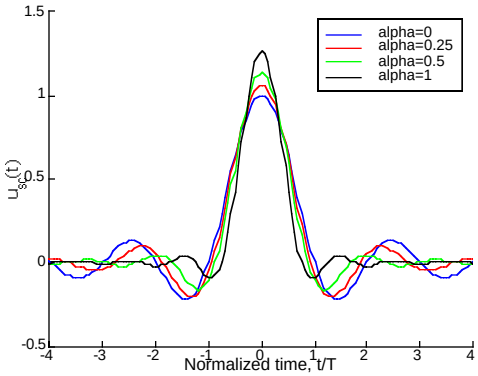
Putting together (10.15) and (10.16) shows that it is desired that a good pulse shape for linear stream modulations both must have a compact spectrum and an autocorrelation function that goes to zeros at integer multiples of the symbol time.

The square cosine function again has the desired characteristics but with time and frequency reversed. Using (10.2) we can set

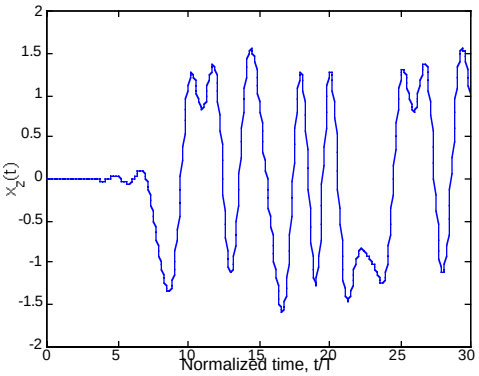
$$V_u(\tau) = \frac{\cos\left(\frac{\pi\alpha\tau}{T}\right)}{1-\left(\frac{2\alpha\tau}{T}\right)^2}\text{sinc}\left(\frac{\tau}{T}\right) \tag{10.17}$$

where $0 \leq \alpha \leq 1$. The $\text{sinc}\left(\frac{\tau}{T}\right)$ term in (10.17) is what ensures that Nyquist's criterion for zero ISI is satisfied. If the autocorrelation function of the pulse is chosen as in (10.17) then the average energy spectrum bit of the linear stream modulation is then given as

$$D_z(f) = E_bG_u(f) \;=\; \begin{cases} E_bT & |f| \leq \dfrac{1-\alpha}{2T} \\ E_bT\cos^2\left(\dfrac{\pi T}{2\alpha}\left(|f|-\dfrac{(1-\alpha)}{2T}\right)\right) & \dfrac{1-\alpha}{2T} \leq |f| \leq \dfrac{1+\alpha}{2T} \\ 0 & \text{elsewhere} \end{cases} \tag{10.18}$$



a) The spectral cosine pulse family



b) A stream modulation with a spectral cosine pulse and $\vec{I} = [11000100111010111100010]$

Figure 10.5: Spectrally efficient stream modulation using BPSK, $\alpha = 0.3$, and $T_u = 8T$.

This spectrum is clearly bandlimited (i.e., $B_T \approx \frac{1+\alpha}{T}$) and we denote the pulse/transform pair in (10.17) and (10.18) as a spectral squared cosine family of pulse shapes. In direct analog to the discussion the squared cosine pulse family this pulse pair as often denoted in the literature as a spectral raised cosine pulse family. α is usually denoted the excess bandwidth factor as it determines how much the ideal bandwidth expands from the minimum, $B_{Tmin} = \frac{1}{T}$.

The actual pulse used in the linear modulation is a spectral cosine pulse family. The Fourier transform of the pulse is given as

$$U_{sc}(f) = \begin{cases} \sqrt{T} & |f| \leq \dfrac{(1-\alpha)}{2T} \\ \sqrt{T}\cos\left(\dfrac{\pi T}{2\alpha}\left(|f| - \dfrac{(1-\alpha)}{2T}\right)\right) & \dfrac{(1-\alpha)}{2T} \leq |f| \leq \dfrac{(1+\alpha)}{2T} \\ 0 & \text{elsewhere} \end{cases} \tag{10.19}$$

The pulse itself is given as

$$u_{sc}(t) = \frac{1-\alpha}{\sqrt{T}} \frac{\mathrm{sinc}\left(\frac{(1-\alpha)t}{T}\right)}{1-\left(\frac{4\alpha t}{T}\right)^2} + \frac{4\alpha \cos\left(\frac{\pi(1+\alpha)t}{T}\right)}{\pi\sqrt{T}\left(1-\left(\frac{4\alpha t}{T}\right)^2\right)}. \tag{10.20}$$

Fig. 10.5-a) shows plots of the spectral cosine pulse shape for various values of α. It should be noted that these pulses have infinite time support but the pulses die off to zero at large arguments. The rate of decay of the pulses away from the peak of the pulse in this family increase with α. This infinite time support implies that approximations to these pulses must be used in practice. The simplest approach to implementation is to window the pulse in time to produce a finite time support. Again we will denote the truncated pulse length with T_u. The truncation implies that the orthogonality condition will be violated for some or all symbol time offsets. With an appropriate choice of window function, the amount of ISI produced can be made acceptably small. The design projects at the end of the chapter will explore windowing of a spectral cosine pulse for use in stream modulations. A typical time waveform for linear stream modulations is shown in Fig. 10.5-b).

Recall the demodulator for a linear stream modulation samples the output of a filter matched to the pulse shape for each transmitted symbol. The demodulator structure is shown in Fig. 10.6. Recall

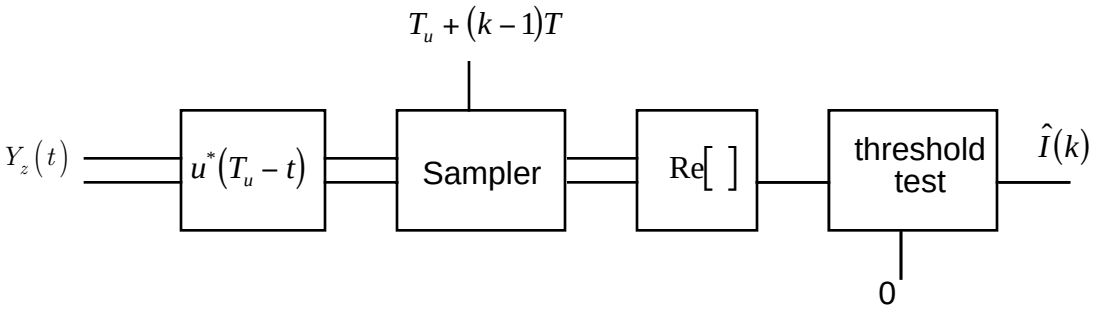
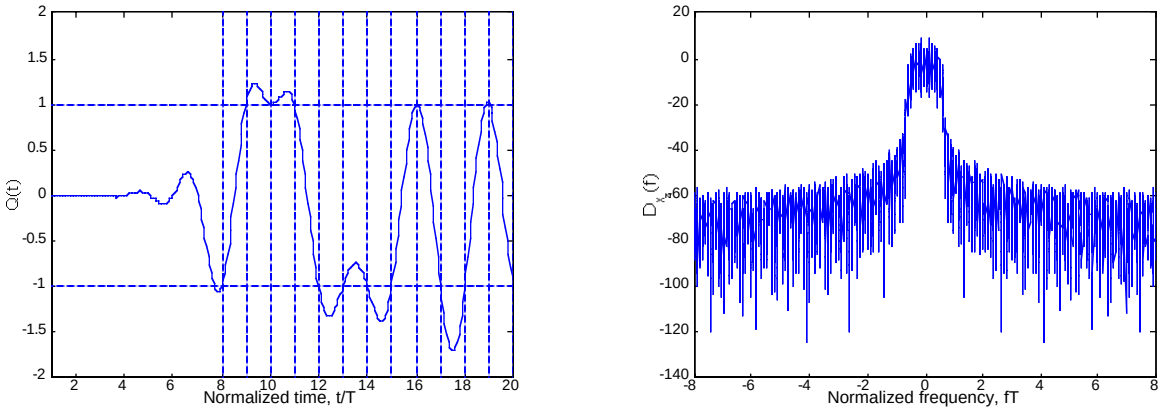


Figure 10.6: The optimal demodulator for stream modulations using BPSK.



a) The matched filter output, b) The measured transmitted spectrum.

Figure 10.7: Spectrally efficient stream modulation using BPSK, $\alpha = 0.3$, and $T_u = 8T$.

that the matched filter output has the form

$$Q(t) = \int_{-\infty}^{\infty} Y_z(\lambda) u^*(\lambda - t + T_u) d\lambda. \tag{10.21}$$

The matched filter output when sampled at $t = T_u + (k-1)T$ will have the form

$$Q(k) = D_z(k)\sqrt{E_b} + \sqrt{E_b} \sum_{l \neq k} D_z(l) V_u(k-l) + N_z(k). \tag{10.22}$$

It is clear that if Nyquist criterion holds for the pulse shape, $u(t)$, then the data bits can be detected without interference from other pulses. Fig. 10.7-a) shows the noiseless output of the matched filter for a spectrally efficient pulse where $T_u = 8T$ and $E_b = 1$. The vertical lines in Fig. 10.7-a) represent the sample times and the horizontal lines represent the modulation symbol values. It is apparent from this figure that even though 8 symbols are overlapped at anytime, designing the pulses to achieve the orthogonality condition allows each BPSK modulation symbol to be detected without interference from the other pulses. Using this shaping on the pulse also achieves the desired spectral efficiency as shown in Fig 10.7-b). This figure is a measurement of the spectrum of the transmitted signal and demonstrates the desired spectral containment.

10.5 Testing Orthogonal Modulations

While orthogonal modulations are used in a wide variety of applications, it is not always possible to keep the waveforms orthogonal at the receiver. The non-orthogonality of the received waveforms can be produced by a wide variety of distortions that arise in practice. A partial list of the causes of the loss in orthogonality is

1. A channel that is frequency selective,
2. Phase noise in the up and down converters in the radio system of a modem,
3. Nonlinear distortion in the radio system of the modem,
4. Complexity constraints on the modem implementation,
5. Noise in the sample timing location for the matched filter output in the modem,

In many situations this loss in orthogonality is significant enough to warrant a reformulation of the demodulation algorithms. This will be pursued in Chapter 11 for the case of frequency selective channels. In many other cases the goal is to characterize and minimize the distortion that causes the non-orthogonality. This section will introduce some common tools used by communication engineers to characterize the non-orthogonality of received waveforms in a demodulator.

When orthogonality is lost at the demodulator this results in interference from other symbols. Recall that OCDM, which is the most general case of orthogonal modulation, has a transmitted signal of the form

$$X_z(t) = \sqrt{E_b} \sum_{l=1}^{K_b} D_z(l) s_l(t). \tag{10.23}$$

For orthogonal modulations when the waveforms for each bit have maintained their orthogonality that the output of the matched filter for the k^{th} bit is

$$Q(k) = \int_0^{T_p} Y_z(t) s_l^*(t) dt = D_z(k)\sqrt{E_b} + N_z(k). \tag{10.24}$$

If distortion has happened in the communication system then the matched filter outputs will have the form

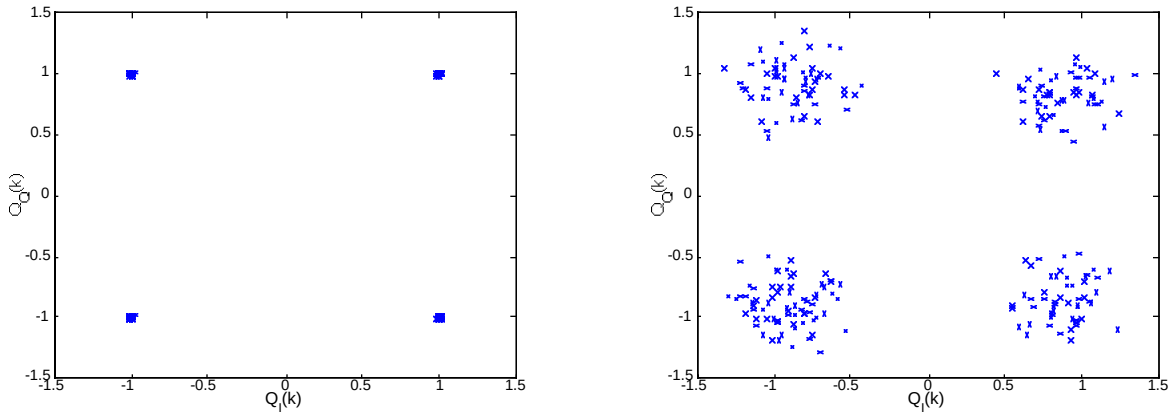
$$Q(k) = \int_0^{T_p} Y_z(t) s_l^*(t) dt = \sqrt{E_b} \sum_{l=1}^{K_b} D_z(l) g(k,l) + N_z(k). \tag{10.25}$$

Hence the distortion in the received waveform causes interference from other symbols in the matched filter output of the desired symbol. Measuring and characterizing the amount of intersymbol interference (ISI) is the goal of this section.

Recall that communication waveforms have three dimensions (I, Q and time) and most of the tools in use in practice are methods to represent the three dimensions of a communication waveform in two dimensions. Examples that we will explore here are

1. The scatter plot (I .vs. Q in the matched filter output),
2. The vector diagram for stream modulations,
3. The eye diagram (I .vs. time or Q .vs. time in the matched filter output) for stream modulation.

These three tools are used frequently in engineering practice.



a) Near ideal orthogonality.

b) Significant ISI.

Figure 10.8: Scatter plots for QPSK modulation.

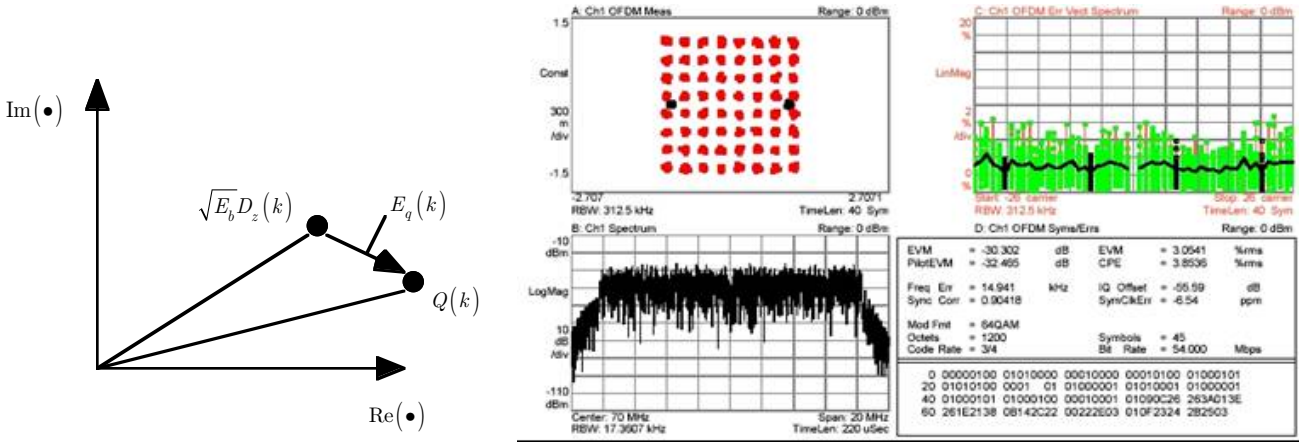
10.5.1 The Scatter Plot

The scatter plot simply plots the I and Q points of the matched filter output. In the absence of noise and ISI the scatter plot should just be a scaled constellation plot as in this case $Q(k) = \sqrt{E_b}D_z(k)$. For example the scatter plot of a QPSK modulation when the amount of non-orthogonality is small is shown in Fig. 10.8-a). A scatter plot for a case where some significant distortion is present in the received signal the is shown in Fig. 10.8-b). The ISI interference present in the matched filter output is manifested as a spreading of each of the constellation points. While the distortion shown in Fig. 10.8-b) appears to not be significant enough to cause errors in the absence of noise it should be apparent that the addition of noise with this distortion will significantly degrade the error rate performance. The scatter plot is a quick visual way for a communication engineer to get a handle on the amount of non-orthogonality in a modulation.

An often used measure of distortion for communication engineers is the error vector magnitude. The error vector for the matched filter output is defined to be $E_q(k) = Q(k) - \sqrt{E_b}D_z(k)$ and a vector diagram demonstrating this concept is shown in Fig.10.9-a). Communication engineers often quote the statistic of error vector magnitude (EVM). EVM is the average error vector magnitude compared to the largest normalized constellation point. Precisely EVM is defined as

$$\text{EVM} = \frac{\sum_{l=1}^{K_b} |E_q(l)|}{K_b \max_l \left|\sqrt{E_b}D_z(l)\right|} \tag{10.26}$$

and is reported in percentage. Most communication systems that use orthogonal modulations strive to achieve an EVM of less than 5%. Systems which use large constellations (64QAM or 256QAM) like digital television often have EVM requirements less than 3%. Most high performance vector analyzer type test equipment have automated EVM computations for common modulations used in accepted telecommunications standards. For example Fig. 10.9-b) shows an example test suite for a radio using one of the 64QAM modes of operation for the IEEE 802.11a standard. Recall as discussed earlier 802.11a is an OFDM based wireless modem and this figure shows a scatter plot for all subcarrier matched filter outputs, the received signal spectrum, and computations on the error vector for each subcarrier. This test suite is automatically produced by an Agilent vector signal analyzer and a great deal of information can be derived from this plot about the performance of a communication system using orthogonal modulation. For example a very useful tutorial article on how to use EVM to troubleshoot an orthogonal modulation system is [Agi00].



a) The error vector b) A test suite for an 802.11a

Figure 10.9: Testing using the error vector.

10.5.2 Stream Modulations

Since historically stream modulations have seen the most utility in practice several techniques have evolved to test the performance of orthogonal stream modulations. The two most common tests in stream modulation are the vector diagram plot and the eye diagram.

Vector Diagram

A generalization of the scatter plot is the vector diagram. The vector diagram was introduced in Chapter 2 as a way to visualize the 3D characteristics of the complex envelope in a 2D graph. For orthogonal stream modulation as with any IQ modulation the vector diagram can be used to get information about the vector modulated signal. For example Fig. 10.10 shows the transmitted vector diagram for a QPSK modulated stream modulation with a spectral cosine pulse with $\alpha = 0.3$ and $T_u = 8T$. A particular important vector diagram is the output of the matched filter to the transmitted pulse shape. This vector diagram with near ideal orthogonality should go through the constellation points at each sample time. Often in a vector diagram of the matched filter output, these sample times are marked with a different marker to emphasize these sample times. For example for a QPSK modulated stream modulation with a spectral cosine pulse with $\alpha = 0.3$ and $T_u = 8T$ with little loss in orthogonality the vector diagram of the matched filter output is shown in Fig. 10.11-a). Note that the matched filter output samples taken by themselves would be the scatter plots shown in Fig. 10.8. Distortion of the stream modulation signal will produce significant ISI and a vector diagram for a case of significant ISI is shown in Fig. 10.11-b). Vector diagrams for stream modulations give engineers insights into the performance of practical implementations of orthogonal stream modulation.

The Eye Diagram

The eye diagram is a technque to view the time waveforms out of the matched filter. Again the matched filter output is a three dimensional signal and the eye diagram is a method to visualize this three dimensional signal in two dimensions. The eye diagram is a repetitive plot of the matched filter output (either I or Q channel) over one symbol in time of the stream modulation. An example of the eye diagram for the case of near ideal orthogonality is shown in Fig. 10.12-a). The original method of producing an eye diagram goes back to the early days of digital communication when analog oscilloscopes were the primary time domain analysis tool available to a communication engineer. The I or Q channel of the matched filter output could be connected to the oscilloscope and the symbol clock could be

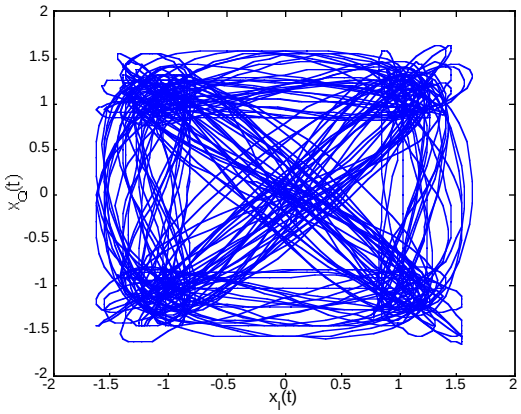
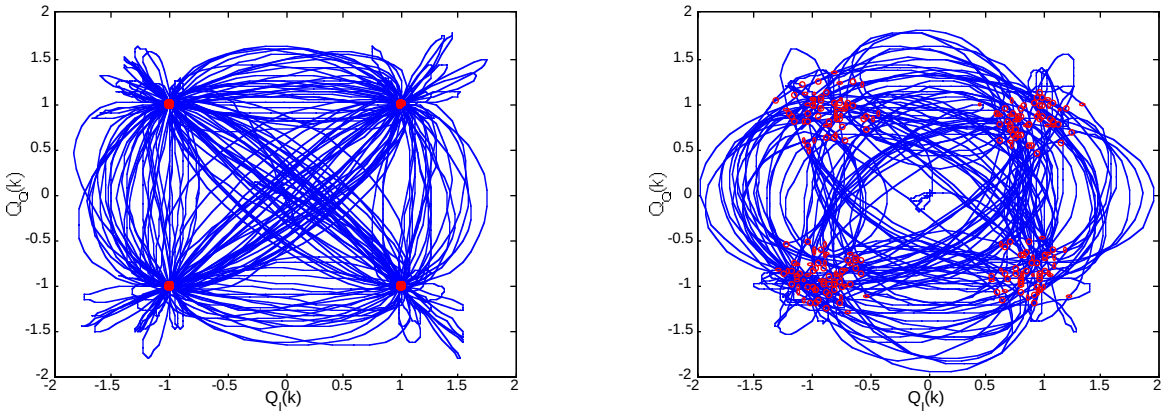
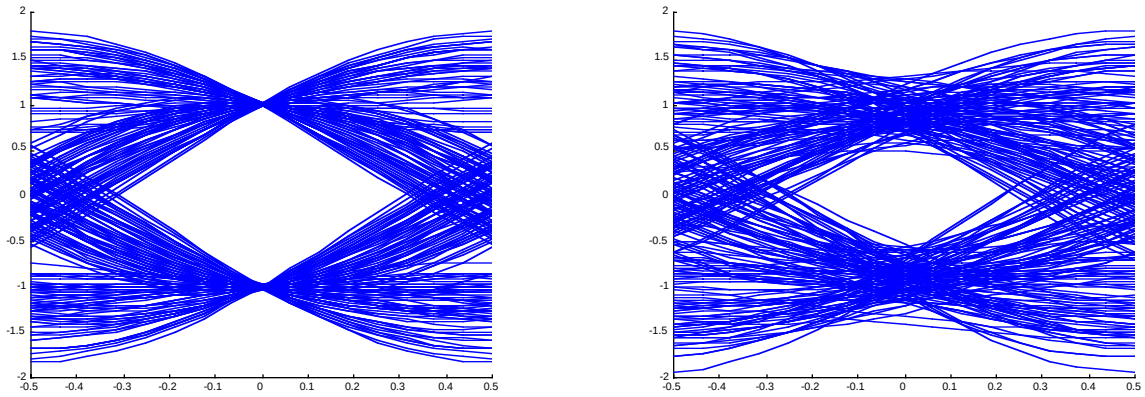


Figure 10.10: Vector diagram for a transmitted signal with QPSK, $\alpha = 0.3$ and $T_u = 8T$.



a) Near ideal orthogonality b) Significant ISI

Figure 10.11: Matched filter output vector diagrams for a spectral cosine pulse with $\alpha = 0.3$ and $T_u = 8T$. Sample times of the matched filter output denoted with a circle.



a) Near ideal orthogonality b) Significant ISI

Figure 10.12: Matched filter output eye diagrams for a spectral cosine pulse with $\alpha = 0.3$ and $T_u = 8T$.

used as the trigger and an eye diagram would then be visible on the display. The eye diagram shows the possible transitions from one symbol to the next and how impacted the orthogonality is by any distortion. An example of the eye diagram for the case of significant ISI is shown in Fig. 10.12-b). The ISI causes the space between the worst case sample points of the matched filter outputs to move closer to the decision threshold. In the case of BPSK modulation as shown in Fig. 10.12 the threshold will be zero. This effect of the ISI moving the optimum sample point closer to the decision boundary is often referred to as the "closing" of the eye. A communication system is often referred to as having a closed eye if there are some ISI patterns which would cause a deterministic error to occur in the simple threshold test demodulator of orthogonal demodulators. Most communications system analysis tools have the capability to produce eye diagrams and vector plots since they prove useful to communication engineers. For example Matlab has an `eyediagram` command. The Agilent vector analyzer can produce eye diagrams for most standard stream modulations and an example of both an eye diagram and a vector diagram for a 16QAM stream modulation is shown in Fig. 10.13. The tools of the eye diagram and the vector plot produce a great deal of insight into the implementation of orthogonal stream modulations and are often used by communication system engineers.

10.6 Conclusions

This chapter has looked at some practical aspects of digital communications using orthogonal modulations. Techniques to shape the spectrum of orthogonal modualtions in response to requirements often mandated in real implementations were overviewed. The concept of the cosine pulse that is popular in practice was introduced. How this cosine pulse can be used in orthogonal modulations was briefly summarized. Finally, a brief overview of tools used in the industry to characterize orthogonal modulations was given. Understanding the concepts in this chapter will help the student get closer to understanding how digital communications is implemented in modern systems.

10.7 Homework Problems

Problem 10.1. Prove

$$V_u(mT) = \begin{cases} E_u & m = 0 \\ 0 & \text{elsewhere} \end{cases} \tag{10.27}$$

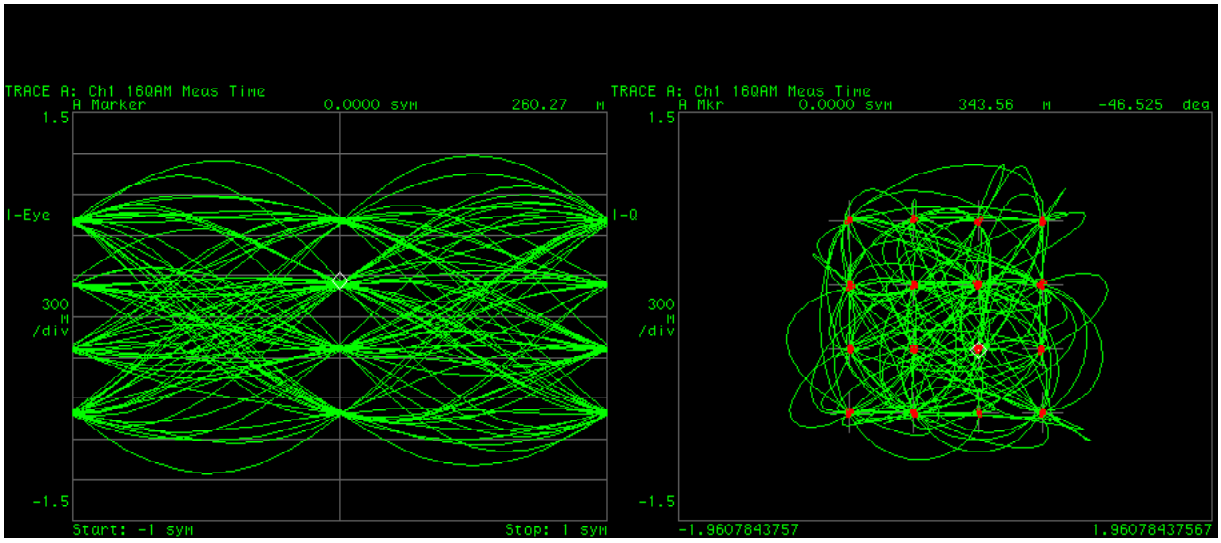


Figure 10.13: Test results for an orthogonal stream modulation from an Agilent vector signal analyzer. 16QAM with $\alpha = 0.25$.

if and only if

$$\sum_{k=-\infty}^{k=\infty} G_u\left(f + \frac{k}{T}\right) = TE_u. \tag{10.28}$$

Hint: Note the proof is based on Nyquist sampling theorem. This property indicates ways to design other pulses besides the spectral cosine pulse family to satisfy the Nyquist criteria for zero ISI.

Problem 10.2. Prove the smallest bandwidth that can be achieved in linear stream modulation is $B_T = 1/T$ where $D_z(k)$ can have an arbitrary constellation. If $D_z(k)$ is restricted to be real valued show the bandwidth can be $B_T = 1/(2T)$.

Problem 10.3. Prove that $u_{sc}(t) = \mathcal{F}^{-1}\{U_{sc}(f)\}$

Problem 10.4. The company you work for, Horizon Wireless, has purchased 1MHz of spectrum from the federal government of Elbonia. The marketing group of Horizon Wireless in Elbonia has decided that the way to make money in Elbonia is to divide the purchased spectrum into 20 equal size channels (of 50kHz) and sell radios that use these 20 channels to the Elbonian government for use by their diplomatic corp.

a) Give an example of a modulation that will achieve a 40kHz transmission rate in a spectrally efficient and power efficient manner on one of these channels while sending packets of $K_b = 256$ bits and achieving a demodulation complexity that is reasonably close to $O(256)$.

b) Give an example of a modulation that will achieve a 120kHz transmission rate in a spectrally efficient and power efficient manner on one of these channels while sending packets of $K_b = 768$ bits and achieving a demodulation complexity that is reasonably close to $O(256)$.

c) Assume the propagation loss in the channel is $L_p = aR^{-2}$ where R is the range of communication and a is a constant and the transmitted power remains the same for both transmission rates. If your design in part a) can achieve a range of 1km with a low frame error rate, what range will your design in b) achieve and give the same frame error rate performance. (assume the noise stays constant).

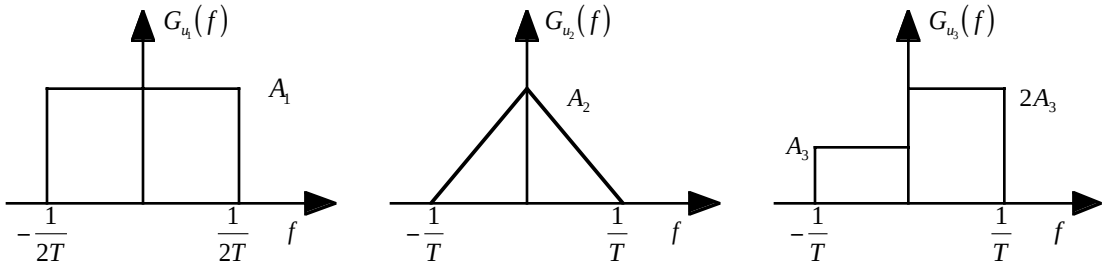


Figure 10.14: The energy spectrum of three possible pulses for linear stream modulation.

Problem 10.5. The energy spectrum of three pulse shapes that can be used in a linear stream modulation is shown in Fig. 10.14. Recall linear stream modulation has the form

$$X_{z,i}(t) = \sqrt{E_b} \sum_{l=1}^{K_b} D_z(l) u_i \left(t - (l-1)T\right). \tag{10.29}$$

where $u_i(t)$ corresponds to one of the three possible pulse shapes.

a) Find the amplitude A_i for $i = 1, 2, 3$ such that $E_u = 1$.

b) Prove that each of these three pulses can be used in a stream modulation and have simple bit demodulation result. *Hint:* These pulses have the same characteristic as the squared cosine pulse that makes the squared cosine pulse shape useful for stream modulation.

c) Find the form of $u_3(t)$.

d) Which pulse shape will provide the best spectral efficiency if the measure of bandwith is the 10dB bandwidth, B_{10}?

Problem 10.6. Your boss has tasked you with designing an OFDM cable modem to send packets of $K_b = 256$ bits in less than 300μs. She wants you to use less than 1 MHz of spectrum and cause little interference to transmissions from adjacent houses (each house is frequency multiplexed). The cable is assumed to be a frequency flat device and the goal given by your boss is to have a simple demodulator. Specify all the important parameters and pulse shapes.

10.8 Example Solutions

Not completed this edition

10.9 Mini-Projects

Goal: To give exposure

1. to a small scope engineering design problem in communications
2. to the dynamics of working with a team
3. to the importance of engineering communication skills (in this case oral presentations).

Presentation: The forum will be similar to a design review at a company (only much shorter) The presentation will be of 5 minutes in length with an overview of the given problem and solution. The presentation will be followed by questions from the audience (your classmates and the professor). All team members should be prepared to give the presentation.

10.9.1 Project 1

This project addresses the design of a pulse shape to be used with linear modulations that will meet the spectral emissions mask given in Fig. 10.15 and be nearly optimally demodulated by a matched filter. In other words the signal will have the form

$$X_z(t) = \sum_{k=1}^{K_b} D_z(k)u(t-(k-1)T) \tag{10.30}$$

You are further constrained in that the pulse, $u(t)$, must not extend for longer than $T_u = 40$ μs in time. This spectral emissions mask is the one used for GSM/EDGE handsets. So this problem is one of great practical interest. Design a pulse shape that will satisfy Nyquist criterion for an arbitrary linear modulation.

For the remainder of the project it is assumed that the pulse shape designed above is used with BPSK modulation. Any finite length pulse shape which achieves this spectral mask will have produced some intersymbol interference (ISI) (i.e., Nyquist criterion cannot exactly be met). Hence the matched filter demodulator given in Fig. 10.16 will not be exactly optimal but can be made to be very close to optimal with a well designed pulse shape.

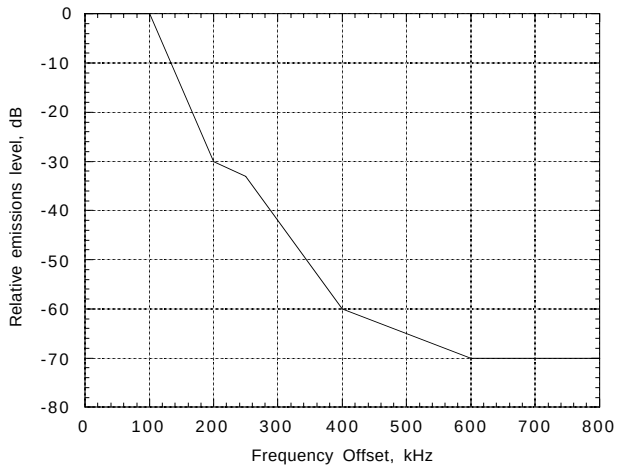


Figure 10.15: Spectral emissions mask for a GSM mobile.

a) Assume $D_z(k)$=1, compute the worse case degradation for an arbitrary pulse shape to the bit error probability for the demodulator given in Fig. 10.16. Specifically consider the case of the kth bit where $k = K_b/2$ where K_b is large and compute the effects of ISI. Identify the values of $D_z(j),\ j \neq k$ which achieve this worse case degradation for an arbitrary pulse shape with BPSK modulation.

b) Design a pulse shape and corresponding symbol rate such that the worse case (over all ISI patterns) degradation to the effective SNR is less than 0.25dB and the spectral mask given in Fig. 10.15 is

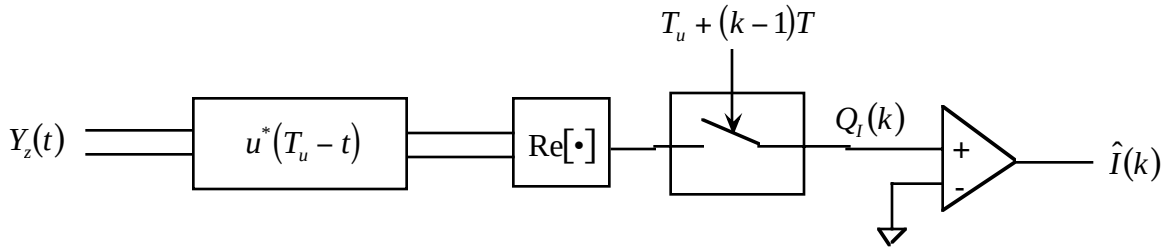


Figure 10.16: The baseband demodulator for binary pulse amplitude modulation.

met by the transmission. Grades will be assigned proportional to the achieved symbol rate (higher the better!). Anyone who beats the posted solution will get a bonus of 10 points for the project!

c) Plot the transmitted complex envelope for the pulse shape that was designed in b). Plot the real component of the matched filter output for a modulation sequence of $[1-111-11-1-1-11]$ for your design. Identify the symbol time sampling points.

Your grade will be penalized if a common solution to yours is found among your classmates solutions in direct proportion to the number of identical solutions. This is to encourage each of you to work independently. **Any computer code (e.g., Matlab) should be turned in with the project write-up.**

10.9.2 Project 2

This project addresses the design of a pulse shape to be used with a linear modulations where the constellation only takes real values that will meet the spectral emissions mask given in Fig. 10.15 and be nearly optimally demodulated by a matched filter. The pulse shape designed for a real valued constellation can be spectrally more efficient than the pulse design for the arbitrary constellation. *Hint: Use the $\Im\{V_u(\tau)\}$ to improve the spectral efficiency.*

The signal will have the form

$$X_z(t) = \sum_{k=1}^{K_b} D_I(k)u(t-(k-1)T) \tag{10.31}$$

You are further constrained in that the pulse, $u(t)$, must not extend for longer than 80μs in time. This spectral emissions mask is the one used for GSM/EDGE handsets. So this problem is one of great practical interest.

For the remainder of the project it is assumed that the pulse shape designed above is used with BPSK modulation. Any finite length pulse shape which achieves this spectral mask will have produced some intersymbol interference (ISI) (i.e., Nyquist criterion cannot exactly be met). Hence the matched filter demodulator given in Fig. 10.16 will not be exactly optimal but can be made to be very close to optimal with a well designed pulse shape.

a) Assume $D_z(k)$=1, compute the worse case degradation for an arbitrary pulse shape to the bit error probability for the demodulator given in Fig. 10.16. Specifically consider the case of the kth bit where $k = K_b/2$ where K_b is large and compute the effects of ISI. Identify the values of $D_z(j),\ j \neq k$ which achieve this worse case degradation for an arbitrary pulse shape with BPSK modulation.

b) Design a pulse shape and corresponding symbol rate such that the worse case (over all ISI patterns) degradation to the effective SNR is less than 0.25dB and the spectral mask given in Fig. 10.15 is met by the transmission. Grades will be assigned proportional to the achieved symbol rate (higher the better!). Anyone who beats the posted solution will get a bonus of 10 points for the project!

c) Plot the transmitted complex envelope for both pulse shapes. Plot the real component of the matched filter output for a modulation sequence of $[1 - 111 - 11 - 1 - 1 - 11]$ for your design. Identify the symbol time sampling points.

Your grade will be penalized if a common solution to yours is found among your classmates solutions in direct proportion to the number of identical solutions. This is to encourage each of you to work independently. **Any computer code (e.g., Matlab) should be turned in with the project write-up.**

Part III

Advanced Digital Communication

Chapter 11

Demodulation in Frequency Selective Channels

11.1 Current Status

Up to this point in this text we have introduced two general methods to communicate K_b bits

1. General M-ary modulations
   - The advantage of a general M-ary modulation is that it can achieve very good performance and arbitrary spectral efficiency.
   - The disadvantage of a general M-ary modulation is that without more structure the optimal demodulator has complexity $O(2^{K_b})$.
2. Orthogonal modulations (including stream modulations, OFDM, OCDM)
   - The advantage of orthogonal modulation is that the optimum demodulator has complexity $O(K_b)$ and a desired spectral efficiency can be achieved with a proper design of the modulation signals.
   - The disadvantage is that the performance is limited to that achievable with a single symbol transmission and the Nyquist criterion for zero intersymbol interference places restrictions on the transmitted signals that often cannot be realized in practice.

The goals for the remainder of the text is to provide communication techniques that bridge some of the areas in the tradeoffs in performance, complexity, and spectral efficiency between general modulations and orthogonal modulations. Specifically the remainder of the text will focus on

- Loosening the orthogonality requirement with a goal of maintaining a demodulation complexity that is $O(K_b)$. Note orthogonality between transmitted bits at the demodulator cannot be maintained in many important communication applications.
- Improving the performance or the spectral characteristics compared to orthogonal modulations with a goal of maintaining a demodulation complexity that is $O(K_b)$

This chapter will focus on the first of these two concepts as often it will be necessary to communicate over a frequency selective channel. A frequency selective channel is a linear time invariant filter whose transfer function varies as a function of frequency. The frequency selective nature of the channel will cause the output to be a "distorted" version of the input. These channels could cause a modulation to lose orthgonality in the received waveforms. We will investigate a wide variety of techniques to

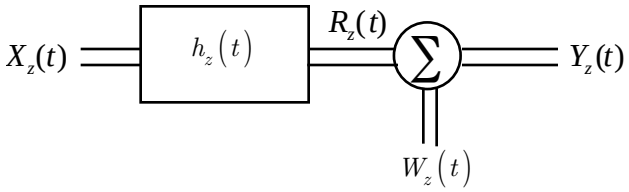


Figure 11.1: The block diagram for a frequency selective channels.

communicate over frequency selective channels. The effects of the channel on the receive signal will be detailed. The optimum demodulation structures will again be studied and performance quantified. The goal will be to identify demodulation algorithms that have complexity $O(K_b)$. Though in many cases the desired linear complexity can be achieved, the complexity of the optimum demodulator in a frequency selective channel is greater than in a frequency flat channel. Since the increase in complexity is often significant it will prove useful to examine lower complexity options for demodulation. These structures and the tradeoffs involved will be investigated for the variety of modulation schemes we have detailed up to this point.

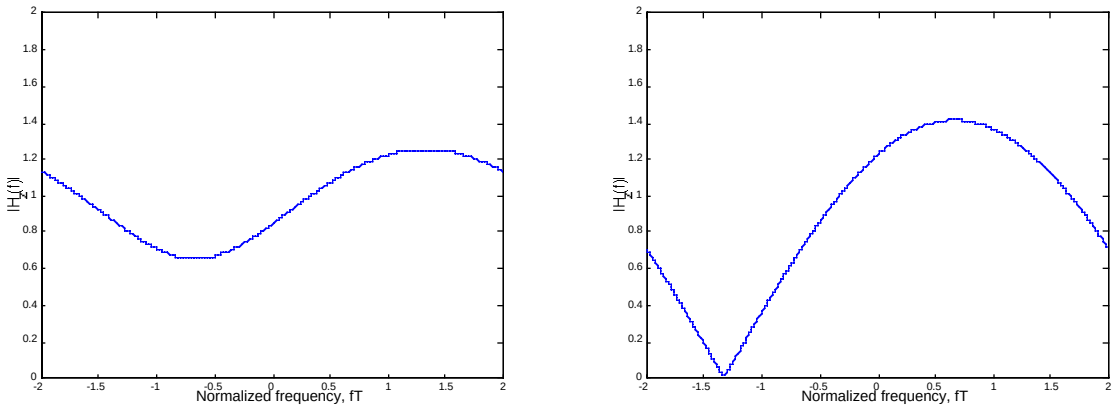
The case that will be considered in this chapter is where the frequency selective channel is known at the receiver but unknown at the transmitter and/or the communication is one-to-many (broadcast)[1]. In this case the transmitted signal cannot be designed to accomodate the frequency selectivity but the receiver will process the signal in a way that accounts for the frequency selectivity. The frequency selective transmission model is given in Fig. 11.1 and parameterized by a channel impulse response, $h_z(t)$ and an AWGN, $W_z(t)$. Clearly if the channel impulse response, $h_z(t)$ is known at the transmitter then the transmitted signal can be designed in such a way that the resulting $R_z(t)$ is any of the modulation techniques that were previously discussed. Unfortunately often in practice the channel might vary rapidly enough that feedback to the transmitter might not be effective or broadcast communications is desired (one to many) so that each user sees a different channel. In these cases little control can be maintained over the form of $R_z(t)$ and there is a need to identify the optimum demodulator in the presence of the channel distortion.

To formalize the discussion the goal will be to communicate K_b bits on a channel with impulse response $h_z(t)$ that is known at the receiver but not known at the transmitter. For a channel to be frequency selective, the channel must be time dispersive. For the clarity of discussion the following definition is given

Definition 11.1 *The delay spread of the frequency selective channel is the length of the support of the channel impulse response.*

In this chapter T_h will denote the delay spread and the support of the transmitted signal will again be denoted T_p. With little loss in generality for the remainder of the chapter it will be assumed that the support of $h_z(t)$ is $[0, T_h]$. The goal in this chapter is to understand how optimum receiver processing must be modified to accomodate the frequency selectivity or equivalently the dispersive nature of the channel.

[1]Some of the communication theory aspects when the channel is known at the transmitter will be explored in the homework.



a) $\alpha_1 = 0.3\exp[j2\pi/3]$ and $\tau_1 = 0.25T_s$. b) $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = 0.25T_s$.

Figure 11.2: The frequency response for two channels of the form given in (11.2).

Example 11.1: Consider a channel with

$$h_z(t) = \sqrt{1-|\alpha_1|^2}\delta(t) + \alpha_1\delta(t-\tau_1) \tag{11.1}$$

where α_1 is a complex constant and τ_1 is a positive constant. The frequency response of such a channel is given as

$$H_z(f) = \sqrt{1-|\alpha_1|^2} + \alpha_1\exp[-j2\pi f\tau_1]. \tag{11.2}$$

The frequency response of two channels of this form are shown in Fig. 11.2. Clearly the larger the magnitude of α_1 and the value of τ_1 the more frequency selective the channel becomes. It is clear that $T_h - \tau_1$.

11.2 General M-ary Modulations

The first thing to note is that the general optimum M-ary demodulation scheme changes very little in form with frequency selective channels compared to frequency flat channels. The only change that is made to the MLWD is that the matched filters are matched to each of the potential received signals, i.e.,

$$\begin{aligned}\hat{\vec{I}} &= \arg\max_{i\in\{0,\ldots,M-1\}} T_i \\ &= \arg\max_{i\in\{0,\ldots,M-1\}} \Re\left[\int_0^{T_p+T_h} Y_z(\tau)r_i^*(\tau)d\tau\right] - \frac{\tilde{E}_i}{2}\end{aligned} \tag{11.3}$$

where $r_i(t)$ is the postulated received signal when $\vec{I} = i$, i.e.,

$$r_i(t) = x_i(t) * h_z(t) = \int_0^{T_h} h_z(\tau)x_i(t-\tau)d\tau \tag{11.4}$$

and the energy correction term is a function of the energy of each of the potential received signals, i.e.,

$$\tilde{E}_i = \int_0^{T_p+T_h} |r_i(t)|^2 \, dt. \tag{11.5}$$

The average received energy per bit is now given as

$$E_b = \frac{1}{K_b M} \sum_{i=0}^{M-1} \tilde{E}_i. \tag{11.6}$$

It is important to note that the dispersive nature of the channel has made the actual received signal longer in time (T_p+T_h) and this increased length of the received signal must be included in the processing in the optimum demodulator. Hence without further structure in the signal little changes in optimum demodulation.

The performance of general M-ary signalling in frequency selective channels will again be a function of the Euclidean distance spectrum. The Euclidean distance between signals is now defined relative to the received signal, i.e.,

$$\Delta_E(i,j) = \int_0^{T_p+T_h} |r_i(t) - r_j(t)|^2 \, dt \tag{11.7}$$

So while the frequency selective channel does not change significantly the demodulator for general M-ary signaling, the channel can significantly change performance.

This discussion of the general M-ary modulation in a frequency selective channel can be concluded by noting that not much changes in this case. The major difference is that the matched filter is matched to the potential outputs of the channel as opposed to the transmitted signal. Similarly the Euclidean squared distance between possible signals is measured at the channel output. Consequently the real problem of interest is how a frequency selective channel changes demodulation and the performance for orthogonal modulations.

11.3 Frequency Selectivity and OCDM

In general OCDM in frequency selective channels has a more complex demodulator than presented in Chapter 9. In fact, the optimum demodulation complexity grows from $O(K_b)$ to $O(2^{K_b})$. Since this is an introductory course we cannot hope to cover all the details but the demodulation and performance of OCDM in frequency selective channels is considered extensively in [Ver98]. A tutorial article that is readable and comprehensive is [Mos96]. In this section we will concentrate on the transmitted signal being a binary OCDM signal where

$$X_z(t) = \sum_{l=1}^{K_b} D_z(l) s_l(t) \tag{11.8}$$

with modulation symbols $D_z(l) = a\left(I(l)\right)$ with $a(\bullet)$ being the constellation mapping and $s_l(t)$ being the unit energy spreading waveform which has a support of length T_p. Again assume the channel response, $h_z(t)$, is known at the receiver and dispersive, the data bits, $I(l)$, $l \in \{1, \ldots, K_b\}$, are equally likely and independent, and for simplicity of discussion that BPSK modulation is used. The output signal has the form

$$\begin{aligned} Y_z(t) &= R_z(t) + W_z(t) = \sum_{l=1}^{K_b} D_z(l) \int_0^{T_h} h_z(\lambda) s_l(t-\lambda) d\lambda + W(t) \\ &= \sum_{l=1}^{K_b} D_z(l) \tilde{s}_l(t) + W_z(t). \end{aligned} \tag{11.9}$$

where the effective spreading waveform for each bit is given as

$$\tilde{s}_l(t) = \int_0^{T_h} h_z(\lambda) s_l(t-\lambda) d\lambda. \tag{11.10}$$

Examining (11.9) shows that the effect of a frequency selective channel on an OCDM received signal is only to change the effective spreading waveform for each bit. The independence of the data bits imply that in the frequency selective channel the average energy of the effective spreading waveforms is equal to the received average energy per bit, E_b, i.e.,

$$E_b = \frac{1}{K_b} \sum_{l=1}^{K_b} E_{\tilde{s}_l} \tag{11.11}$$

where

$$E_{\tilde{s}_l} = \int_0^{T_p+T_h} |\tilde{s}_l(t)|^2 \, dt. \tag{11.12}$$

Example 11.2: Consider the OCDM waveforms that were considered in Fig. 9.12 where each spreading waveform has unit energy and the channel given in Example 11.1with $\tau_1 = T_p/4$. Since

$$h_z(t) = \alpha_0 \delta(t) + \alpha_1 \delta(t - \tau_1) \tag{11.13}$$

where $\alpha_0 = \sqrt{1 - |\alpha_1|^2}$, the equivalent spreading waveforms are

$$\tilde{s}_l(t) = \alpha_0 s_l(t) + \alpha_1 s_l(t - \tau_1). \tag{11.14}$$

Two of the equivalent spreading waveforms are plotted in Fig. 11.3 (note: $l = 1$ refers to the first spreading waveform and $l = 3$ refers to the third spreading waveform). The time dispersive nature of the channel has made the equivalent spreading waveform longer in time and has significantly changed the time waveform. It should be clear that while the example waveforms are clearly orthogonal in a frequency flat channel as considered in Chapter 9, the frequency selectivity nature of the channel has eliminated the orthogonality between the waveforms generated by each spreading code. This lack of orthogonality must be accounted for in the optimum demodulator.

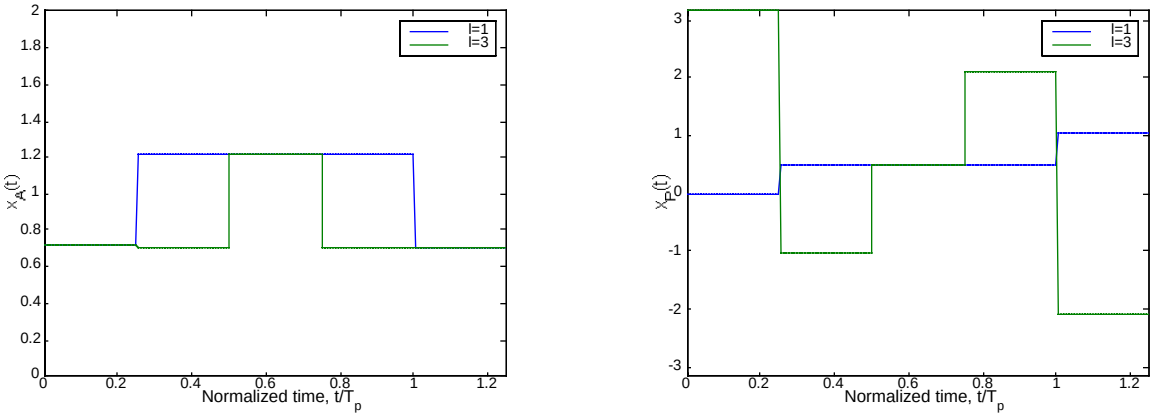
11.3.1 MLWD for OCDM in Frequency Selective Channels

The optimum demodulator has the same matched filter and energy correction term that has been prevalent in our discussion so far. Specifically for MLWD

$$\begin{aligned} \hat{\vec{I}} &= \arg \max_{i \in \{0,\ldots,M-1\}} T_i \\ &= \arg \max_{i \in \{0,\ldots,M-1\}} \Re \left[\int_0^{T_p+T_h} Y_z(t) r_i^*(t) dt \right] - \frac{\tilde{E}_i}{2} \end{aligned} \tag{11.15}$$

where $r_i(t) = \sum d_i(l) \tilde{s}_l(t)$ and $\tilde{E}_i = \int_0^{T_p+T_h} |r_i(t)|^2 \, dt$. The matched filter to the k^{th} effective spreading waveform will be denoted $\tilde{Q}(k)$ and has the form

$$\tilde{Q}(k) = \int_0^{T_p+T_h} Y_z(t) \tilde{s}_k^*(t) dt. \tag{11.16}$$



a) Amplitude b) Phase

Figure 11.3: Two equivalent spreading waveforms, $\tilde{s}_1(t)$ and $\tilde{s}_3(t)$, from the OCDM example of Chapter 9. $E_T = T_p$, $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = 0.25T$.

Recall the energy correction term will have the form

$$\begin{aligned}\tilde{E}_i &= \int_0^{T_p+T_h} \left| \sum_{l_1=1}^{K_b} d_i(l)\tilde{s}_l(t) \right|^2 dt \\ &= \sum_{l_1=1}^{K_b}\sum_{l_2=1}^{K_b} d_i(l_1)d_i^*(l_2) \int_0^{T_p+T_h} \tilde{s}_{l_1}(t)\tilde{s}_{l_2}^*(t)dt \\ &= \sum_{l_1=1}^{K_b}\sum_{l_2=1}^{K_b} d_i(l_1)d_i^*(l_2)V_{\tilde{s}}(l_1,l_2). \end{aligned} \tag{11.17}$$

Here we denote the cross-correlation function between the possible effective waveforms with $V_{\tilde{s}}(l_1, l_2)$, i.e.,

$$V_{\tilde{s}}(l_1,l_2) = \int_0^{T_p+T_h} \tilde{s}_{l_1}(t)\tilde{s}_{l_2}^*(t)dt. \tag{11.18}$$

Hence the energy correction term is entirely a function of the postulated transmitted symbols, $d_i(l)$ $l \in \{1, \ldots, K_b\}$ and the cross-correlation functions between the effective spreading waveforms. Using (11.16) and (11.17) in (11.15) gives the MLWD as

$$\hat{\vec{I}} = \arg \max_{i\in\{0,\ldots,M-1\}} \sum_{k=1}^{K_b} \Re\left[\tilde{Q}(k)d_i^*(k)\right] - \frac{1}{2}\sum_{l_1=1}^{K_b}\sum_{l_2=1}^{K_b} d_i(l_1)d_i^*(l_2)V_{\tilde{s}}(l_1,l_2) \tag{11.19}$$

The effect of the frequency selective channel is to eliminate the orthogonality condition and mandate that the optimum demodulator must have an exponential complexity in the number of bits transmitted. All possible 2^{K_b} possible matched filters outputs and energy correction terms must be computed to produce the optimum decision. It should be noted that if a channel does not produce a large amount of non-orthogonality the simple demodulator for the flat channel is often implemented in practice. The low complexity implementation is well worth the small loss in performance that is produced. As the amount of frequency selectivity grows the viability of using the simple demodulator for the flat fading channel with a resulting small degradation decreases.

The form for the output of the matched filter provides insight into why the exponential complexity in the number of bits transmitted is necessary. If $\vec{I} = j$ then $Y_z(t) = \sum d_j(l)\tilde{s}_l(t) + W_z(t)$ and then

$$\tilde{Q}(k) = \sum_{l=1}^{K_b} d_j(l) V_{\tilde{s}}(l,k) + \tilde{N}(k), \tag{11.20}$$

where

$$\tilde{N}(k) = \int_0^{T_p+T_h} W_z(t)\tilde{s}_k^*(t)dt. \tag{11.21}$$

Because $h_z(t)$ is dispersive, the orthogonality condition, $V_{\tilde{s}}(l,k) = E_{\tilde{s}}(l)\delta_{l-k}$, cannot be guaranteed. Hence the terms in the summation, $\sum_{l\neq k} d_j(l)V_{\tilde{s}}(l,k)$, can be viewed as the interference from other bits, $d_j(l)$ $l \neq k$, not corresponding to the spreading waveform k. This interference is often denoted intersymbol interference (ISI) or multi-user interference[2]. Because of this interference the optimum demodulator for a frequency selective channel is generally more complicated than the demodulator for the non-selective channel.

Using vector and matrix notation allows a very simple form for the matched filter output. Defining the vectors

$$\vec{\tilde{Q}} = \begin{bmatrix} \tilde{Q}(1) \\ \vdots \\ \tilde{Q}(K_b) \end{bmatrix} \qquad \vec{D} = \begin{bmatrix} D_z(1) \\ \vdots \\ D_z(K_b) \end{bmatrix} \qquad \vec{\tilde{N}} = \begin{bmatrix} \tilde{N}(1) \\ \vdots \\ \tilde{N}(K_b) \end{bmatrix}, \tag{11.22}$$

the matched filter outputs are given as

$$\vec{\tilde{Q}} = E_b \mathbf{G}\vec{D} + \vec{\tilde{N}}. \tag{11.23}$$

The matrix $\mathbf{G}$ is of size $K_b \times K_b$ where the elements are given as $[\mathbf{G}]_{mn} = V_{\tilde{s}}(n,m)/E_b$. It should be noted that $\mathbf{G}$ is a Hermitian symmetric correlation matrix with trace $(G) = K_b$ (see 11.11) and that if $h_r(t)$ is not frequency selective then $G = I_{K_b}$ where I_N is the identity matrix of size N. The noise vector has a covariance matrix of (see Problem 11.2)

$$\mathbf{R}_{\tilde{N}} = E\left[\vec{\tilde{N}}\vec{\tilde{N}}^H\right] = E_b N_0 \mathbf{G}. \tag{11.24}$$

An equivalent linear algebraic representation of the problem formulation is given in Fig. 11.4. In examing the model in Fig. 11.4 it is obvious that OCDM detection must accomodate the variable gain given to the symbol for each bit (non-equal values on the diagonal of $\mathbf{G}$), the intersymbol interference between symbols (nonzero off diagonal elements), and the colored noise corrupting the detection.

With this notation the optimum demodulator takes a simple form. The MLWD is

$$\hat{\vec{I}} = \arg \max_{i\in\{0,\ldots,M-1\}} \Re\left[\vec{d}_i^H \vec{\tilde{Q}}\right] - E_b \frac{\vec{d}_i^H \mathbf{G}\vec{d}_i}{2}. \tag{11.25}$$

Here is can be seen that the energy correction term in OCDM is a quadratic form of the matrix $\mathbf{G}$. While the form for the optimum demodulator in frequency selective fading is compact there is still a need to search for the maximum ML metric over $M = 2^{K_b}$ possible hypotheses.

[2]In some multiple access systems each bit is assigned to a separate user in OCDM [Ver98].

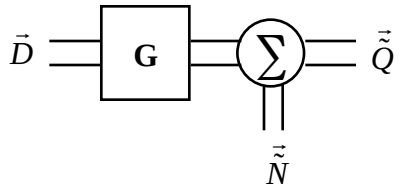


Figure 11.4: The linear algebraic model of OCDM demodulation in a frequency selective channel.

Example 11.3: Consider the channel model in Example 11.1 and the signals set considered in Example 11.2. Since

$$\tilde{s}_l(t) = \alpha_0 s_l(t) + \alpha_1 s_l(t-\tau_1) \tag{11.26}$$

the cross correlation between the effective spreading waveforms is

$$V_{\tilde{s}}(l_1,l_2) = E_T\delta_{l_1-l_2} + \alpha_0\alpha_1^* V_{s_{l_1}s_{l_2}}(\tau_1) + \alpha_0^*\alpha_1 V_{s_{l_1}s_{l_2}}(-\tau_1) \tag{11.27}$$

where $V_{s_{l_1}s_{l_2}}(\tau_1)$ is cross correlation function of the original spreading waveforms evaluated at a time offset of τ_1. When $E_T = T_p$, $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = 0.25T_p$ the matrix in the energy correction term is given as

$$\mathbf{G} = \begin{bmatrix} 1+0.7498\cos\left(\frac{\pi}{3}\right) & -j0.2499\sin\left(\frac{\pi}{3}\right) & -j0.2499\sin\left(\frac{\pi}{3}\right) & -0.2499\cos\left(\frac{\pi}{3}\right) \\ j0.2499\sin\left(\frac{\pi}{3}\right) & 1+0.2499\cos\left(\frac{\pi}{3}\right) & 0.2499\cos\left(\frac{\pi}{3}\right) & -j0.7498\sin\left(\frac{\pi}{3}\right) \\ j0.2499\sin\left(\frac{\pi}{3}\right) & 0.2499\cos\left(\frac{\pi}{3}\right) & 1-0.7498\cos\left(\frac{\pi}{3}\right) & j0.2499\sin\left(\frac{\pi}{3}\right) \\ -0.2499\cos\left(\frac{\pi}{3}\right) & j0.7498\sin\left(\frac{\pi}{3}\right) & -j0.2499\sin\left(\frac{\pi}{3}\right) & 1-0.2499\cos\left(\frac{\pi}{3}\right) \end{bmatrix}.$$

For example the vector diagram of the noiseless matched filter outputs are plotted in Fig. 11.5. The non-orthogonality of the effective spreading waveforms is evident in the intersymbol interference that is apparent in each matched filter output.

The performance analysis of OCDM in a frequency selective channel follows the same procedure as a general M-ary modulation. The union bound can be formed by enumerating all the pair-wise Euclidean distances. For OCDM with a frequency selective channel this can be enumerated with

$$\begin{aligned} \Delta_E(i,j) &= \int_0^{T_p+T_h} |r_i(t)-r_j(t)|^2\,dt = \int_0^{T_p+T_h} \left|\sum_{l_1=1}^{K_b} d_i(l)\tilde{s}_l(t) - \sum_{l_1=1}^{K_b} d_j(l)\tilde{s}_l(t)\right|^2 dt \\ &= \int_0^{T_p+T_h} \left|\sum_{l_1=1}^{K_b} \left[d_i(l)-d_j(l)\right]\tilde{s}_l(t)\right|^2 dt. \end{aligned} \tag{11.28}$$

Noting the similarity of (11.28) to the energy correction term in (11.17) this Euclidean distance can be simplified down to

$$\begin{aligned} \Delta_E(i,j) &= \sum_{l_1=1}^{K_b}\sum_{l_2=1}^{K_b} \left(d_i(l_1)-d_j(l_1)\right)\left(d_i(l_2)-d_j(l_2)\right)^* V_{\tilde{s}}(l_1,l_2) \\ &= E_b\left(\vec{d}_i-\vec{d}_j\right)^H \mathbf{G}\left(\vec{d}_i-\vec{d}_j\right) \end{aligned} \tag{11.29}$$

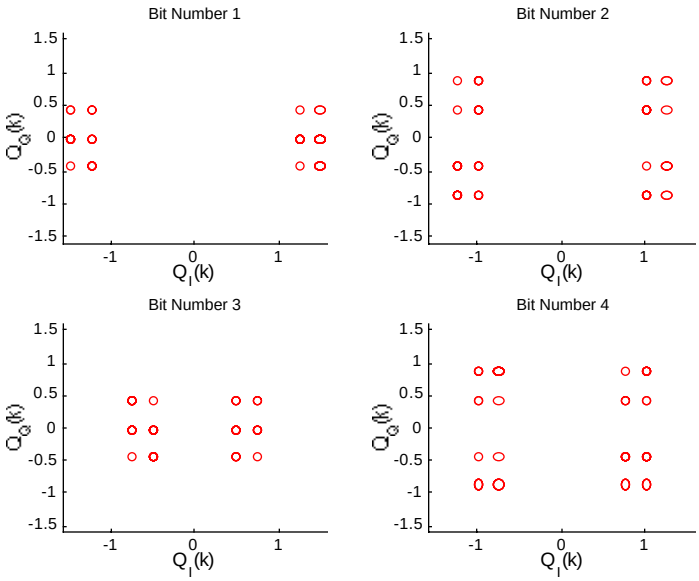


Figure 11.5: The vector diagram for the matched filter outputs for the channel considered in Example 11.3.

Consequently it is apparent from (11.29) that the only two things that determine the performance of the MLWD of OCDM in frequency selective fading are the pair–wise symbol differences and the auto and cross correlations of the distorted spreading waveforms.

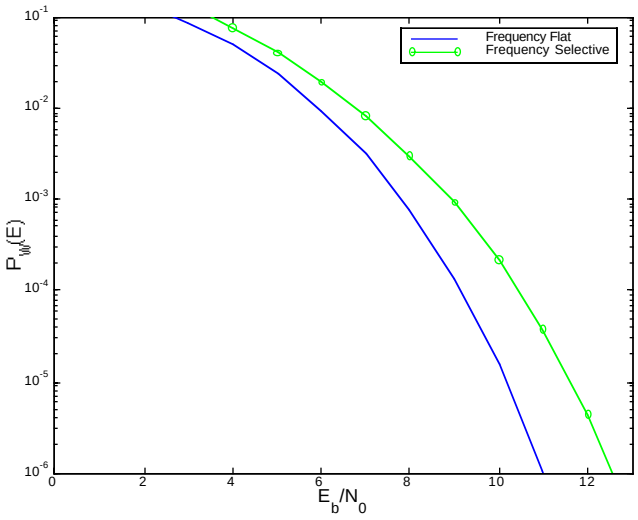


Figure 11.6: The union bound to the probability of word error for a frequency selective channel of Example 11.4:

Example 11.4: All the Euclidean distances for the OCDM modulation given in Example 11.3 are enumerated in a table

	$\Delta_E(i,j)$															
	$\vec{I}=i$															
$\vec{I}=j$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	5.5	4.5	10.0	2.5	8.0	8.0	13.5	3.5	8.0	8.0	12.5	6.0	10.5	11.5	16.0
1	5.5	0	10.0	4.5	8.0	2.5	13.5	8.0	10.0	3.5	14.5	8.0	12.5	6.0	18.0	11.5
2	4.5	10.0	0	5.5	6.0	11.5	2.5	8.0	8.0	12.5	3.5	8.0	9.5	14.0	6.0	10.5
3	10.0	4.5	5.5	0	11.5	6.0	8.0	2.5	14.5	8.0	10.0	3.5	16.0	9.5	12.5	6.0
4	2.5	8.0	6.0	11.5	0	5.5	4.5	10.0	6.0	10.5	9.5	14.0	3.5	8.0	8.0	12.5
5	8.0	2.5	11.5	6.0	5.5	0	10.0	4.5	12.5	6.0	16.0	9.5	10.0	3.5	14.5	8.0
6	8.0	13.5	2.5	8.0	4.5	10.0	0	5.5	11.5	16.0	6.0	10.5	8.0	12.5	3.5	8.0
7	13.5	8.0	8.0	2.5	10.0	4.5	5.5	0	18.0	11.5	12.5	6.0	14.5	8.0	10.0	3.5
8	3.5	10.	8.	14.5	6.	12.5	11.5	18.	0	5.5	4.5	10.	2.5	8.0	8.	13.5
9	8.0	3.5	12.5	8.	10.5	6.	16.0	11.5	5.5	0	10.	4.5	8.0	2.5	13.5	8.0
10	8.0	14.5	3.5	10.	9.5	16.	6.0	12.5	4.5	10.	0	5.5	6.0	11.5	2.5	8.0
11	12.5	8.	8.	3.5	14.	9.5	10.5	6.	10.	4.5	5.5	0	11.5	6.0	8.0	2.5
12	6.	12.5	9.5	16.	3.5	10.	8.0	14.5	2.5	8.	6.	11.5	0	5.5	4.5	10.0
13	10.5	6.	14.	9.5	8.	3.5	12.5	8.	8.	2.5	11.5	6.	5.5	0	10.0	4.5
14	11.5	18.	6.	12.5	8.	14.5	3.5	10.	8.	13.5	2.5	8.	4.5	10.0	0	5.5
15	16.	11.5	10.5	6.	12.5	8.0	8.0	3.5	13.5	8.	8.	2.5	10.	4.5	5.5	0

For this particular example the minimum Euclidean distance results when there is a difference only in the symbol that is modulated on lowest energy effective spreading waveform ($\tilde{s}_3(t)$). The union bound to the word error probability in the frequency selective channel and the word error probability for the frequency flat channel are plotted in Fig. 11.6. This particular frequency selective channel causes about 1.5dB degradation in the performance of OCDM and the optimum demodulator complexity grows significantly to accomodate this frequency selectivity.

11.3.2 Suboptimal OCDM Demodulation

Suboptimal demodulators are often used in engineering practice. Two reasons typically drive engineers to consider suboptimum demodulators. First, when K_b is large the complexity of (11.25) can become prohibitive for implementation due to the exponential complexity in K_b. The second situation is that the amount of nonorthogonality is relatively mild. Very good performance with mild nonorthogonality can be achieved without implementing a full MLWD or MLBD. Hence, communication engineers searched for alternate demodulation structures that were not exponentially complex and yet provided good performance. The field of multi-user detection is a rich and mature one. This text only hopes to introduce the concepts while the interested reader is referred to more comprehensive treatments (e.g., [Ver98])

Decorrelating Detector

The first idea that was proposed was a linear transformation, $\mathbf{W}^H$ on $\vec{\tilde{Q}}$ known as the decorrelating detector [Shn67, Ver98]. This detector ignores noise and tries to restore the orthogonality condition by setting $\mathbf{W}^H = \mathbf{G}^{-1}$ which results in

$$\hat{\vec{D}} = \mathbf{W}^H \vec{\tilde{Q}} = \mathbf{G}^{-1}\vec{\tilde{Q}} = E_b\vec{D} + \mathbf{G}^{-1}\vec{\tilde{N}} = E_b\vec{D} + \vec{N}_d. \tag{11.30}$$

Each bit decision is then made with a simple threshold test on each component of $\hat{\vec{D}}$ as is done in orthogonal modulation, i.e.,

$$\Re\left[\hat{D}(k)\right] \underset{\hat{I}(k)=1}{\overset{\hat{I}(k)=0}{\gtrless}} 0. \tag{11.31}$$

It should be noted that $\mathbf{G}^{-1}$ might not necessarily be well defined so care must be taken to evaluate whether $\mathbf{G}$ is singular. This detector has a complexity that is $O(K_b^2)$ so there is a significant complexity savings compared to the complexity of the optimal demodulator $O(2^{K_b})$. Since the decorrelator output has a noise covariance of (see Problem 11.3)

$$\mathbf{R}_{\tilde{N}_d} = E\left[\vec{N}_d\vec{N}_d^H\right] = E_bN_0\mathbf{G}^{-1}, \tag{11.32}$$

the probability of error for each bit and the average probability of error is given as

$$P_B(E,k) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0 w_{k,k}}}\right) \qquad P_B(E) = \frac{1}{K_b}\sum_{k=1}^{K_b} P_B(E,k) \tag{11.33}$$

where $w_{k,k} = \left[\mathbf{W}^H\right]_{k,k} = \left[\mathbf{G}^{-1}\right]_{k,k}$. It is possible to show that the decorrelator never improves the average probability of bit error though decisions for some bit indexes might have a higher fidelity than on the frequency flat channel (e.g., $w_{k,k} \leq 1$, see Problem 11.3). The word error probability of the decorrelator is more difficult to calculate due to the correlated noise corrupting each of the K_b threshold tests.

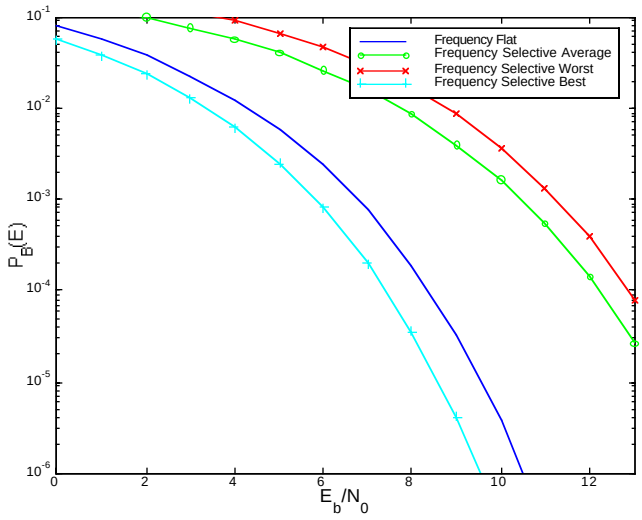


Figure 11.7: The performance of the decorrelating detector for the channel given in Example 11.3.

Example 11.5: For the system considered in Example 11.3 the decorrelating detector is (to 3 significant digits)

$$\begin{aligned}\mathbf{W}^H &= \mathbf{G}^{-1} \\ &= \begin{bmatrix} 0.8 & 0 & j0.346 & 0.2 \\ 0 & 2 & -1 & j1.730 \\ -j0.346 & -1 & 2.4 & -j1.384 \\ 0.2 & -j1.730 & j1.384 & 2.8 \end{bmatrix}.\end{aligned} \tag{11.34}$$

The performance for the decorrelating detector is shown in Fig. 11.7. The first bit has performance better than that obtained in a frequency flat channel but the fourth bit has a performance much worse (approximately 4.5dB worse). The average performance is dominated by the worst case.

The advantages of the decorrelating detector are a relatively low complexity, no need to know the SNR, a return to orthogonality between data bits, and a tractable performance analysis. The disadvantages of the decorrelating detector are noise enhancement in certain situations which can produce performance far from a MLWD and the complexity of the detector is $O(K_b^2)$ and not linear in the number of bits transmitted.

MMSE Linear Detector

Another linear detector has the same structure as the decorrelating detector but can give better performance. The general linear detector has the form

$$\hat{\vec{D}} = \mathbf{W}^H \vec{Q} = \mathbf{W}^H E_b \mathbf{G} \vec{D} + \mathbf{W}^H \vec{N} \tag{11.35}$$

In the decorrelating detector the ISI is completely eliminated but the effects on the noise of this ISI removal is ignored. The problem can be reformulated to minimize the combined effects residual ISI and the filtered noise. Defining the estimation error as

$$\vec{E} = E_b \vec{D} - \hat{\vec{D}} = E_b \left(\mathbf{I}_{K_b} - \mathbf{W}^H \mathbf{G} \right) \vec{D} + \mathbf{W}^H. \tag{11.36}$$

The first term of the estimation error is due to the residual ISI after estimation and the second term is due to the additive noise. The goal in choosing a linear estimator is to find the matrix, $\mathbf{W}$ such that $\hat{\vec{D}} = \mathbf{W}^H \vec{\tilde{Q}}$ and that $R_E = E\left[\vec{E}\vec{E}^H\right]$ is minimized. This is classical linear minimum mean square error (MMSE) estimation [Poo88, Sch91]. A brief review of MMSE estimation is given in Appendix B. In order to use this tool the data has to be modeled as a random vector. The statistical model most consistent with the development in this book is that each data symbol is independent and with each constellation element equally likely to be transmitted. In addition it is assumed that the data symbols are all independent of the noise time series. This model reduces down to $\mathbf{R}_D = E\left[\vec{D}\vec{D}^H\right] = \mathbf{I}_{K_b}$ and $E\left[\vec{D}\vec{\tilde{N}}^H\right] = \mathbf{0}_{K_b}$. The MMSE filter is achieved when the observations are orthogonal to the error, i.e., $E\left[\vec{\tilde{Q}}\vec{E}^H\right] = 0$. Solving for $\mathbf{W}$ based on the orthogonality condition gives

$$\mathbf{W}^H E\left[\vec{\tilde{Q}}\vec{\tilde{Q}}^H\right] = E\left[E_b\vec{D}\vec{\tilde{Q}}^H\right] \tag{11.37}$$

Using the assumed data models we have

$$E\left[\vec{\tilde{Q}}\vec{\tilde{Q}}^H\right] = E_b^2\mathbf{G}\mathbf{G}^H + E_bN_0\mathbf{G} \tag{11.38}$$

and

$$E\left[E_b\vec{D}\vec{\tilde{Q}}^H\right] = E_b^2\mathbf{G}^H. \tag{11.39}$$

Consequently the orthogonality condition reduces down to

$$\mathbf{W}^H\left(E_b^2\mathbf{G}\mathbf{G}^H + E_bN_0\mathbf{G}\right) = E_b^2\mathbf{G}^H. \tag{11.40}$$

Noting that $\mathbf{G}^H = \mathbf{G}$ gives the MMSE filter as

$$\mathbf{W}^H = \left(\mathbf{G} + \frac{N_0}{E_b}I_{K_b}\right)^{-1}. \tag{11.41}$$

Each bit decision is then made with a simple threshold test on each component of $\hat{\vec{D}}$ as in the decorrelator, i.e.,

$$\Re\left[\hat{D}(k)\right] \underset{\hat{I}(k)=1}{\overset{\hat{I}(k)=0}{\gtrless}} 0. \tag{11.42}$$

This filter and detector is often referred to as the MMSE detector [Ver98, XSR90, MH94]. This detector has the same complexity as the decorrelator. The probability of error is not as easily analyzed as with the decorrelating detector since the error in $\hat{D}(k)$ is a combination of filtered Gaussian noise (easy to analyze) and the residual ISI (dependent on the transmitted data). The procedure to analyze a simple case is considered in Problem 11.4. It should also be noted that if $N_0 \to 0$ then the MMSE detector converges to the decorrelating detector.

The advantage of the MMSE detector is a significant performance gain versus the decorrelating detector at the same complexity. The disadvantages of the MMSE detector are that the SNR must be known, noise enhancement still occurs, and the complexity of the detector is $O(K_b^2)$ and not linear in the number of bits transmitted.

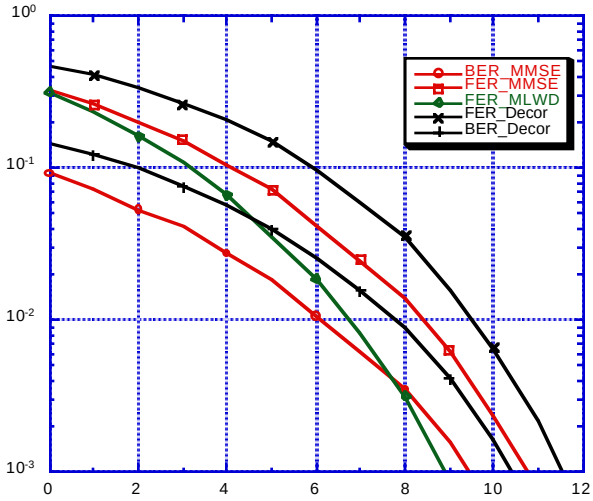


Figure 11.8: The performance of the MMSE detector for Example 11.3.

Example 11.6: Consider the system given in Example 11.3 where $E_b/N_0 = 10$dB. The MMSE linear detector is given as

$$\begin{aligned} \mathbf{W}^H &= \left(\mathbf{G} + \frac{N_0}{E_b} I_{K_b}\right)^{-1} \qquad (11.43) \\ &= \begin{bmatrix} 0.7305 & j0.03852 & j0.2482 & 0.1231 \\ -j0.0385 & 1.4609 & -0.5669 & j1.0939 \\ -j0.2482 & -0.5669 & 1.7919 & -j0.8072 \\ 0.1231 & -j1.0939 & j0.8072 & 1.9492 \end{bmatrix}. \end{aligned}$$

The overall response for the signal will be due to the matrix

$$\mathbf{W}^H\mathbf{G} = \begin{bmatrix} 0.9269 & -j0.0039 & -j0.0248 & -0.0123 \\ j0.0039 & 0.8539 & 0.0567 & -j0.1094 \\ j0.0248 & 0.0567 & 0.8208 & j0.0807 \\ -0.012 & j0.1094 & -j0.0807 & 0.8051 \end{bmatrix}. \qquad (11.44)$$

Consequently the MMSE linear detector produces ISI ($\mathbf{W}^H\mathbf{G} \neq \mathbf{I}_{K_b}$), but it does it in such a way so that the sum of the ISI power and noise power is minimized. It should also be noted that the MMSE detector is not a conditionally unbiased estimator (i.e., $E\left(\hat{D}(i)\,|D(i) = d\right) \neq E_b d$. For the estimator to be conditionally unbiased the diagonal elements of $\mathbf{W}^H\mathbf{G}$ would have to be unity. The performance for the MMSE detector is shown in Fig. 11.8. At low SNR the MMSE detector has as much as a 2dB gain compared to the decorrelating detector. The gain slowly decreases as the SNR increases.

Successive Interference Canceller

The simple linear structures of the decorrelating and the MMSE detectors often does not give good enough performance in certain channels. This occurs when the interference is such that a number of the

bits cannot be reliably detected. The fortunate thing is that usually some of the bits can be reliably detected and interference can be removed from the received signal. A simple addition to either of the abovementioned detectors (decorrelator and MMSE) is a component wise detection process with a decision directed interference cancellation operation.

The successive interference cancelling strategy is such that the row of the $\mathbf{W}^H$ matrix that provides the best performance is identified. That row of $\mathbf{W}^H$ is applied to $\vec{\tilde{Q}}$ and a hard decision is made. For instance assume that the k_1^{th} row gives the best performance then the following bit detector is formed

$$\Re\left[\hat{D}(k_1)\right] = \Re\left[\vec{W}(k_1)\vec{\tilde{Q}}\right] \underset{\hat{I}(k_1)=1}{\overset{\hat{I}(k_1)=0}{\gtrless}} 0. \tag{11.45}$$

Since $\mathbf{G}$ is known and

$$\tilde{Q}(k) = E_b \sum_{l=1}^{K_b} D(l) g_{k,l} + \tilde{N}(k) = E_b \sum_{l\neq k_1}^{K_b} D(l) g_{k,l} + E_b D(k_1) g_{k,k_1} + \tilde{N}(k) \qquad k \neq k_1, \tag{11.46}$$

this decision can reduce the interference by forming

$$\tilde{Q}^{(1)}(k) = \tilde{Q}(k) - E_b g_{k,k_1} a\left(\hat{I}(k_1)\right) \qquad k \neq k_1. \tag{11.47}$$

Assuming that $I(k_1) = \hat{I}(k_1)$ (perfect decisions), one can note that after forming $\tilde{Q}^{(1)}(k) \quad k \neq k_1$ the signal model for the vector $\vec{\tilde{Q}}^{(1)}$ (size $K_b - 1 \times 1$) is exactly the same as for $\tilde{Q}$ except with one less dimension. Specifically $\mathbf{G}^{(1)}$ is $\mathbf{G}$ with the k_1^{th} row and column removed. The MMSE (or decorrelator) detector is recomputed for the smaller dimensoional problem and the detection process is repeated. This successive application of a linear detector and interference subtraction is continued until a decision is made on all the bits. This demodulator is known as a successive interference canceller (SIC) and Fig. 11.9 shows the block diagram of the SIC. Several variants of this general structure exist but with the same general idea [Ver98]. Also decision errors can cause significant degradation to the performance of such an interference canceller so in practice this algorithm should be used only when the most reliable bit/user can be detected with high fidelity in a linear decoder.

The SIC demodulator has a complexity about the same as the linear detectors. The linear detector at each stage, $k \in \{1, \ldots, K_b\}$, has a complexity of $O(K_b - k + 1)$. The interference removal/subtraction has a complexity of $O(k-1)$. Together the combination of filtering and interference supression will again result in a complexity that is $O(K_b^2)$. This is another case of intelligent engineering resulting in improved performance without an increase in complexity.

The advantages of the SIC detector are a significant performance gain versus linear detectors. The disadvantages of the SIC detector are that the complexity of the detector is $O(K_b^2)$ and not $O(K_b)$ as desired. Also decision errors fed back in the cancellation process can cause significant performance degradations in the bit error probability.

Example 11.7: Consider the system given in Example 11.4 where the linear stage in the SIC is chosen to be a decorrelator. Since $w_{1,1} = 0.8$ the first bit can be detected with the highest fidelity,

$$\hat{D}(1) = \vec{W}(1)^H \vec{\tilde{Q}} \tag{11.48}$$

where $\vec{W}(1)^H$ is the first row of the decorrelating detector at little increase in complexity. It should be noted that the performance for this first detected bit is exactly the same as the decorrelator detector. The SIC assumes that the decision, $\hat{I}(1)$ is correct and subtracts out the contribution of this bit, i.e.,

$$\tilde{Q}^{(1)}(k) = \tilde{Q}(k+1) - E_b g_{k+1,1} a\left(\hat{I}(1)\right) \qquad k \in \{1,2,3\}\,. \tag{11.49}$$

The new signal model is

$$\vec{\tilde{Q}}^{(1)} = \begin{bmatrix} 1.125 & 0.125 & -j0.649 \\ 0.125 & 0.625 & j0.216 \\ j0.649 & -j0.216 & 0.875 \end{bmatrix} \begin{bmatrix} D(2) \\ D(3) \\ D(4) \end{bmatrix} + \vec{\tilde{N}}^{(1)} \tag{11.50}$$

A new decorrelating detector can be formed and the best decision ($\hat{D}(2)$) selected for feedback to produce further model reduction. Again assuming correct decisions the model after two stages is

$$\vec{\tilde{Q}}^{(2)} = \begin{bmatrix} 0.625 & j0.216 \\ -j0.216 & 0.875 \end{bmatrix} \begin{bmatrix} D(3) \\ D(4) \end{bmatrix} + \vec{\tilde{N}}^{(2)}. \tag{11.51}$$

A new decorrelating detector can be formed and the best decision ($\hat{D}(4)$) selected for feedback to produce further model reduction. Again assuming correct decisions the model for the final stage is

$$\vec{\tilde{Q}}^{(3)} = \begin{bmatrix} 0.625 \end{bmatrix} \begin{bmatrix} D(3) \end{bmatrix} + \vec{\tilde{N}}^{(3)}. \tag{11.52}$$

Under the assumption of perfect decision feedback the first detected bit will have the same performance as the linear decorrelator and the subsequent three bits will be detected with a higher fidelity than the linear decorrelator. The performance for the decorrelator SIC detector is shown in Fig. 11.10. This detector has almost 2dB gain versus the stand alone decorrelator. It should be noted that Fig. 11.10 has plots of the decisions being fedback and the true data used for the feedback. It should be noted that the perfect feedback case is most often denoted genie aided detection in the communication theory literature hence the labels "SIC_Genie". It is interesting to note that the frame/word error rate of the SIC is not impacted by errors fedback in the interference cancellation (for obvious reasons). Even the bit error rate is only minorly impacted as in this example the decisions are taken in performance ranked order, i.e., the best decorrelator decision is taken first and the worst decorrelator decision is taken last.

11.4 Frequency Selectivity and OFDM

In general OFDM in frequency selective channels will also have a more complex demodulator than presented in Chapter 9 but certain modifications can be made to permit simple suboptimal bit-by-bit demodulation. In this section we will concentrate on the the transmitted signal being a binary OFDM signal where

$$X_z(t) = \sum_{l=1}^{K_b} D_z(l) u(t) \exp\left[j2\pi f_d (2l - K_b - 1)t\right] = \sum_{l=1}^{K_b} D_z(l) u(t) \exp\left[j2\pi f_l t\right] \tag{11.53}$$

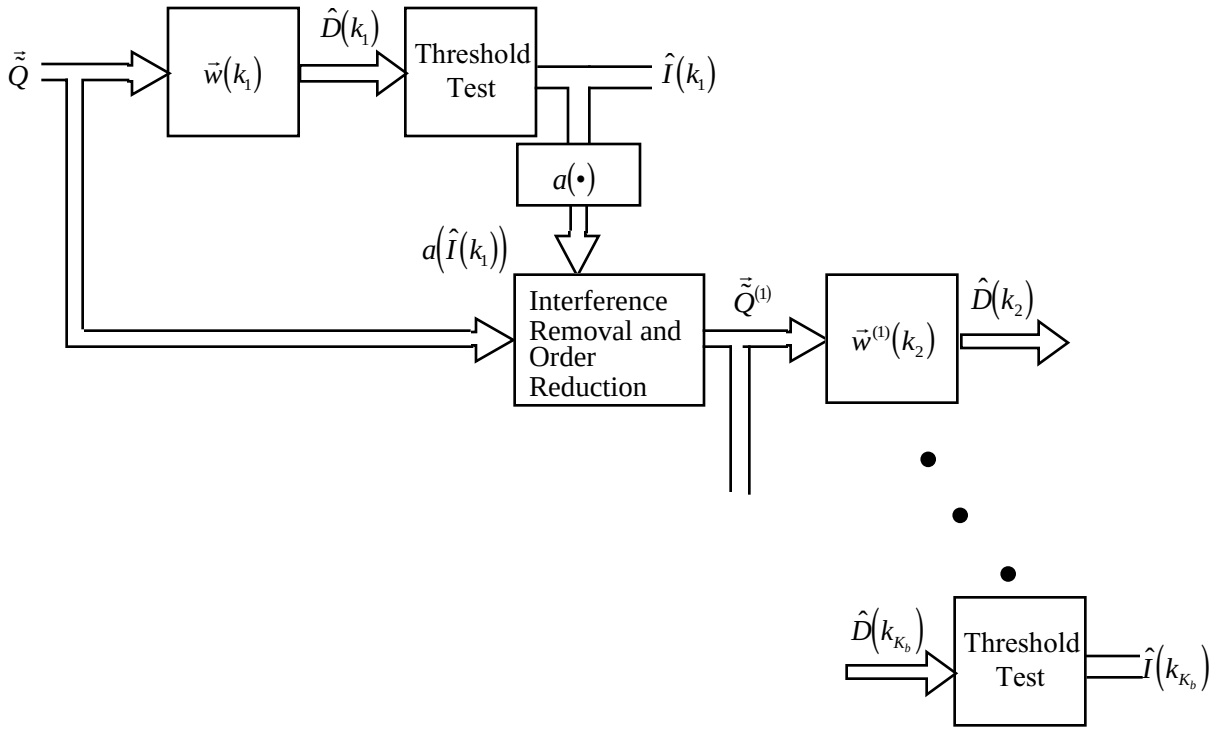


Figure 11.9: The block diagram of the successive interference canceller.

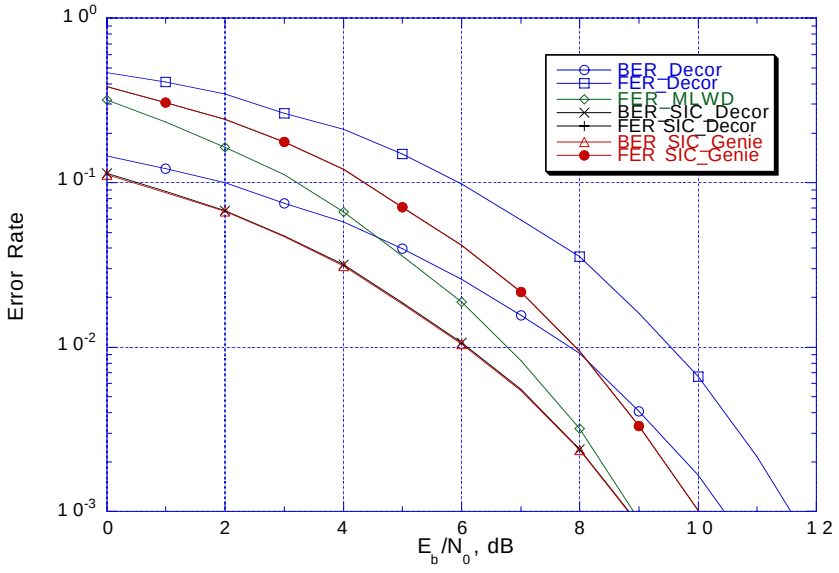


Figure 11.10: The performance of the SIC detector for Example 11.7.

with modulation symbols $D_z(l) = a\left(I(l)\right)$ with $a(\bullet)$ being the constellation mapping and $u(t)$ being the unit energy pulse shape which has a support of length T_p. Again assume the channel response, $h_z(t)$, is known at the receiver and dispersive. The output signal has the form

$$Y_z(t) = R_z(t) + W_z(t) = \sum_{l=1}^{K_b} D_z(l) \int_0^{T_h} h_z(\lambda) u(t-\lambda) \exp\left[j2\pi f_l (t-\lambda)\right] d\lambda + W_z(t). \tag{11.54}$$

By factoring the exponential term into a function of t and a function of λ, this channel output is given as

$$\begin{aligned} Y_z(t) &= \sum_{l=1}^{K_b} D_z(l) \exp\left[j2\pi f_l t\right] \int_0^{T_h} h_z(\lambda) u(t-\lambda) \exp\left[-j2\pi f_l \lambda\right] d\lambda + W_z(t). \\ &= \sum_{l=1}^{K_b} D_z(l) \exp\left[j2\pi f_l t\right] \tilde{u}(t,l) + W_z(t) = \sum_{l=1}^{K_b} D_z(l) \tilde{s}_l(t) + W_z(t). \end{aligned} \tag{11.55}$$

where $\tilde{s}_l(t)$ is the effective pulse waveform and $\tilde{u}(t,l)$ is the effective subcarrier pulse shape for each bit. The effective pulse shape for each subcarrier is given as

$$\tilde{u}(t,l) = \int_0^{T_h} h_z(\lambda) u(t-\lambda) \exp\left[-j2\pi f_l \lambda\right] d\lambda \tag{11.56}$$

and the effective spreading waveform is

$$\tilde{s}_l(t) = \tilde{u}(t,l) \exp\left[j2\pi f_l t\right]. \tag{11.57}$$

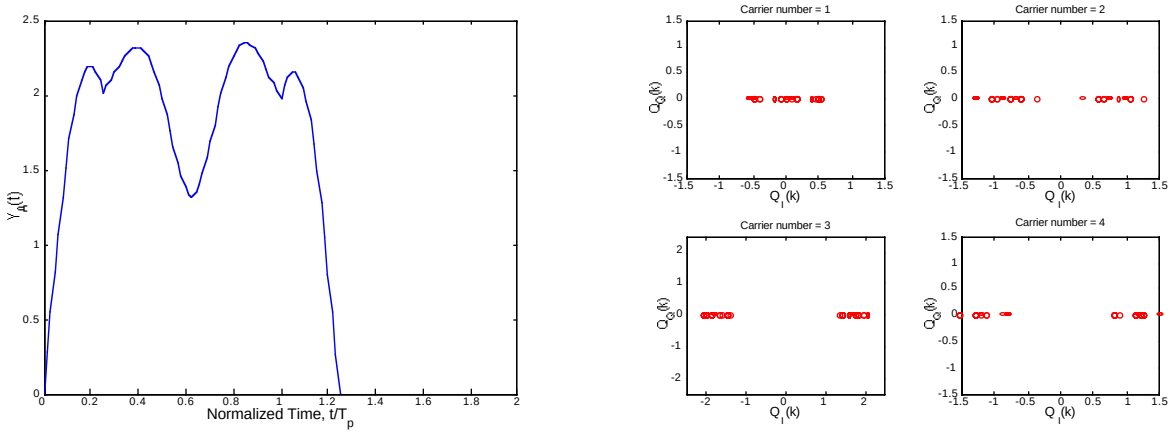
Examining (11.55) shows that the effect of a frequency selective channel on an OFDM transmitted signal is only to change the effective pulse shape for each of the subcarriers in exactly the same way as OCDM.

Example 11.8: Consider an OFDM signal with $K_b = 4$ and $f_d T_p = 0.5$ with BPSK modulation that is used in the channel introduced in Example 11.1. A sample path of the received amplitude of the OFDM signal in the frequency selective channel is shown in Fig. 11.11-a). Clearly the received signal is longer than T_p due to the dispersive nature of the frequency selective channel. The vector diagram of the four matched filter outputs is shown in Fig. 11.11-b) in the absence of noise. This figure clearly shows the loss in orthogonality of the OFDM waveform due to the frequency selectivity. The intersymbol interference is large enough for the first bit matched filter that error free bit decision cannot be made even in the absence of noise. Note by referring to Figure 11.2, that this worst performing bit is associated with the subcarrier whose frequency has the lowest channel gain.

It should be apparent from the results in Chapter 9 that as long as the orthogonality condition exists between carriers, i.e.,

$$V_{\tilde{u}}(l,k) = \int_0^{T_p+T_h} \exp\left[j2\pi(f_l - f_k)t\right] \tilde{u}(t,l)\tilde{u}^*(t,k)dt = 0 \tag{11.58}$$

then simple OFDM demodulation is possible. If $h_z(t)$ is allowed to take a general form then it is very difficult to ensure that this orthogonality condition is satisfied. As OFDM is a special case of OCDM the general discussion of optimum demodulators and simplified demodulators holds directly. In practice though none of those techniques are typically used as there is a very simple modification of OFDM transmitted waveforms that permits a very simple suboptimum demodulator on frequency selective channels.



a) Received envelope of a sample path of OFDM in a frequency selective channel.

b) Matched filter output vector diagrams.

Figure 11.11: The received OFDM signal in a frequency selective channel. $\alpha = 0.7 \exp\left[\frac{j\pi}{3}\right]$, $\tau_1 = 0.25T_p$.

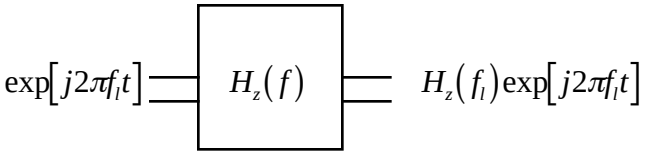


Figure 11.12: A long subcarrier pulse into a linear time–invariant system.

11.4.1 Simplified Demodulator for OFDM

A simplified but suboptimum demodulator is available for OFDM signals on frequency selective channels by augmenting the transmitted signal slightly. Insight into this suboptimum demodulator is obtained by assuming that $u(t) = 1 \forall t$. This very long pulse can be viewed as having each subcarrier being an infinite length complex sinusoid. The effective pulse for the l^{th} subcarrier then becomes

$$\tilde{u}(t,l) = \int_0^{T_h} h_z(\lambda) \exp\left[-j2\pi f_l \lambda\right] d\lambda \exp\left[j2\pi f_l t\right] = H_z(f_l) \exp\left[j2\pi f_l t\right]. \tag{11.59}$$

Note that $H_z(f_l)$ is the channel transfer function evaluated at f_l and that it is a constant with respect to time. The intuition into this is easily gained by looking at Fig. 11.12. A linear time invariant system with a complex sinusoidal input will produce a complex sinusoid at the output and the complex gain on this sinusoid will be the transfer function of the linear system evaluated at the frequency of the complex sinusoid. Hence if we have a long pulse the only change in the OFDM pulse (except for the up and down transients) for a particular subcarrier will be a multiplicative distortion due to the frequency response of the channel at the subcarrier frequency.

With this assumption of a constant pulse, the received signal will be much like the OFDM signal in a frequency flat channel. The received signal is given as

$$Y_z(t) = \sum_{l=1}^{K_b} D_z(l) H_z(f_l) \exp\left[j2\pi f_l t\right] + W_z(t). \tag{11.60}$$

Consequently the form in (11.60) is exactly the same as the original OFDM modulation with $D_z(l)$ replaced by $D_z(l)H_z(f_l)$. If a time interval, T_d is selected such that the subcarriers are orthogonal, i.e.,

$$\int_0^{T_d} \exp\left[j2\pi(f_l - f_k)t\right] dt = 0 \qquad k \neq l \tag{11.61}$$

then simple bit demodulation is possible. This time interval T_d and the frequency separation of the subcarriers is chosen exactly as in the frequency flat channel. It should be noted that a spacing of $f_d T_d = 0.25$ cannot be used with a real modulation (e.g., BPSK) on a frequency selective channel. This will become apparent as the details of the demodulator are revealed. This simple demodulator for BPSK modulation then has the form

$$\Re\left[H_z^*(f_k)\int_0^{T_d} Y_z(t)\exp\left[-j2\pi f_k t\right] dt\right] = \Re\left[\bar{Q}(k)\right] = \Re\left[H_z^*(f_k)Q(k)\right] \underset{\hat{I}(k)=1}{\overset{\hat{I}(k)=0}{\gtrless}} 0. \tag{11.62}$$

Note that we use the notation $Q(k)$ since this is the same matched filter as was used in the frequency flat channel in Chapter 9.

Why this condition results in a simplified demodulation form is shown by expanding out matched filter (to the transmitted signal) as

$$\begin{aligned}
\bar{Q}(k) &= H_z^*(f_k)Q(k) = H_z^*(f_k)\int_0^{T_d} Y_z(t)\exp\left[-j2\pi f_k t\right] dt \\
&= H_z^*(f_k)\sum_{l=1}^{K_b} D_z(l)H_z^*(f_l)\int_0^{T_d}\exp\left[j2\pi(f_l - f_k)t\right] dt + \bar{N}(k) \\
&= |H_z(f_k)|^2 D_z(k)T_d + \bar{N}(k).
\end{aligned} \tag{11.63}$$

When the pulse shape is constant and an integration interval is chosen where the subcarriers are orthogonal, all the Intercarrier Interference (ICI) is removed from each subcarrier output. It should be noted that a spacing of $f_d T_d = 0.25$ results in

$$0 = \Re\left[\int_0^{T_d}\exp\left[j2\pi(f_l - f_k)t\right] dt\right] \tag{11.64}$$

but for operation on an arbitrary frequency selective channel the necessary condition is

$$0 + j0 = \int_0^{T_d}\exp\left[j2\pi(f_l - f_k)t\right] dt \tag{11.65}$$

Consequently the only difference between the OFDM demodulator introduced in Chapter 9 and the one for the frequency selective channel is the derotation by the channel gain. The overall block diagram is shown in Fig. 11.13 and should be compared to Fig. 9.10.

Since $\bar{Q}(k)$ has the form in (11.63), a vector can be formed with analogy to the OCDM case where the output samples have the form

$$\vec{\bar{Q}} = \bar{E}_b \mathbf{G}\vec{D} + \vec{\bar{N}} = \bar{E}_b \mathrm{diag}\left[\vec{g}\right]\vec{D} + \vec{\bar{N}} \tag{11.66}$$

where

$$\bar{E}_b = \frac{T_d}{K_b}\sum_{k=1}^{K_b}|H_z(f_k)|^2 \qquad g(k) = \frac{|H_z(f_k)|^2 T_d}{\bar{E}_b}. \tag{11.67}$$

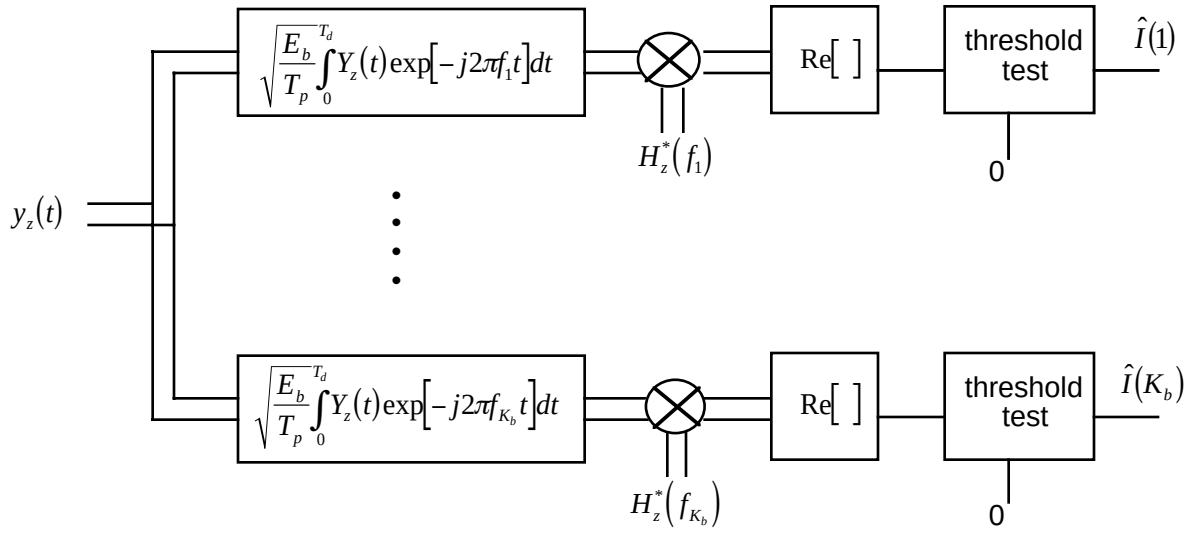


Figure 11.13: The simplified demodulator for OFDM in a frequency selective channel.

The diagonal nature of the **G** shows that the orthogonality of the subcarriers has been restored. Because of this orthogonality, except for the derotation by the channel gain this reduced complexity demodulator on the frequency selective channel has exactly the same form as the optimum demodulator for OFDM in a frequency non-selective channel.

The question remains of how the effective pulse shape, $\tilde{u}(t,l)$, can be made a constant over a time interval $[0,T_d]$. Returning to the definition of $\tilde{u}(t,l)$, it can be seen that since

$$\tilde{u}(t,l) = \int_0^{T_h} h_z(\lambda)u(t-\lambda)\exp\left[-j2\pi f_l\lambda\right]d\lambda. \tag{11.68}$$

For example if $u(t)$ is a rectangular pulse of length T_p then

$$\tilde{u}(t,l) = \int_{\max(0,t-T_p)}^{\min(t,T_h)} h_z\left(\lambda\right)\exp\left[-j2\pi f_l\lambda\right]d\lambda. \tag{11.69}$$

By examining (11.69) it is apparent that $\tilde{u}(t,l)$ will not be a function of time over the interval $[0,T_d]$ only if the limits of integration are not a function of time. The limits of integration are not a function of time only if $u(t)$ is constant over the interval $[-T_h,T_d]$. If this condition holds then the the effective pulse becomes

$$\tilde{u}(t,l) = \int_0^{T_h} h_z\left(\lambda\right)\exp\left[-j2\pi f_l\lambda\right]d\lambda = H_z\left(f_l\right) \qquad t\in[0,T_d]\,. \tag{11.70}$$

In essence the OFDM pulse must be extended by T_h in length. As is shown in Fig. 11.14 the first T_h part of the output defining the effective pulse will transient due to the impulse response of the channel. Once all of the impulse response is covered in the convolution integral the effective pulse becomes the transfer function of the channel evaluated at the subcarrier frequency. When the pulse is removed from the input then the output will have a transient response back to zero.

Consequently if $u(t)$ is made constant over the interval $[-T_h,T_d]$ then the suboptimal demodulation in (11.62) can be implemented and there will be no ICI. An example transmitted waveform that achieves

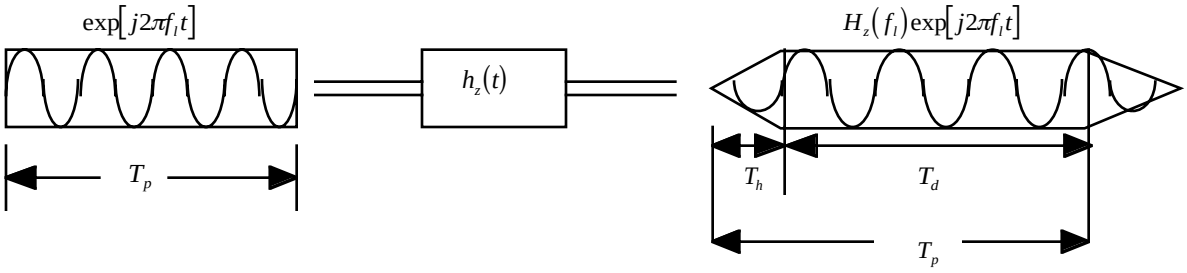


Figure 11.14: The effective pulse in a frequency selective channel.

this characteristic is

$$X_z(t) = \sum_{l=1}^{K_b} D_z(l) \exp\left[j2\pi f_l t\right] \qquad -T_h \le t \le T_d. \tag{11.71}$$

i.e.,

$$u(t) = \begin{cases} 1 & -T_h \le t \le T_d \\ 0 & \text{elsewhere} \end{cases} \tag{11.72}$$

So simply by extending the rectangular pulse used in OFDM by at least the length of the delay spread of the channel the orthogonality condition can be retained over the interval $t \in [0, T_d]$ for any frequency selective channel. In practice the cyclic prefix length is chosen to be bigger than the longest delay spread that is normally encountered in the operation of the communication system, i.e., $T_p = T_d + T_h(\max)$.

Example 11.9: Consider again an OFDM signal with $K_b = 4$ and $f_d T_d = 0.5$ with BPSK modulation that is used in the channel introduced in Example 11.1 where $\tau_1 = T_d/4 = T_h$. The effective pulse shapes for the case when the pulses are rectangular and of length T_d are plotted in Fig. 11.15. The output pulses from the channel show a transient effect until all the impulse response of the channel is covered by the input pulse ($t = T_h$). After this point the effective pulse becomes constant until the transmitted pulse goes away ($t = T_d$) at which point another transient in the effective pulse is noticable. For this channel $T_h = 0.25T_d$ and the cyclic prefix will increase the standard OFDM transmitted signal in length by 25% (i.e., $T_p = T_d + T_h$). A sample path of the received amplitude of the OFDM signal in the frequency selective channel is shown in Fig. 11.16-a). Clearly the received signal is longer than T_d ($T_p + T_h = T_d + 2T_h$) due to the dispersive nature of the frequency selective channel and the added cyclic prefix. The vector diagrams of the four matched filter outputs are shown in Fig. 11.16-b) in the absence of noise. The channel gains at each of the subcarriers are $H_z(-1.5/T_p) = 0.18e^{-j78^\circ}$, $H_z(-0.5/T_p) = 0.86e^{j51^\circ}$, $H_z(0.5/T_p) = 1.4e^{j7^\circ}$, and $H_z(1.5/T_p) = 1.1e^{-j37^\circ}$ and the constellations are exactly as predicted by theory.

The performance of the simplified demodulator for OFDM in a frequency selective channel is easily computed due to the achieved orthogonality between the subcarriers. Since the matched filter output has the form

$$Q(k) = H_z(f_k) T_d D_z(k) + N_z(k) \tag{11.73}$$

where $N_z(k)$ is a white process with $\text{var}\left(N_z(k)\right) = T_d N_0$. Assuming BPSK modulation, the optimum

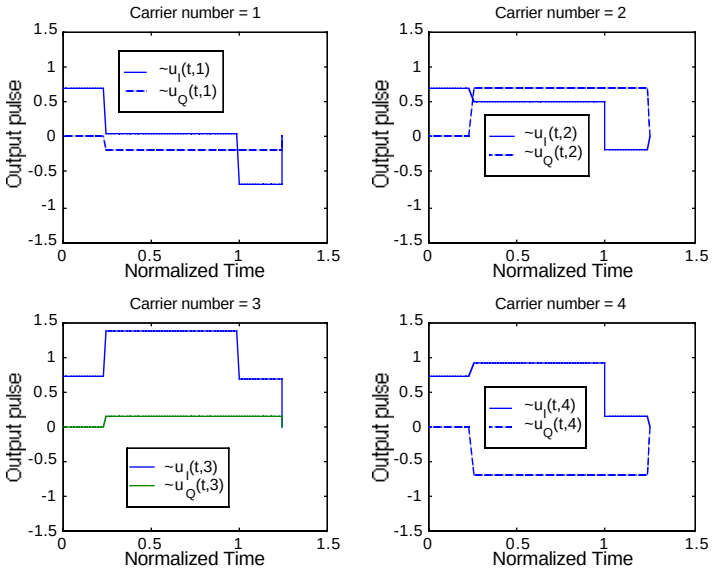
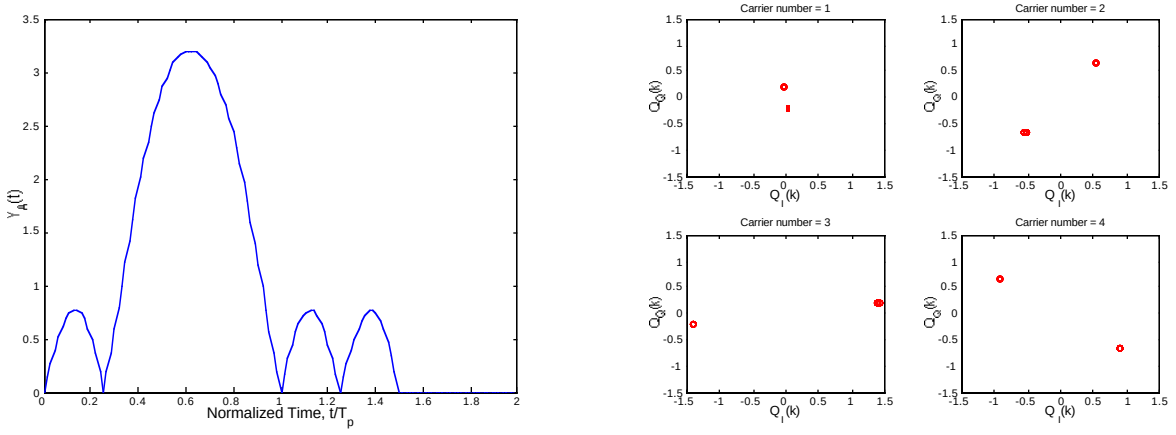


Figure 11.15: The output pulses for each OFDM subcarrier.



a) Received envelope of a sample path of OFDM in a frequency selective channel.

b) Matched filter output vector diagrams.

Figure 11.16: The received OFDM signal in a frequency selective channel with a cyclic prefix. $\alpha = 0.7 \exp\left[\frac{j\pi}{3}\right]$, $\tau_1 = 0.25T_d$.

threshold test will be

$$\Re\left[\bar{Q}(k)\right] \underset{\hat{I}(k)=1}{\overset{\hat{I}(k)=0}{\gtrless}} 0. \tag{11.74}$$

Noting that $\text{var}\left(\bar{N}(k)\right) = |H_z(f_k)|^2 T_d N_0$, it is clear that the probability of error for the k^{th} subcarrier is

$$P_B(E,k) = \frac{1}{2}\text{erfc}\left(|H_z(f_k)|\sqrt{\frac{T_d}{N_0}}\right) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{g(k)\bar{E}_b}{N_0}}\right). \tag{11.75}$$

Likewise due to the orthogonality between the noise, the word error probability is

$$P_W(E) = 1 - \prod_{k=1}^{K_b}\left(1 - P_B(E,k)\right). \tag{11.76}$$

Examining (11.75) and (11.76), it is obvious that the performance of the simplified demodulator is dominated by the subcarrier with the lowest $|H_z(f_k)|$.

Note the only difference between the original OFDM transmitted waveform introduced in Chapter 9 and the one presented in (11.71) is that the transmission time has been extended by T_h seconds. Since mathematically it makes sense to think about putting this on the frontend of the OFDM symbol (i.e., the time support has changed from $[0, T_d]$ to $[-T_h, T_d]$) this addition to the OFDM signal is most often denoted a cyclic prefix. Clearly the demodulation algorithm proposed in (11.62), while simple and eliminating ICI from each subcarrier decision metric, is suboptimal. The suboptimality results from not considering the whole received signal in the processing. The cyclic prefix is also a redundant transmission that conveys no information. Adding a cyclic prefix and using the simple demodulator results in both a loss in bandwidth efficiency and performance but in many applications the gain in demodulator simplicity in a frequency selective channel more than makes up for these losses. It should be noted that the complexity of this simple decoder for OFDM is $O(K_b)$ which is even less than the simplified demodulators presented for general OCDM. The homework will explore both the loss in performance and the loss in bandwidth efficiency of this demodulation technique for OFDM in frequency selective channels.

Example 11.10: The wireless local area network (WLAN) standard denoted IEEE 802.11a uses OFDM modulation for indoor communications. The indoor wireless channel is a frequency selective channel due to the multipath nature typical of unguided radio propagation. The 802.11a standard is set with $T_d = 3.2\mu$s, $T_h = 0.8\mu$s, and $f_d = 156.25$kHz. This design allows a path length difference between the significant multipaths of 240m before orthogonality is again lost between the subcarriers. Since WLANs are normally operated within a building 240m $= c \times T_h$ is usually much greater than the coverage area.

11.5 Conclusion

This chapter examined the structure and performance of digital communications in the presence of a frequency selective channel. The first point was that general M-ary modulations essentially had a demodulation structure that was relatively unchanged in a frequency selective channel. The only difference was that the matched filter needed to be matched to the received signal (the transmitted signal filtered by the channel). The performance of the general M-ary is again a function of the received Euclidean distance spectrum. Hence the major focus of the chapter was on orthogonal modulations. OCDM was

	Modulation Type	
	OCDM	OFDM
Optimum Demodulation	$O\left(2^{K_b}\right)$	$O\left(2^{K_b}\right)$
Suboptimal Demodulation	$O\left(K_b^2\right)$	$O\left(K_b\right)$

Table 11.1: A complexity comparison of the various orthogonal modulations on frequency selective fading channels.

examined as this is the most general form of modulation and the MLWD was derived. The loss of orthogonality due to the frequency selective channel implies that a general M-ary demodulator is needed which has a complexity $O(2^{K_b})$. This complexity is often considered too high for many implementations so several forms of reduced complexity demodulators were considered. These suboptimum demodulators had a complexity that were $O\left(K_b^2\right)$. These demodulators offer a variety of tradeoffs between performance and complexity. The special case of OFDM was considered and it was shown that a very simple suboptimum demodulator can be built for use in the frequency selective channel. This demodulator is exactly the same as the OFDM demodulator in a frequency flat channel and has a complexity $O\left(K_b\right)$. The case of stream modulations will be considered in the next chapter in detail. The complexity for the various demodulation options available with the orthogonal modulations considered in this chapter is summarized in Table 12.1.

11.6 Homework Problems

Problem 11.1. Consider 4-ary FSK as considered in Chapter 8 that is to be communicated over the two channels considered in Example 11-1 with $T_s = T_p$. Assume the frequency separation is that of the optimum separation in a frequency flat channel (see Problem 8.10).

a) Detail out the optimum demodulator for each channel.

b) Compute the union bound to the probability of word error for each of the channels and plot for $E_b/N_0 = 0 - 10$dB.

Problem 11.2. Assume that $K_b = 2$ bits of information is to be sent with a linear modulation with a QPSK constellation and a pulse shape $u(t)$ having a support of T_u. This transmission is to be on a frequency selective channel with

$$H_z(f) = \sqrt{0.51} + 0.7 \exp\left[\frac{-j\pi f T_u}{2} + \frac{j\pi}{3}\right]. \tag{11.77}$$

a) Find the optimum demodulator.

b) Find the word error probability of the optimum demodulator.

c) Find one pulse shape, $u(t)$, that would have a better word error probability performance than the rectangular pulse shape

$$u_r(t) = \begin{cases} \dfrac{1}{\sqrt{T_p}} & 0 \le t \le T_p \\ 0 & \text{elsewhere} \end{cases} \tag{11.78}$$

without increasing the 3dB transmission bandwidth by more than 50%.

Problem 11.3. This problem explores the correlation of the noise in the demodulation of OCDM in a frequency selective channel. Recall that the matched filter outputs have the form

$$\vec{\tilde{Q}} = E_b \mathbf{G} \vec{D} + \vec{\tilde{N}}. \tag{11.79}$$

For any random vector, $\vec{X}$, the correlation matrix is defined as $\mathbf{R}_X = E\left[\vec{X}\vec{X}^H\right]$.

a) Show $\mathbf{R}_{\tilde{N}} = E_b \mathbf{G} N_0$.

b) Show that the correlation matrix of the noise out of the decorrelating detector is $\mathbf{R}_{N_d} = E_b \mathbf{G}^{-1} N_0$.

c) Assume BPSK is used on each spreading waveform find the probability of error for the k^{th} bit.

d) Prove that the average probability of error is higher for a decorrelating detector in a frequency selective channel than for the optimum detector in a frequency flat channel at high SNR. *Hint: Jensen's Inequality.*

e) Give an example when the probability of error will have an error floor (a probability of error that does not decrease exponentially with SNR)

Problem 11.4. This problem explores the correlation of the noise in the MMSE demodulator of OCDM in a frequency selective channel. Recall that the matched filter outputs have the form

$$\vec{\tilde{Q}} = E_b \mathbf{G} \vec{D} + \vec{\tilde{N}}. \tag{11.80}$$

For any random vector, $\vec{X}$, the correlation matrix is defined as $\mathbf{R}_X = E\left[\vec{X}\vec{X}^H\right]$.

a) Find $\mathbf{R}_{\tilde{N}}$.

b) Find $\mathbf{R}_{N_m}$, the correlation matrix of the noise out of the MMSE detector where

$$\hat{\vec{D}} = \mathbf{W}^H \vec{\tilde{Q}} = \mathbf{W}^H \mathbf{G} \vec{D} + \vec{N}_m. \tag{11.81}$$

c) Consider Example 11.6 and assume BPSK is used on each spreading waveform with $E_b/N_0 = 10$dB. Find the probability of error of the MMSE detector for the first bit. Note this will require averaging over the 8 possible bit patterns that could be generated by $I(k),\ k \in \{2,3,4\}$.

Problem 11.5. Assume OFDM is used with a cyclic prefix and

$$h_z(t) = \frac{1}{\sqrt{2}}\delta(t) + \frac{1}{\sqrt{2}} \exp\left[j\theta\right] \delta\left(t - T_p/4\right), \tag{11.82}$$

and the OFDM scheme has $K_b = 4$, $f_d = \frac{1}{2T_p}$, and uses BPSK. Assume the demodulator is implemented using the simple scheme in Section 11.4.1

a) Compute the channel frequency response, $H_z(f)$.

b) Compute exact output for each of the four frequency matched filters.

c) Identify the simplified algorithm for making bit decisions.

d) Compute the probability of bit error for each subcarrier. Choose the value of θ such that optimizes the error performance of the second tone.

Problem 11.6. There are certain situations where transmitting a cyclic prefix with OFDM would be bandwidth inefficient (e.g., short packets in channels with long delay spreads) and this problem considers the optimal demodulator in this case. Assume

$$h_z(t) = \frac{1}{\sqrt{2}}\delta(t) + \frac{1}{\sqrt{2}}\delta\left(t - T_p/4\right), \tag{11.83}$$

and the OFDM scheme has $K_b = 4$, $f_d = \frac{1}{2T_p}$, and uses BPSK.

a) Compute the channel frequency response, $H_z(f)$.

b) Compute $V_{\tilde{u}}(1,1)$ and $V_{\tilde{u}}(2,1)$

c) Find the optimum word decision when $\vec{\tilde{Q}}_I = [0.9 \ \ -0.3 \ 0.1 \ 0.7]$. Note since the modulation is only on the I-channel the quadrature component of the matched filter output does not effect the decision.

Problem 11.7. This problem is concerned with the transmission of one bit of information, I, by BPSK modulation with a pulse shape of $u(t)$. Specifically $x_0(t) = u(t)$ and $x_1(t) = -u(t)$ during the time interval $[0, T_u]$. The transmitted signal is passed through a frequency selective channel with an impulse response given as

$$h_z(t) = \sum_{k=1}^{m_p} h_k \delta(t - \tau_k). \tag{11.84}$$

where the complex numbers h_k and the real numbers τ_k and m_p are known constants. The output of this channel is subjected to a complex additive white Gaussian noise with a one-sided spectral density of N_0.

a) What is the transfer function of the channel, $H_z(f)$.

b) The received observation when $I = m$ can be modeled as $y_z(t) = r_m(t) + W_z(t)$ where $W_z(t)$ is a complex additive white noise. Find the form for $r_0(t)$ and $r_1(t)$.

c) The optimum demodulator for the case is shown in Fig. 11.17 and has been denoted the RAKE demodulator due the fact the block diagram has a form much like a common garden rake. In this block diagram specify the operations labeled with ????.

d) What is the bit error probability performance of this system? Simplify the result as much as you can.

Problem 11.8. The company you work for, Lurid Technologies, produces a modem that transmits $K_b = 3$ bits using OFDM and BPSK on each subcarrier. The transmitted signal has the form

$$X_z(t) = \begin{cases} \displaystyle\sum_{l=1}^{3} D_z(l) \exp\left[j2\pi f_l t\right] & 0 \le t \le 10^{-3} \\ 0 & \text{elsewhere} \end{cases} \tag{11.85}$$

The demodulator structure for the modem is given in Fig. 11.18 where $T_p = 10^{-3}$.

A new client would like to use this modem to communicate over a frequency selective channel corrupted by an AWGN of one sided spectral density N_0. Your company has spent a significant amount of money to build an integrated circuit that computes the integrals in Fig. 11.18 for $T_p = 10^{-3}$ and would like to reuse the circuit in the new design.

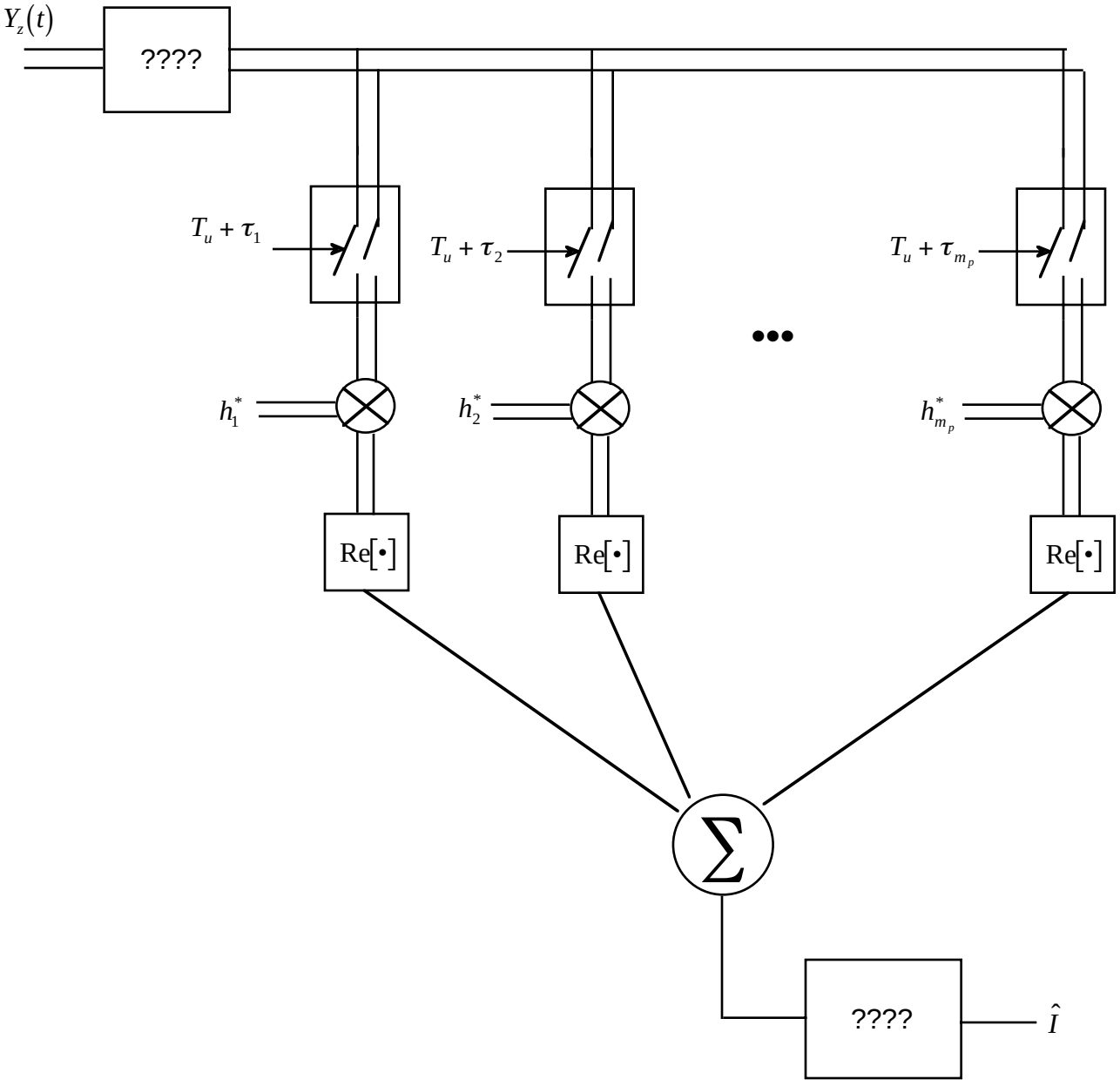


Figure 11.17: The RAKE demodulator.

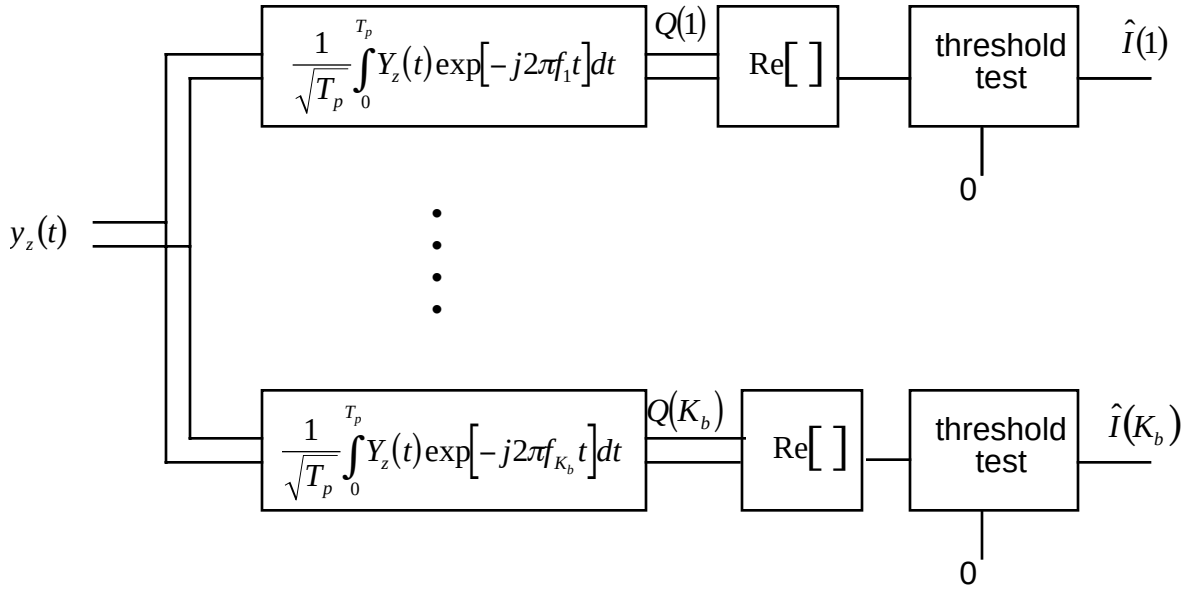


Figure 11.18: The OFDM demodulator structure.

a) A testing of the channel discovered that $T_h = 10^{-4}$ and a channel impulse response given as $h_z(t)$. Give this simplest way to change the transmitted signal and/or the decision processing to achieve the desired communication.

b) For the system specified in a) using BPSK modulation and for $H_z(f_1) = 1$, $H_z(f_2) = 0.1 + j0.1$, and $H_z(f_3) = j0.5$ detail out the decision processing.

c) For the channel given in b) find the probability of error for each bit as a function of $\bar{E}_b/N_0$. What is the word error probability?

d) Assume that you know the channel before transmitting and that you are limited to using MPSK modulation with amplitude $|D_z(k)| = \sqrt{\log_2(M)}$ on each subcarrier of an OFDM system, would there be a way to transmit three bits that would achieve a better word error probability than to put one bit on each subcarrier with BPSK modulation? If so give one example and provide the word error probability of the resulting system.

Problem 11.9. This problem is concerned with the transmission of $K_b = 4$ bits of information, $\vec{I}$, by orthogonal code division multiplexing using BPSK on each spreading waveform. Assume that the transmitted spreading waveforms, $s_l(t)$, are given in Fig. 11.19. The channel is frequency selective and the impulse response is given as

$$h_z(t) = 0.707\delta(t) + 0.707\delta(t - T_p/4) \tag{11.86}$$

a) The effective spreading waveform at the receiver is given as

$$\tilde{s}_l(t) = \int_0^{T_p/4} s_l(t-\lambda)h_z(\lambda)d\lambda. \tag{11.87}$$

Find $\tilde{s}_l(t)$ for $l \in \{1, \ldots, 4\}$.

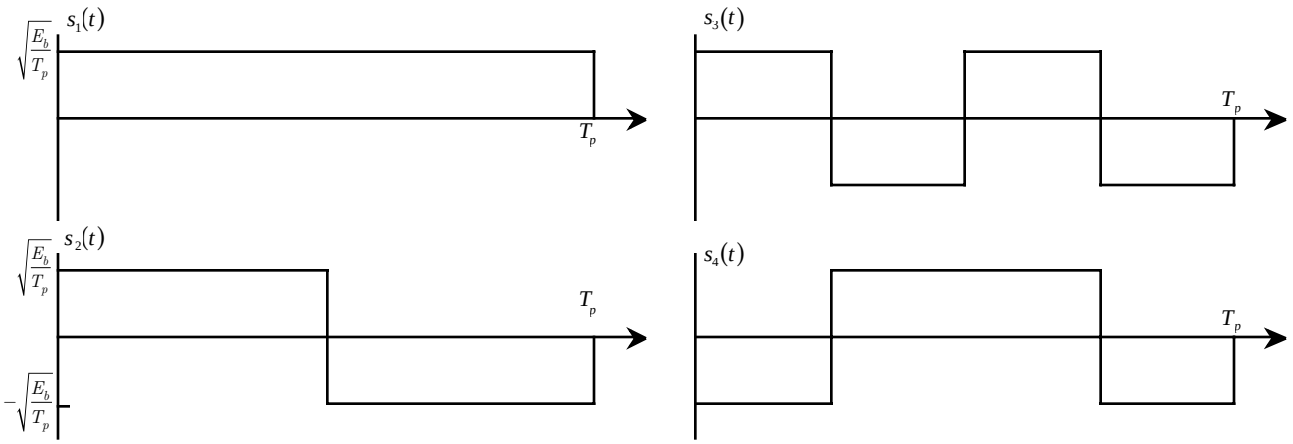


Figure 11.19: Spreading waveforms, $s_l(t)$ for OCDM and $K_b = 4$.

b) What is E_b?.

c) Show $\tilde{s}_1(t)$ still remains orthogonal to $\tilde{s}_2(t)$ on this channel.

d) Define $\tilde{Q}(k)$ as the output of the matched filter to the k^{th} effective spreading waveform and show that $\tilde{Q}(k)$ $k \in \{1, \ldots, 4\}$ are sufficient statistics for maximum likelihood word demodulation.

e) Defining the vectors

$$\vec{\tilde{Q}} = \begin{bmatrix} \tilde{Q}(1) \\ \vdots \\ \tilde{Q}(K_b) \end{bmatrix} \qquad \vec{D} = \begin{bmatrix} D(1) \\ \vdots \\ D(K_b) \end{bmatrix} \qquad \vec{\tilde{N}} = \begin{bmatrix} \tilde{N}(1) \\ \vdots \\ \tilde{N}(K_b) \end{bmatrix}, \tag{11.88}$$

the matched filter outputs are given as

$$\vec{\tilde{Q}} = E_b \mathbf{G} \vec{D} + \vec{\tilde{N}}. \tag{11.89}$$

Identify the general form for elements of the matrix $\mathbf{G}$, i.e., find $[\mathbf{G}]_{i,j}$

f) In the case of a frequency flat channel the following demodulator is optimal

$$\Re\left[Q(k)\right] \underset{\hat{I}(k)=1}{\overset{\hat{I}(k)=0}{\gtrless}} 0 \tag{11.90}$$

where $Q(k)$ is the matched filter output for the k^{th} spreading waveform. If $\tilde{Q}(k)$ is used in this same demodulator. What will be the probability of error for $k = 1$.

Problem 11.10. This problem is concerned with the transmission of $K_b = 4$ bits of information, $\vec{I}$, by orthogonal code division multiplexing using BPSK on each spreading waveform. Assume that the transmitted spreading waveforms, $s_l(t)$, are given in Fig. 11.19. The channel is frequency selective and the impulse response is given as

$$h_z(t) = 0.707\delta(t) + 0.707\delta(t - T_p/4) \tag{11.91}$$

a) Identify $\tilde{s}_l(t)$ for each $l \in \{1, \ldots, 4\}$.

b) Defining the vectors

$$\vec{\tilde{Q}} = \begin{bmatrix} \tilde{Q}(1) \\ \vdots \\ \tilde{Q}(K_b) \end{bmatrix} \qquad \vec{D} = \begin{bmatrix} D(1) \\ \vdots \\ D(K_b) \end{bmatrix} \qquad \vec{\tilde{N}} = \begin{bmatrix} \tilde{N}(1) \\ \vdots \\ \tilde{N}(K_b) \end{bmatrix}, \tag{11.92}$$

the matched filter outputs are given as

$$\vec{\tilde{Q}} = E_b \mathbf{G} \vec{D} + \vec{\tilde{N}}. \tag{11.93}$$

Identify the general form for elements of the matrix $\mathbf{G}$, i.e., find $[\mathbf{G}]_{i,j}$.

c) Show that the MLWD for this problem can be computed with a complexity that is $O(8)$ versus $O(16)$ as would be generally the case for a frequency selective channel.

d) Find the union bound for the MLWD.

e) Detail out the MLBD for $I(1)$.

Problem 11.11. Consider the problem formulation in Problem 11.19. Assume that you only have to transmit 2 bits. This 2 bit transmission is done by selecting two of the four spreading waveforms from Problem 11.19 and still using BPSK modulation.

a) Assume that you will implement a MLWD. Identify the Euclidean distance spectrum for an arbitrary 2 of the 4 possible received waveforms as a function of the elements of $\mathbf{G}$.

b) Select the two waveforms that will give the largest minimum Euclidean distance

c) Assume that you will implement a decorrelator. Identify the decorrelator structure for an arbitrary 2 of the 4 possible received waveforms as a function of the elements of $\mathbf{G}$.

d) Select the two waveforms such that the average probability of bit error for the decorrelator is minimized.

Problem 11.12. In comparing word error demodulation performance of two systems, sucessive interference cancellation with perfect feedback of the true modulation symbol and with feedback of modulation symbols decisions, select one of the following three statements as true.

1. True modulation symbol feedback performs better.
2. Both systems perform the same.
3. Modulation symbol decision feedback performs better.

Full justification of the choice must be provided.

Problem 11.13. You are asked to design an OFDM system by your boss at ModemsRus that uses independent BPSK modulation on each subcarrier operating on a frequency selective channel. She told you that for complexity purposes you should constrain the design to have a cyclic prefix and a simple bit demodulator. The channel is such that $T_h = 4\mu$s. The goal set for you by your boss is to achieve $W_b = K_b/T_p = 10$Mbps and $\eta_B = 90$ %. Assume that the transmission bandwidth is measured via the 3dB bandwidth and that rectangular pulse shapes are used on each subcarrier in the OFDM modulation. Specify all the important system parameters, e.g., f_d, T_d, and K_b.

Problem 11.14. Show that

$$\arg \max_{i=0,\ldots,M-1} \Re\left[\vec{d}_i^H \vec{\tilde{Q}}\right] - \frac{E_b \vec{d}_i^H \mathbf{G} \vec{d}_i}{2} = \arg \min_{i=0,\ldots,M-1} \left[\vec{\tilde{Q}} - E_b \mathbf{G}\vec{d}_i\right]^H \mathbf{G}^{-1} \left[\vec{\tilde{Q}} - E_b \mathbf{G}\vec{d}_i\right]. \quad (11.94)$$

This shows that the optimal demodulator is essentially a generalized minimum distance decoder in the presences of a colored noise.

Problem 11.15. Show that for OCDM when an MMSE detector is used that

$$R_E = E_b^2 \left(I_{K_b} - \mathbf{G}^H \left(\mathbf{G}\mathbf{G}^H + \frac{N_0}{E_b}\mathbf{G} \right)^{-1} \mathbf{G} \right). \quad (11.95)$$

Problem 11.16. Your boss has tasked you with designing an OFDM cable modem to send packets of $K_b = 256$ bits in less than 300μs. She wants you to use less than 1 MHz of spectrum and cause little interference to transmissions from adjacent houses (each house is frequency multiplexed). It turns out that when cable lines are shared in apartment buildings the channel becomes frequency selective. The goal given by your boss is to have a simple demodulator. Specify all the important parameters and pulse shapes.

Problem 11.17. Include decomposible errors and union bound next time.

11.7 Projects

Not included this edition.

Chapter 12

Frequency Selectivity and Stream Modulations

12.1 Current Status

Up to this point in this text we have introduced two general methods to communicate K_b bits

1. General M-ary modulations
   - The advantage of a general M-ary modulation is that it can achieve very good performance and arbitrary spectral efficiency.
   - The disadvantage of a general M-ary modulation is that without more structure the optimal demodulator has complexity $O(2^{K_b})$.
2. Orthogonal modulations (including stream modulations, OFDM, OCDM)
   - The advantage of orthogonal modulation is that the optimum demodulator has complexity $O(K_b)$ and a desired spectral efficiency can be achieved with a proper design of the modulation signals.
   - The last chapter showed how if suboptimal demodulators are considered the loss of orthogonality can be compensated by structures that have complexity $O(K_b)$ for OFDM signals and $O(K_b^2)$ for the most general signals (OCDM).

This chapter will focus on stream modulation in frequency selective channels. In particular this chapter will show

- Optimal stream demodulation on a frequency selective channel has a complexity that is $O(K_b)$.
- There are simple suboptimal demodulation structures that are practical to implement in a wide variety of communication applications.

The distortion of a frequency selective channel forces optimum bit demodulation for stream modulations away from the simple form that was presented in Chapter 9. In particular this chapter examines a class of communication problems where binary linear stream modulation is used but Nyquist criterion for zero ISI cannot be satisfied (the orthogonality condition). This subject is the focus of a separate chapter because stream modulation has traditionally been most frequently used in applications. In this formulation the transmitted signal is a binary linear stream modulation

$$X_z(t) = \sum_{l=1}^{K_b} D_z(l) u(t - (l-1)T) \tag{12.1}$$

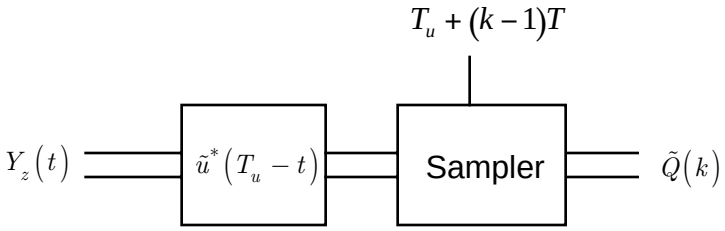


Figure 12.1: The recursive front-end processor for stream modulations on frequency selective channels.

where $D_z(l) = a\left(I(l)\right)$ with $a(\bullet)$ being the constellation mapping and $u(t)$ being the unit energy pulse shape with support of length T_u and $I(k)$ are modeled as independent and identically distributed random variables. Recall for binary linear stream modulations $T_p = (K_b - 1)T + T_u$. As before the MLWD has the form

$$\hat{\vec{I}} = \arg \max_{i \in \{0,\ldots,M-1\}} \Re \left[\int_0^{T_p+T_h} Y_z(t) r_i^*(t) dt \right] - \frac{\tilde{E}_i}{2} \tag{12.2}$$

and the text will examine ways of implementing this demodulator efficiently. Again stream modulation is a special case of OCDM so the optimum and suboptimum demodulation discussion could carry through exactly as in Section 11.3. Since the data is being carried on time offset pulses, the optimum and suboptimum demodulation structures can be reinterpreted as filters and recursive processors in the time domain. Consequently this section will examine the signal models and demodulators presented in Section 11.3 in more detail to get this insight.

The output signal has the form

$$Y_z(t) = R_z(t) + W_z(t) = \sum_{l=1}^{K_b} D_z(l) \tilde{u}(t - (l-1)T) + W_z(t) \tag{12.3}$$

where $\tilde{u}(t) = \int_0^{T_h} h_z(\lambda) u(t-\lambda) d\lambda$. $\tilde{u}(t)$ is denoted the effective pulse shape in the frequency selective channel. It can easily be shown that with the assumption of the independent data bits that $E_b = E_{\tilde{u}}$. In the same manner as above the effective pulse shape matched filter sampled outputs are the sufficient statistics, i.e.,

$$\tilde{Q}(k) = \int_0^{T_p+T_h} Y_z(t) \tilde{u}^*(t - (k-1)T) dt. \qquad k \in \{1, \ldots, K_b\}. \tag{12.4}$$

These K_b samples are sufficient statistics since a matched filter to any received waveform can be derived from these samples, i.e.,

$$\int_0^{T_p+T_h} Y_z(\tau) r_i^*(\tau) d\tau = \sum_{l=1}^{K_b} D_z^*(l) \int_0^{T_p+T_h} Y_z(t) \tilde{u}^*(t - (l-1)T) dt = \sum_{l=1}^{K_b} D_z^*(l) \tilde{Q}(l). \tag{12.5}$$

Finally these sufficient statistics can be generated in a time recursive fashion using one filter. Fig. 12.1 shows this time recursive front-end processor. One reason for the historically dominance stream modulation in practice is because the matched filter processing of stream modulation can be done sequentially in time.

It is instructive to examine the form of these matched filter outputs to the effective pulse shape to understand how frequency selective channels affect stream modulations. Substituting (12.3) into (12.4)

shows the matched filter output has the form

$$\begin{aligned}\tilde{Q}(k) &= \sum_{l=1}^{K_b} D_z(l) V_{\tilde{u}}((k-l)T) + \tilde{N}(k) \\ &= E_b \sum_{l=1}^{K_b} D_z(l) g_{k,l} + \tilde{N}(k) \end{aligned} \tag{12.6}$$

where $\tilde{N}(k)$ is an additive non-white[1] Gaussian noise and $V_{\tilde{u}}(mT)$ is the autocorrelation function of the effective pulse shape. Again resorting to the vector notation of Section 12.3 the matched filter outputs are given as

$$\vec{\tilde{Q}} = E_b \mathbf{G} \vec{D} + \vec{\tilde{N}}. \tag{12.7}$$

Because $h_z(t)$ is dispersive, Nyquist criterion, $V_{\tilde{u}}(mT) = E_{\tilde{u}}\delta_m$, cannot be guaranteed. Since $[\mathbf{G}]_{k,l} = g_{k,l} = V_{\tilde{u}}((k-l)T)/E_{\tilde{u}}$ the matrix $\mathbf{G}$ is Hermitian symmetric and Toeplitz [LT85]. Consequently the rows of $\mathbf{G}$ are mostly different only in a horizontal shift. Also because of the energy normalization, $g_{k,k} = 1$ $k \in \{1, \ldots, K_b\}$.

[1]The correlation of the noise is explored in the homework.

Example 12.1: Consider a linear stream modulated signal with $K_b = 4$ with BPSK modulation and

$$u(t) = \begin{cases} \sqrt{\frac{1}{T_u}} & 0 \le t \le T_u \\ 0 & \text{elsewhere} \end{cases} \tag{12.8}$$

where $T_u = T_p/4 = T$ that is used in the channel introduced in Example 12.1. The effective pulse for the stream modulation is given as

$$\tilde{u}(t) = \alpha_0 u(t) + \alpha_1 u(t - \tau_1) \tag{12.9}$$

The correlation function that determines the energy correction term is given as

$$\begin{aligned} V_{\tilde{u}}(\tau) &= \int_{-\infty}^{\infty} \tilde{u}(t)\tilde{u}^*(t-\tau)dt \\ &= V_u(\tau) + \alpha_0\alpha_1^* V_u(\tau + \tau_1) + \alpha_1\alpha_0^* V_u(\tau - \tau_1) \end{aligned} \tag{12.10}$$

where

$$V_u(\tau) = \begin{cases} 1 - \frac{|\tau|}{T_u} & |\tau| \le T_u \\ 0 & \text{elsewhere.} \end{cases} \tag{12.11}$$

Again assuming that $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = 0.25T_p = T_u = T$ the matrix in the energy correction term is given as

$$\begin{aligned} G &= \begin{bmatrix} 1 & \alpha_0^*\alpha_1 & 0 & 0 \\ \alpha_0\alpha_1^* & 1 & \alpha_0^*\alpha_1 & 0 \\ 0 & \alpha_0\alpha_1^* & 1 & \alpha_0^*\alpha_1 \\ 0 & 0 & \alpha_0\alpha_1^* & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0.4999\exp[j\pi/3] & 0 & 0 \\ 0.4999\exp[-j\pi/3] & 1 & 0.4999\exp[j\pi/3] & 0 \\ 0 & 0.4999\exp[-j\pi/3] & 1 & 0.4999\exp[j\pi/3] \\ 0 & 0 & 0.4999\exp[-j\pi/3] & 1 \end{bmatrix} \end{aligned} \tag{12.12}$$

The energy correction term for the MLWD using (12.3) is

$$\begin{aligned} \tilde{E}_i &= \int_0^{T_p+T_h} |r_i(t)|^2\, dt = \sum_{l_1=1}^{K_b}\sum_{l_2=1}^{K_b} d_i(l_1)d_i^*(l_2)\int_0^{T_p+T_h} \tilde{u}(t-(l_1-1)T)\tilde{u}^*(t-(l_2-1)T)dt \\ &= \sum_{l_1=1}^{K_b}\sum_{l_2=1}^{K_b} d_i(l_1)d_i^*(l_2)V_{\tilde{u}}\left((l_2-l_1)T\right) \end{aligned} \tag{12.13}$$

In a way similar to the development for OCDM the energy correction term for stream modulations can be expressed in matrix notation as

$$\tilde{E}_i = E_b \vec{d}_i^H \mathbf{G} \vec{d}_i. \tag{12.14}$$

Typically when a stream modulation is used in practice both the length of the pulse for transmission, T_u, and the multipath impulse response length, T_h, are much shorter than the transmission time, T_p.

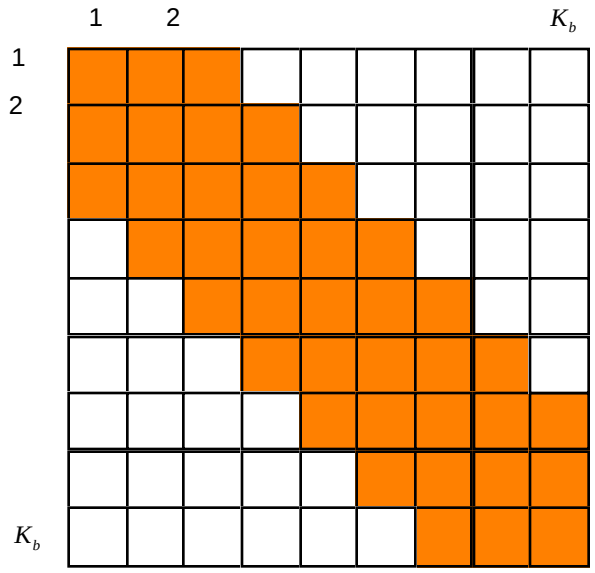


Figure 12.2: The form of **G** for $K_b = 9$ and $N_u = 2$. The potentially non-zero elements are shaded

This fact implies that the ISI extends only over a fairly short interval since

$$V_{\tilde{u}}(\tau) = 0 \qquad |\tau| > T_u + T_h. \tag{12.15}$$

Consequently **G** will be a matrix having a band diagonal property. Banded diagonal implies that a matrix only takes non-zero values in a finite region around the diagonal of the matrix. The symbol support of $\tilde{u}(t)$ determines how many off diagonal elements will occur in each row.

Definition 12.1 *The symbol support, N_u, of the $\tilde{u}(t)$ is the largest integer such that $T_u + T_h > N_u T$.*

Now precisely **G** will have N_u off diagonal elements. Fig. 12.2 graphically shows the case for $K_b = 9$ and $N_u = 2$

Example 12.2: Consider the situation presented in Example 12.1 where $T_u = T_h = T$. Here $T_u + T_h = 2T_u > T$ but $T_u + T_h = 2T_u = 2T$ so that $N_u = 1$.

12.2 MLWD for Stream Modulations

Since the matched filter outputs, $\tilde{Q}(k)$, are computed sequentially in time using the same filter, it would be useful to formulate a method of computing the MLWD recursively in time. A technique that accomplishes this was identified by Ungerboeck [Ung74]. To derive this recursive MLSD recall the form for the MLWD is

$$\hat{\vec{I}} = \arg \max_{i \in \{0,\ldots,M-1\}} T_i \tag{12.16}$$

$$= \arg \max_{i \in \{0,\ldots,M-1\}} \sum_{k=1}^{K_b} \Re\left[d_i^*(k)\tilde{Q}(k)\right] - \frac{E_b}{2} \sum_{l_1=1}^{K_b} \sum_{l_2=1}^{K_b} d_i(l_1) d_i^*(l_2) g_{l_2 l_1}$$

$$= \arg \max_{i \in \{0,\ldots,M-1\}} \Re\left[\vec{d}_i^H \vec{\tilde{Q}}\right] - \frac{E_b \vec{d}_i^H \mathbf{G} \vec{d}_i}{2} \tag{12.17}$$

where $d_i(l)$ is the data symbol at the l^{th} symbol time when $\vec{I} = i$. It should be noted that the matched filter term in (12.16) is already in a form that can be computed in a time recursive fashion. Consequently the effort needs to be focused on computing the energy correction term in a time recursive fashion. Ungerboeck came to a realization that the energy correction term for each postulated sequence is a form that is a sum of terms $d_i(l_1)d_i^*(l_2)g_{l_2 l_1}$. Each of these terms is nonzero only when $g_{l_2 l_1}$ is nonzero. Hence the energy correction term for each postulated transmitted word only has to sum over the shaded components of G as shown in Fig. 12.2. While this simplifies the computation of the ML metric it does not address the exponential complexity of computing the ML metric for each possible transmitted data word.

Example 12.3: This example continues Example 12.1 which has $K_b = 4$, $T_u = T_p/4$, $\alpha_1 = 0.7\exp[j\pi/3]$, and BPSK modulation with $E_b = 1$. Recall that this text uses $d_0(k) = 1$ and $d_1(k) = -1$ for BPSK modulation. This example produces a $\mathbf{G}$ matrix of the form

$$G = \begin{bmatrix} 1 & 0.4999\exp[j\pi/3] & 0 & 0 \\ 0.4999\exp[-j\pi/3] & 1 & 0.4999\exp[j\pi/3] & 0 \\ 0 & 0.4999\exp[-j\pi/3] & 1 & 0.4999\exp[j\pi/3] \\ 0 & 0 & 0.4999\exp[-j\pi/3] & 1 \end{bmatrix}$$

For the example considered here the observations are

$$\vec{\tilde{q}} = [\tilde{q}(1)\ \tilde{q}(2)\ \tilde{q}(3)\ \tilde{q}(4)]^T = [1.5 - j0.5\ 0.3 + j0.7\ 1.2 + j0.1\ -0.9 + j0.1] \tag{12.18}$$

and the goals is to efficiently form the MLWD decision given as

$$T_i = \arg \max_{i \in \{0,\ldots,M-1\}} \Re\left[\vec{d}_i^H \vec{\tilde{q}}\right] - \frac{E_b}{2}\vec{d}_i^H \mathbf{G}\vec{d}_i. \tag{12.19}$$

This optimum demodulator can be viewed as computing the maximum likelihood metric on a binary tree of depth K_b. An example of such a tree for K_b=4 is shown in Fig. 12.3. The form of the tree has a great deal of symmetry. For example the top half and the bottom half of the tree are exactly the same, with the only difference being the first bit. This symmetry can be exploited in forming the MLWD metric. Ungerboeck proposed an algorithm that had a recursion at each time instant which partitioned the K_b bits into four distinct sets as shown in Fig. 12.4. The "past" bits in the algorithm have had conditional decision made on them based on the past observations and the possible future bits. There is a current bit that will be processed along with the current observation to produce a subsequent conditional decision. There are the N_u "future" bits needed to make the subsequent conditional decision. Finally there are the "future" bits not needed in making the conditional decision. The finite memory of the ISI and the symmetry in the tree structure were the characteristics that Ungerboeck exploited to find a demodulation structure that had a complexity $O\left(N_u 2^{N_u+1} K_b\right)$. It should be noted that if $K_b \gg N_u$ then the complexity is essentially again only growing linearly with the number of bits transmitted, as desired. The Ungerboeck algorithm has a two step recursion

1. Forward step

2. Survivor selection

and the remainder of this section will provide the details of this algorithm.

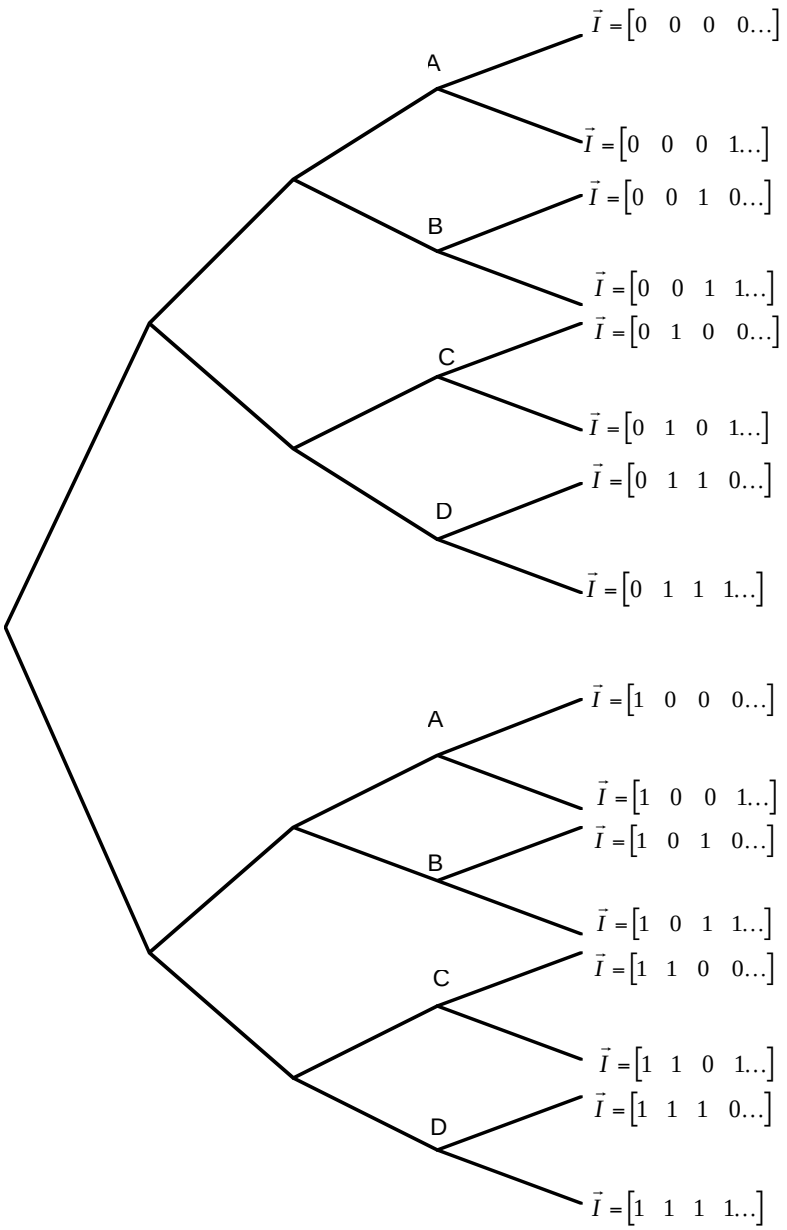


Figure 12.3: The binary tree for enumerating all possible words in an M-ary demodulator. K_b=4.

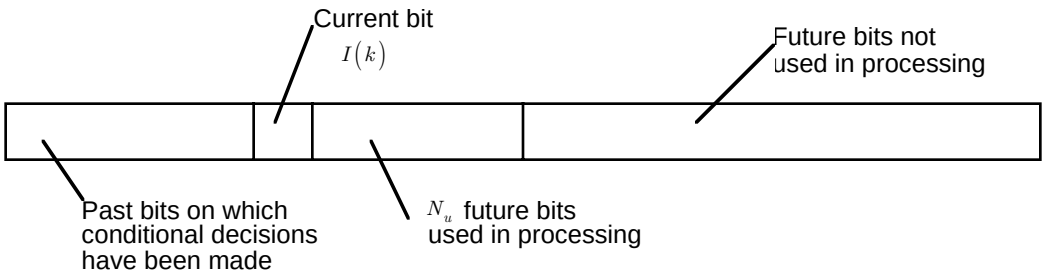


Figure 12.4: The partitioning of the data frame for each step of the MLWD as proposed by Ungerboeck.

12.2.1 Forward Step–Time 1

Since there are no "past" bits at the first step, Ungerboeck's form for the MLWD can be derived by reformulating the MLWD in (12.16) by separating out the maximization over $I(1) = m_1$ and $\vec{I}^{(-)}(1) = n(1)$. Suppose there is a split of the ML metric such that $T_i = T^{(-)}_{n(1)}(1) + T^{(+)}_{m_1,n(1)}(1)$ then

$$\begin{aligned}\hat{\vec{I}} &= \arg \max_{i\in\{0,\ldots,M-1\}} T_i = \arg \max_{\substack{m_1=0,1\\ n(1)=0,M/2-1}} T_{\{m_1,n(1)\}}(1) \\ &= \arg\left\{ \max_{n(1)=0,M/2-1} \left(T^{(-)}_{n(1)}(1) + \max_{m_1=0,1} T^{(+)}_{m_1,n(1)}(1) \right) \right\}.\end{aligned} \tag{12.20}$$

Consequently if we can isolate everything in the ML metric that is a function of the first bit we can make a conditional decision on the first bit once the first matched filter output is observed. Manipulating (12.16) and using the finite memory of $\tilde{u}(t)$ gives

$$\begin{aligned}\hat{\vec{I}} &= \arg\left\{ \max_{n(1)=0,M/2-1} \left(\sum_{k=2}^{K_b} \Re\left[d^*_{n(1)}(k)\tilde{Q}(k) \right] - \frac{E_b}{2} \sum_{l_1=2}^{K_b}\sum_{l_2=2}^{K_b} d_{n(1)}(l_1) d^*_{n(1)}(l_2) g_{l_1,l_2} + \right.\right. \\ &\left.\left. \max_{m_1=0,1} \left\{ \Re\left[d^*_{m_1}(1)\tilde{Q}(1) \right] - \frac{E_b}{2} \left(|d_{m_1}(1)|^2 + \sum_{k_1=2}^{N_u+1} d_{n(1)}(k_1) d^*_{m_1}(1) g_{k_1,1} + \sum_{k_2=2}^{N_u+1} d_{m_1}(1) d^*_{n(1)}(k_2) g_{1,k_2} \right) \right\} \right) \right\}\end{aligned} \tag{12.21}$$

From (12.21) it is apparent that

$$T^{(+)}_{m_1 n(1)}(1) = \Re\left[d^*_{m_1}(1)\tilde{Q}(1) \right] - \frac{E_b}{2} \left(|d_{m_1}(1)|^2 + \sum_{k_1=2}^{N_u+1} d_{n(1)}(k_1) d^*_{m_1}(1) g_{k_1,1} + \sum_{k_2=2}^{N_u+1} d_{m_1}(1) d^*_{n(1)}(k_2) g_{1,k_2} \right). \tag{12.22}$$

$T^{(+)}_{m_1,n(1)}(1)$ is not a function of all the $K_b - 1$ future possible bits but only N_u future bits, $\vec{I}_{N_u}(1) = [I(2)\ I(3) \ldots I(N_u+1)]$. Notationally we will enumerate the possible values that these N_u bits can take with $n_1 = 0, 2^{N_u} - 1$ and $\vec{I}^{(-)}(1) = n(1) = \left\{ n_1, n_1^{(-)} \right\}$.

Example 12.4: Consider when $K_b = 4$ then the notation would be

$I(1)$	$I(2)$	$I(3)$	$I(4)$	m_1	$n(1)$	n_1	$n_1^{(-)}$
0	0	0	0	0	0	0	0
0	0	0	1	0	1	0	1
0	0	1	0	0	2	0	2
0	0	1	1	0	3	0	3
0	1	0	0	0	4	1	0
0	1	0	1	0	5	1	1
0	1	1	0	0	6	1	2
0	1	1	1	0	7	1	3
1	0	0	0	1	0	0	0
1	0	0	1	1	1	0	1
1	0	1	0	1	2	0	2
1	0	1	1	1	3	0	3
1	1	0	0	1	4	1	0
1	1	0	1	1	5	1	1
1	1	1	0	1	6	1	2
1	1	1	1	1	7	1	3

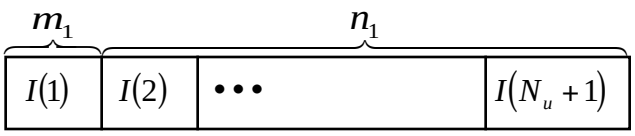


Figure 12.5: The enumeration of the first $N_u + 1$ bits.

The components of the energy correction term that $T^{(+)}_{m_1,n_1}(1)$ corresponds to are shown in Fig. 12.6 with a different shading. These terms are all the possible terms in the double summation that are a function of $d_i(1)$. Since the ISI support in a stream modulation is N_u to enumerate all the values that these components of the energy correction term represent require all of the possible values of the first $N_u + 1$ modulation symbols to be enumerated. Consequently the first step in a MLWD is to compute $T^{(+)}_{m_1,n_1}(1)$ for all possible bit values corresponding to $m_1 = 0, 1,\ n_1 = 0, 2^{N_u} - 1$. It should be noted that m_1 enumerates the value of the first bit, while n_1 enumerates the value of the next N_u bits. A graphic that represents how the notation enumerates the bits is shown in Fig. 12.5. Consequently the partial metric is given as

$$\begin{aligned} T^{(+)}_{m_1 n_1}(1) &= \Re\left[d^*_{m_1}(1)\tilde{Q}(1)\right] \\ &\quad -\frac{E_b}{2}\left(|d_{m_1}(1)|^2 + \sum_{k_1=2}^{N_u+1} d_{n_1}(k_1)d^*_{m_1}(1)g_{k_1,1} + \sum_{k_2=2}^{N_u+1} d_{m_1}(1)d^*_{n_1}(k_2)g_{1,k_2}\right). \end{aligned} \tag{12.23}$$

These 2^{N_u+1} values of $T^{(+)}_{m_1,n_1}(1)$ will be denoted the forward metrics of the Ungerboeck form of the MLWD at the first step. At the first time, these metrics can be thought of as being computed by pushing forward along all branches into the tree that represents the data word (see Fig. 12.3) to a depth of $N_u + 1$. Considering the form of (12.23), the complexity of this forward step is $O(N_u 2^{N_u+1})$.

Example 12.5: Continuing with Example 12.3 the forward step at time 1 can be detailed. Note that since $N_u = 1$, that m_1 enumerates the possible values for the first bit and n_1 enumerates that possible values of the second bit. The partial ML metric for the possible values of n_1 and m_1 are given as

$$T^{(+)}_{m_1 n_1}(1) = \Re\left[d^*_{m_1}(1)\tilde{q}(1)\right] - \frac{E_b}{2}\left[1 + d_{n_1}(2)d^*_{m_1}(1)g_{12} + d_{m_1}(1)d^*_{n_1}(2)g_{21}\right] \tag{12.24}$$

Consequently the four metrics that result in the forward step are

$$\begin{aligned} T^{(+)}_{00}(1) &= \Re\left[\tilde{q}(1)\right] - \frac{E_b}{2}\left[1 + g_{21} + g_{12}\right] = 0.75005 \\ T^{(+)}_{10}(1) &= \Re\left[-\tilde{q}(1)\right] - \frac{E_b}{2}\left[1 - g_{21} - g_{12}\right] = -1.75005 \\ T^{(+)}_{01}(1) &= \Re\left[\tilde{q}(1)\right] - \frac{E_b}{2}\left[1 - g_{21} - g_{12}\right] = 1.24995 \\ T^{(+)}_{11}(1) &= \Re\left[-\tilde{q}(1)\right] - \frac{E_b}{2}\left[1 + g_{21} + g_{12}\right] = -2.24995. \end{aligned} \tag{12.25}$$

Since $N_u = 1$ the first two bits are considered in computing the partial ML metric and a total of four sequences must be computed on the forward step.

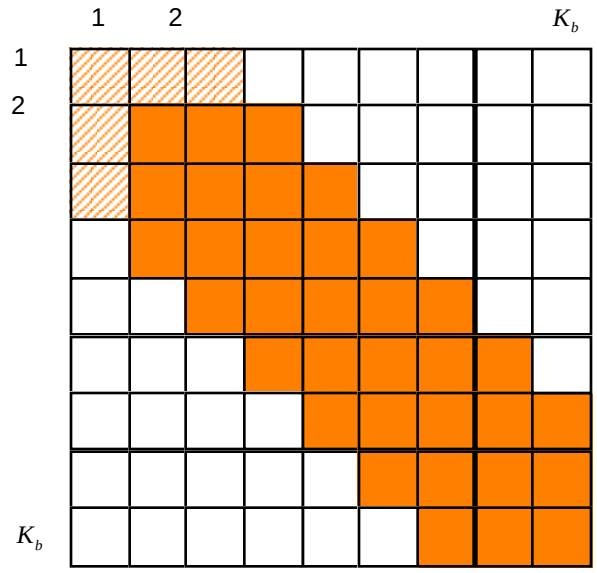


Figure 12.6: The form of $\mathbf{G}$ for $K_b = 9$ and $N_u = 2$. The lightly shaded elements of $\mathbf{G}$ are in $T_{m_1,n}(1)$ and the dark shaded elements of $\mathbf{G}$ are in $T_n^{(-)}(1)$.

12.2.2 Survivor Selection–Time 1

It is possible to prune the number of paths considered after the forward step. Since $T_{n(1)}^{(-)}(1)$ is not a function of m_1 it will be common for all sequences that have a common index $n(1)$ that enumerates the future bits, $[I(2)\dots I(K_b)]$. Since the forward step has pushed N_u+1 branches into the tree, there will be 2^{N_u} pairs of these 2^{N_u+1} paths having a common $T_{n(1)}^{(-)}$. For example the sequence $\vec{I} = [0000\dots0]$ will produce the same value of $T_{n(1)}^{(-)}$ as would the sequence $\vec{I} = [1000\dots0]$. In fact all of these pairs of possible data words that have a common $T_{n(1)}^{(-)}$ can be shown to be extensions of two specific paths created in the forward step. For instance for N_u=2 all the paths in the tree of Fig. 12.3 produced in the forward step labeled with the same letter have a common future, $[I(2)\dots I(K_b)]$. Since the forward paths in the tree are enumerated with $\{m_1, n(1)\}$ the paths with a common future will be paths with a common n_1. This common future implies that to find the ML data word we do not need to keep all of the paths for future processing. For instance if $T_{0,0}^{(+)}(1) > T_{1,0}^{(+)}(1)$ then it will never be possible for a sequence of the form $\vec{I} = [100\dots]$ to have a larger ML metric than a sequence of the form $\vec{I} = [000\dots]$ because of the common future $T_{n(1)}^{(-)}$. The forward step creates 2^{N_u+1} branches in the tree that defines the transmitted data word and a metric associated with each of those branches. It is possible to prune back to 2^{N_u} branches that could possibly become the ML demodulated sequence (make a conditional decision on $I(1)$). Again a graphic that demonstates this idea of making a conditional decision based on the future bits is shown in Fig. 12.7. For each enumeration of n_1 the values of $T_{0,n_1}^{(+)}(1)$ and $T_{1,n_1}^{(+)}(1)$ can be compared. Only the larger of these two partial metrics needs to be saved since no path associated with the smaller of these two metrics can become the MLWD output. Mathematically the surviving conditional metric and the conditional first bit decision is given as

$$\tilde{T}_{n_1}(1) = \max_{m_1=0,1} T_{m_1,n_1}^{(+)}(1) \qquad \hat{I}_{n_1}(1) = \arg\max_{m_1=0,1} T_{m_1,n_1}^{(+)}(1). \tag{12.26}$$

The survivor selection step identifies and saves the maximum $T_{m_1,n_1}^{(+)}(1)$ and the path that is still a possible ML demodulated sequence for each n_1. To define an algorithm we have denoted the largest metric as $\tilde{T}_{n_1}(1)$ and the associated conditional decision as $\hat{\vec{I}}_{n_1}(1)$. This "surviving" metric and path

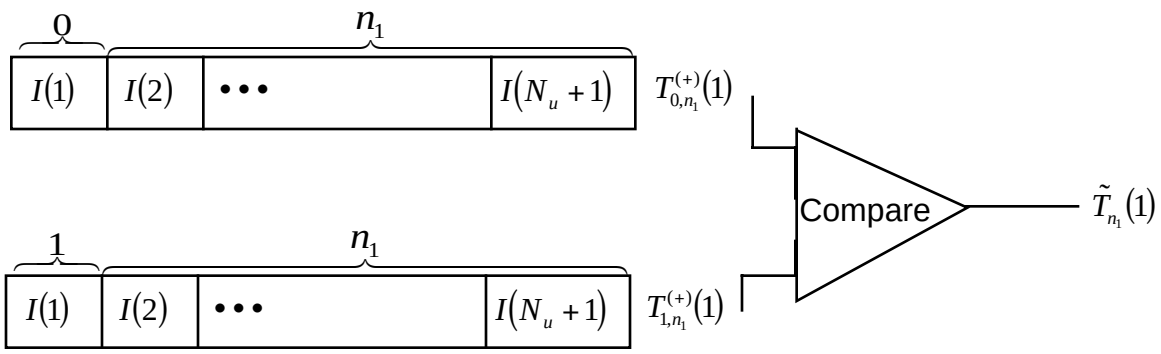


Figure 12.7: Survivor selection stage for time 1.

that achieves this metric are saved for further processing. This completes the first stage of the Ungerboeck MLWD.

Example 12.6: For Example 12.3, the first survivor selection stage can be detailed. Note that $T^{(+)}_{0,0}(1)$ and $T^{(+)}_{1,0}(1)$ will have all the possible futures sequences in common and hence a common value of $T^{(-)}_{n(1)}(1)$. Consequently when $\hat{I}(2) = 0$ the best estimate of $I(1)$ is $\hat{I}(1) = 0$ since

$$T^{(+)}_{0,0}(1) = 0.75005 > T^{(+)}_{1,0}(1) = -1.75005. \tag{12.27}$$

The selected survivor metric for the next stage is

$$\tilde{T}_0(1) = T^{(+)}_{0,0}(1) = 0.75005 \qquad \hat{\vec{I}}_0(1) = [0]\,. \tag{12.28}$$

Similarly since

$$T^{(+)}_{0,1}(1) = 1.24995 > T^{(+)}_{1,1}(1) = -2.24995 \tag{12.29}$$

the second survivor metric for the next stage is

$$\tilde{T}_1(1) = T^{(+)}_{0,1}(1) = 1.24995 \qquad \hat{\vec{I}}_1(1) = [0]\,. \tag{12.30}$$

In this example the decision on the first bit was identical for both of the possibilities of the second bit. This is not true in general. The Ungerboeck equalizer in the survivor seclection has made a conditional decision on the first bit. Consequently two of the four partial ML metrics that have been computed can be discarded with the knowledge that they will not result in the largest likelihood metric. The two surviving partial metrics are saved and will be used in future processing. A graphic that represents the binary tree and the paths searched is given in Fig. 12.8 where the "X" represents the eliminated paths.

12.2.3 Forward Step – General

The goal now is to generalize the algorithm to an arbitrary time, $k = 2, K_b - N_u$. The general algorithm block diagram is given in Fig. 12.9. The discussion in the sequel will focus on $k = 2$ but the extensions to other k should be apparent. At this point we have 2^{N_u} paths of depth $N_u + 1$ in the tree[2] that are

[2]For arbitrary k we still will have 2^{N_u} paths but now with a depth of $N_u + k - 1$.

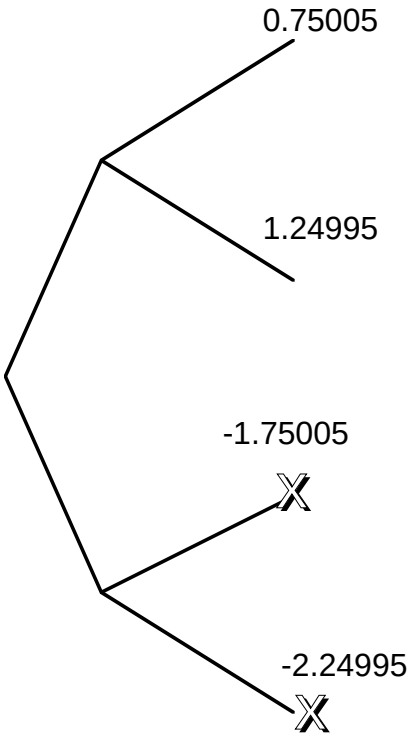


Figure 12.8: Step 1 of the Ungerboeck recursion for $N_u = 1$ and Example 12.13.

still possible ML demodulated sequences. These paths correspond to the most likely sequence for each possible value of $[I(2)\ I(3) \ldots I(N_u+1)]$ ($[I(k)\ I(k+1) \ldots I(k+N_u-1)]$ for general k). Associated with each of these paths we have the largest accumulated partial ML metric for this set of bits and the past bit decisions corresponding to the most likely sequence associated with this set of bits. The goal now is to extend this process one time iteration and to do this we note that $\vec{I}^{(-)}(1) = n(1) = \left\{n_1, n_1^{(-)}\right\} = \{m_2, n(2)\}$ and $n_1 = \{m_2, \tilde{n}_1\}$. This notation for the partitioning of the bits is graphically illustrated in Fig. 12.10 for the case of the second time step.

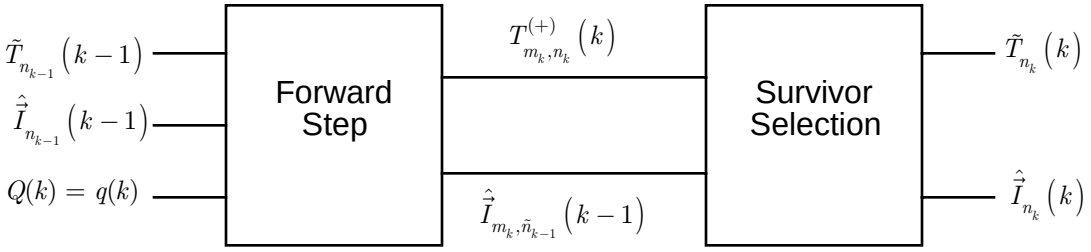


Figure 12.9: The Ungerboeck MLWD recursion.

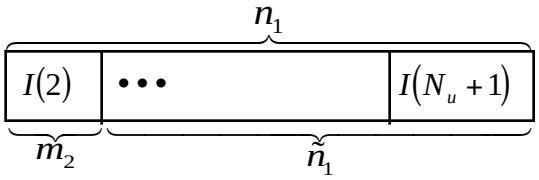


Figure 12.10: The notation for the partitioning of the bits in the second time step.

Example 12.7: Consider when $N_u = 2$ then the notation would be

n_1	$I(2)$	$I(3)$	m_2	$\tilde{n}_1$
0	0	0	0	0
1	0	1	0	1
2	1	0	1	0
3	1	1	1	1

The MLWD can be rewritten by separating out the terms that are a function of $I(2)$, i.e.,

$$\begin{aligned}
\hat{\vec{I}} &= \arg \max_{i\in\{0,\ldots,M-1\}} T_i = \arg \max_{n(1)\in\{0,\ldots,M/2-1\}} \left(T^{(-)}_{n(1)}(1) + \max_{m_1=0,1} T_{m_1,n_1}(1) \right) \\
&= \arg \max_{\substack{m_2=0,1 \\ n(2)\in\{0,\ldots,M/4-1\}}} T^{(-)}_{m_2,n(2)}(1) + \tilde{T}_{m_2,\tilde{n}_1}(1) \qquad (12.31) \\
&= \arg \max_{n(2)\in\{0,\ldots,M/4-1\}} \left(T^{(-)}_{n(2)}(2) + \max_{m_2=0,1} \left(T_{m_2,n(2)}(2) + \tilde{T}_{m_2,\tilde{n}_1}(1) \right) \right) \\
&= \arg \max_{n(2)\in\{0,\ldots,M/4-1\}} \left(T^{(-)}_{n(2)}(2) + \max_{m_2=0,1} T^{(+)}_{m_2,n(2)}(2) \right)
\end{aligned}$$

where we have defined

$$T^{(+)}_{m_2,n(2)}(2) = T_{m_2,n(2)}(2) + \tilde{T}_{m_2,\tilde{n}_1}(1) = T_{m_2,n_2}(2) + \tilde{T}_{m_2,\tilde{n}_1}(1). \qquad (12.32)$$

The second step forward pushes the enumerated paths in the tree one branch further (up to $I(2+N_u)$) hence the term $T_{m_2,n(2)}(2)$ does not have to be computed for all $[I(2)\ I(3)\ldots I(K_b)]$ but only for all $[I(2)\ I(3)\ldots I(N_u+2)]$. We will again enumerate all the sequences in $[I(2)\ I(3)\ldots I(N_u+2)]$ with $m_2 = 0, 1,\ n_2 = 0, 2^{N_u}-1$. It should be again noted that m_2 enumerates the value of the second bit, while n_2 enumerates the value of the next N_u bits. Consequently by computing the values of $T^{(+)}_{m_2,n(2)}(2)$, or equivalently again $T^{(+)}_{m_2,n_2}(2)$ due to the finite length of the ISI, the MLWD metric can be found. This enumeration of all possible bit sequences in this time step 2 update is illustrated in Fig. 12.11. The partial metrics computed the previous stage are only a function of $I(2)$ through $I(N_u+1)$ (enumerated by m_2 and $\tilde{n}_1$). The new terms that need to be added to the partial metric are a function of $I(2)$ through $I(N_u+2)$ (enumerated by m_2 and n_2). Consequently the forward step recursion in Fig. 12.9 is completely determined by computing the function $T_{m_2,n_2}(2)$ and adding it to the appropriate metric from the previous stage as in (12.32).

The term $T_{m_2,n_2}(2)$ is often denoted the branch metric. Again it is apparent that the new terms needed in $T^{(+)}_{m_2,n(2)}(2)$ compared to what were computed the first iteration of the algorithm are all the

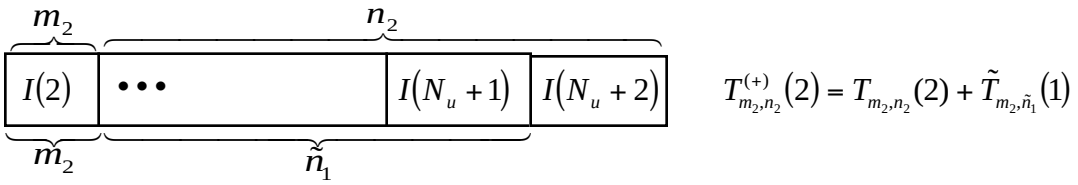


Figure 12.11: The time step 2 update of the MLWD.

terms involving $D_z(2)$ or equivalently $I(2)$ that were not included in $\tilde{T}_{n_1}(1)$ and these are given as

$$T_{m_2,n_2}(2) = \Re\left[d^*_{m_2}(2)\tilde{Q}(2)\right] - \frac{E_b}{2}\left(\sum_{k_1=1}^{N_u+1} d_{n_2}(k_1+1)d^*_{m_1}(2)g_{k_1+1,2} + \sum_{k_2=2}^{N_u+1} d_{m_2}(2)d^*_{n_2}(k_2+1)g_{2,k_2+1}\right). \tag{12.33}$$

The terms in $T_{m_2,n_2}(2)$ are the terms containing the second matched filter output and the energy correction terms represented by the checkered pattern in Fig. 12.12. Consequently we have again grown the size of the number of considered sequences to 2^{N_u+1} and again the complexity of this forward step is $O(N_u 2^{N_u+1})$. This completes the forward step as shown in Fig. 12.9 as $T^{(+)}_{m_2,n_2}(2)$ now contains all the terms of the ML metric that are related to $I(1)$ and $I(2)$. It should be noted that m_2 enumerates the value of the second bit, while n_2 enumerates the value of the next N_u bits. The survivor selection can again make conditional decisions to shrink down the number of possible words that might be the maximum likelihood word.

It is worth it to detail out the forward step algorithm for the case for a general k for clarity of discussion. The partial accumulated metric for n_k becomes

$$T^{(+)}_{m_k,n_k}(k) = T_{m_k,n_k}(k) + \tilde{T}_{m_k,\tilde{n}_{k-1}}(k-1). \tag{12.34}$$

This forward step only requires the computation of the branch metric, $T_{m_k,n_k}(k)$ and the adding of this branch metric to the appropriate survivor partial metric from the previous stage, $\tilde{T}_{m_k,\tilde{n}_{k-1}}(k-1)$. Again we now have 2^{N_u+1} surviving metrics corresponding to the possible values of m_k and n_k and a surviving word of the form $\hat{\vec{I}}(k) = [m_k \ \hat{\vec{I}}_{\tilde{n}_{k-1}}(k-1)]$. These surviving partial metrics and surviving possible ML sequences are passed to the survivor selection portion of the algorithm.

Example 12.8: For Example 12.3, the second forward stage is detailed in this example. Since $N_u = 1$ in this example, n_1 only enumerates the value of the second bit and there is no $\tilde{n}_1$. Recall the surviving metrics from the first stage are

$$\begin{aligned} \tilde{T}_0(1) &= 0.75005 \qquad \hat{\vec{I}}_0(1) = [0] \\ \tilde{T}_1(1) &= 1.24995 \qquad \hat{\vec{I}}_1(1) = [0]. \end{aligned} \tag{12.35}$$

In a similar development as Example 12.14 we have

$$T_{m_2,n_2}(2) = \Re\left[d^*_{m_2}(2)\tilde{q}(2)\right] - \frac{E_b}{2}\left[1 + d_{n_2}(3)d^*_{m_2}(2)g_{23} + d_{m_2}(2)d^*_{n_2}(3)g_{32}\right]. \tag{12.36}$$

Due to the Hermitian symmetric and Toeplitz nature $\mathbf{G}$, the four branch metrics for the second update are

$$\begin{aligned} T_{0,0}(2) &= \Re\left[\tilde{q}(2)\right] - \frac{E_b}{2}\left[1 + g_{21} + g_{12}\right] = -0.44995 \\ T_{1,0}(2) &= \Re\left[-\tilde{q}(2)\right] - \frac{E_b}{2}\left[1 - g_{21} - g_{12}\right] = -0.55005 \\ T_{0,1}(2) &= \Re\left[\tilde{q}(2)\right] - \frac{E_b}{2}\left[1 - g_{21} - g_{12}\right] = 0.04995 \\ T_{1,1}(2) &= \Re\left[-\tilde{q}(2)\right] - \frac{E_b}{2}\left[1 + g_{21} + g_{12}\right] = -1.04995. \end{aligned}$$

Adding these updates to the partial metrics saved from the previous stages provides new total partial metrics for stage two, i.e.,

$$\begin{aligned} T^{(+)}_{0,0}(2) &= T_{0,0}(2) + \tilde{T}_0(1) = 0.3001 \\ T^{(+)}_{1,0}(2) &= T_{1,0}(2) + \tilde{T}_1(1) = 0.6999 \\ T^{(+)}_{0,1}(2) &= T_{0,1}(2) + \tilde{T}_0(1) = 0.8 \\ T^{(+)}_{1,1}(2) &= T_{1,1}(2) + \tilde{T}_1(1) = 0.2. \end{aligned}$$

We now have four partial metrics corresponding to the most likely data sequences. All values of $I(2)$ and $I(3)$ are considered and conditional decisions have been made on $I(1)$ in computing these partial metrics. Note the algorithm has effectively returned to the same situation as was reached in (12.25).

12.2.4 Survivor Selection – General

It is again possible to prune the number of paths considered after the forward step. The forward step now has created 2^{N_u+1} metrics corresponding to the branches in the tree, $[I(2)\ I(3)\ldots I(2+N_u)]$ ($[I(k)\ I(k+1)\ldots I(k+N_u)]$ for general k) and past conditional decisions associated with each of those branches. It is possible to prune back to 2^{N_u} branches that could possibly become the ML demodulated sequence by again realizing that $[0\ I(3)\ldots I(N_u+2)]$ and $[1\ I(3)\ldots I(N_u+2)]$ have a common future and values of $T^{(-)}_n(2)$. The partial accumulated metric for n_2 becomes

$$\tilde{T}_{n_2}(2) = \max_{m_2=0,1}\left(T_{m_2,n_2}(2) + \tilde{T}_{m_2,\tilde{n}_1}(1)\right) = \max\left(T^{(+)}_{0,n_2}(2), T^{(+)}_{1,n_2}(2)\right). \tag{12.37}$$

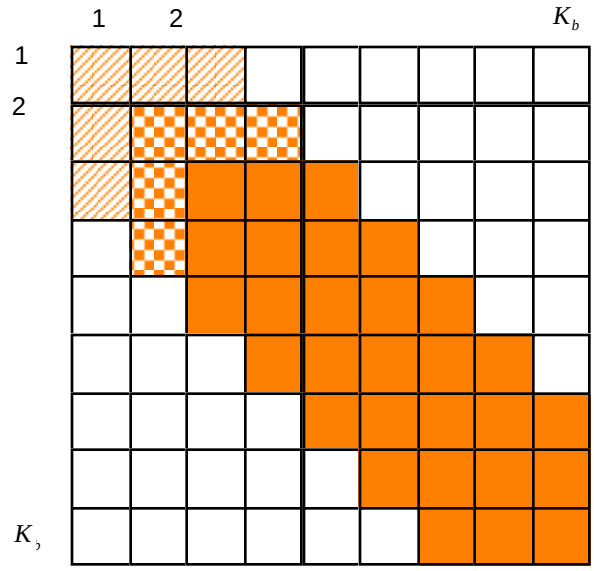


Figure 12.12: The form of $\mathbf{G}$ at $k = 2$ where the cross–hatched shaded elements of $\mathbf{G}$ are in $\tilde{T}_{n_1}$ and the checkerboard shaded elements are in $T_{m_2,n(2)}(2)$ and the dark shaded elements of $\mathbf{G}$ are in $T_n^{(-)}(2)$.

For instance, when $N_u = 2$,

$$\tilde{T}_0(2) = \max\left(T_{0,0}(2) + \tilde{T}_0(1), T_{1,0}(2) + \tilde{T}_1(1)\right). \tag{12.38}$$

Similarly the conditional decisions can be made on the second bit by defining

$$\hat{I}_{n_2}(2) = \arg \max_{m_2=0,1} T_{m_2,n_2}(2) + \tilde{T}_{m_2,\tilde{n}_1}(1) \tag{12.39}$$

so that $\hat{\vec{I}}_{n_2}(2) = [\hat{I}_{\hat{I}_{n_2}(2),\tilde{n}_1}(1)\ \hat{I}_{n_2}(2)]$. This selection corresponds to picking the largest partial ML metric conditioned on the last $K_b - 2$ bits of the transmitted word. The accumulated two conditionally decoded bits are the ones that won in both the first and second stage of the survivor selection. This "surviving" metric and path that achieves this metric are consequently saved for further processing. This completes the survivor selection and the recursion of the Ungerboeck MLWD.

Example 12.9: For Example 12.3, the second survivor selection stage is detailed in this example. Proceeding as in the first survivor selection stage, the surviving two metrics are

$$\tilde{T}_0(2) = \max\left\{T_{0,0}^{(+)}(2), T_{1,0}^{(+)}(2)\right\} = 0.6999 \qquad \tilde{T}_1(2) = \max\left\{T_{0,1}^{(+)}(2), T_{1,1}^{(+)}(2)\right\} = 0.8. \tag{12.40}$$

The surviving two sequences are

$$\hat{\vec{I}}_0 = [0\ 1] \qquad \hat{\vec{I}}_1 = [0\ 0] \tag{12.41}$$

Note in this step the two surviving sequences are different. A graphic that represents the binary tree and the paths searched is given in Fig. 12.13 where the "X" represents the eliminated paths.

It is worth it to detail out the survivor selection algorithm for the case for a general k for clarity of

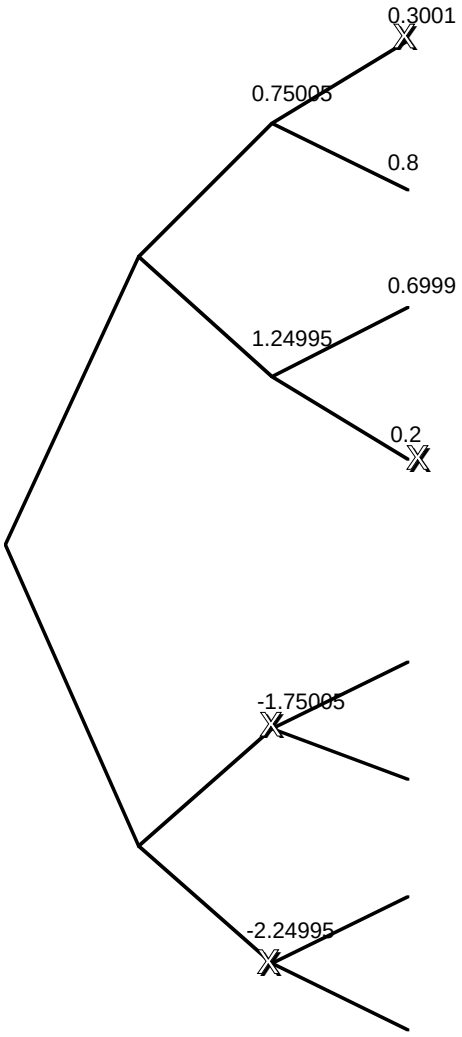


Figure 12.13: Step 2 of the Ungerboeck recursion for $N_u = 1$ and Example 12.13.

discussion. The partial accumulated metric for n_2 becomes

$$\tilde{T}_{n_k}(k) = \max_{m_k=0,1} \left(T_{m_k,n_k}(k) + \tilde{T}_{m_k,\tilde{n}_{k-1}}(k-1) \right) = \max_{m_k=0,1} T^{(+)}_{m_k,n_k}(k) = \max \left\{ T^{(+)}_{0,n_k}(k), T^{(+)}_{1,n_k}(k) \right\}. \tag{12.42}$$

The conditional decisions can be made on the second bit by defining

$$\hat{I}_{n_k}(k) = \arg \max_{m_k=0,1} T_{m_k,n_k}(k) + \tilde{T}_{m_k,\tilde{n}_{k-1)}}(k-1) = \arg \max_{m_k=0,1} T^{(+)}_{m_k,n_k}(k) \tag{12.43}$$

This can be combined with the previous conditional decisions as

$$\hat{\vec{I}}_{n_k}(k) = [\hat{\vec{I}}_{\hat{I}_{n_k}(k),\tilde{n}_{k-1}}(k-1) \ \hat{I}_{n_k}(k)] \tag{12.44}$$

This conditional decision is added to the previous decisions that have provided the largest partial MLWD metric up until this point in the recursion. Again we now have 2^{N_u} surviving metrics and surviving sequences that can be passed on to the next stage of the recursive algorithm.

The Ungerboeck MLWD has a recursion that exhibits an "add, compare, select" (ACS) structure. The add corresponds to the addition of the previous stage surviving partial metrics with the updates corresponding to the new bit. This "add" stage is detailed in (12.32). The "compare" stage makes the conditional decision on which partial metric is largest and is detailed in (12.42). The "select" stage is given in (12.43) and selects and adds to the surviving bit sequence (see (12.44)). This ACS architecture is very important in digital communications and is often included in digital processors as a fundamental operation.

Example 12.10: For Example 12.3, the third and final recursion is detailed in this example. The four branch metrics for the third update are

$$\begin{aligned} T_{0,0}(3) &= 0.45005 \qquad T_{1,0}(3) = -1.45005 \\ T_{0,1}(3) &= 0.94995 \qquad T_{1,1}(3) = -1.94995 \end{aligned}$$

Adding these updates to the partial metrics saved from the previous stages provides new total partial metrics for stage two, i.e.,

$$\begin{aligned} T^{(+)}_{0,0}(3) &= T_{0,0}(3) + \tilde{T}_0(2) = 1.14995 \\ T^{(+)}_{1,0}(3) &= T_{1,0}(3) + \tilde{T}_1(2) = -0.65005 \\ T^{(+)}_{0,1}(3) &= T_{0,1}(3) + \tilde{T}_0(2) = 1.64985 \\ T^{(+)}_{1,1}(3) &= T_{1,1}(3) + \tilde{T}_1(2) = -1.14995. \end{aligned}$$

The survivor selection step results in surviving metrics of

$$\tilde{T}_0(3) = 1.14995 \qquad \tilde{T}_1(3) = 1.64985. \tag{12.45}$$

and surviving sequences of

$$\hat{\vec{I}}_0 = [0 \ 1 \ 0] \qquad \hat{\vec{I}}_1 = [0 \ 1 \ 0]. \tag{12.46}$$

12.2.5 Termination

After the last forward step (corresponding to $k = K_b - N_u$) the entire tree has been enumerated and the remainder of the ML metrics must be computed to find the ML word decision. At this point there remains 2^{N_u} words that could possibly be the ML decision. For simplicity of notation we enumerate these 2^{N_u} words with n_t and the remaining components of the ML metric that have not been computed are given as

$$T_{n_t}(K_b - N_u) = \sum_{k=K_b-N_u+1}^{K_b} \Re\left[d_{n_t}^*(k)\tilde{Q}(k)\right] - \frac{E_b}{2} \sum_{l_1=K_b-N_u+1}^{K_b} \sum_{l_2=K_b-N_u+1}^{K_b} d_{n_t}(l_1) d_{n_t}^*(l_2) g_{l_1,l_2}. \quad (12.47)$$

The values of $T_{n_t}(K_b - N_u)$ should be computed and added to $\tilde{T}_{n_t}$ to get the ML metric for each of the remaining sequences. Identifying the word corresponding to the maximum ML metric is the most likely transmitted word. It should be noted that the termination computations can be accomplished in a more efficient manner than the enumeration discussed here. The termination can be accomplished in N_u stages that eliminates half of the words at each stage. This will be investigated in the homework.

Example 12.11: For Example 12.3, the termination is detailed in this example. Since $N_u = 1$ the only term left in the ML metric computation is

$$T_{m_4}(4) = \Re\left[d_{m_4} q(4)\right] - \frac{E_b}{2} \quad (12.48)$$

and this becomes the final branch metric. At this point there are only two word left that could be the ML word, [0 1 0 0] and [0 1 0 1]. The final ML metrics for these two words are

$$\begin{aligned} T_{[0\ 1\ 0\ 0]} &= \tilde{T}_0(3) + T_0(4) = 1.14995 - 1.4 = -0.025005 \\ T_{[0\ 1\ 0\ 1]} &= \tilde{T}_1(3) + T_1(4) = 1.64985 + 0.4 = 2.04985. \end{aligned} \quad (12.49)$$

Consequently the ML word decision is

$$\hat{\vec{I}} = [0\ 1\ 0\ 1]. \quad (12.50)$$

There were four steps in the MLWD algorithm, one for each of the K_b bits. Each step added a matched filter output. The energy correction term has 10 components in the summation as represented by the shaded regions in Fig. 12.14. Each step adds the terms as numbered in Fig. 12.14. A graphic that represents the binary tree and the paths searched is given in Fig. 12.15 where the "X" represents the eliminated paths.

12.2.6 Ungerboeck MLWD Summary

The MLWD algorithm for binary linear stream modulations in a frequency selective channel has a computationally efficient implementation. In computing the ML metric contributions at time k, a forward step is performed where 2^{N_u+1} paths in the tree defining the data word will be enumerated and a partial ML metric will be computed for each path. The complexity of this forward step is $O(N_u 2^{N_u+1})$. After forming these partial ML metrics, conditional decisions can be made on the oldest bit of the bits considered. This conditional decision reduces the number of possible words that can be the ML word back to 2^{N_u}. This process of a forward step and a survivor selection continues until N_u symbols before the end of the frame where a termination process is initiated to find the ML word. The overall complexity of this algorithm is $O(K_b N_u 2^{N_u+1})$ and when $K_b \gg N_u$ this has the desired form

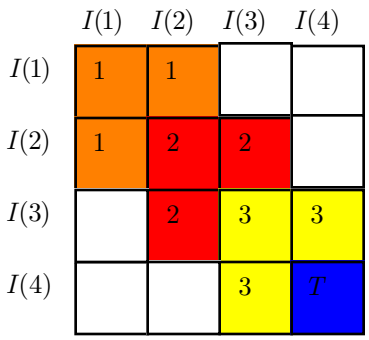


Figure 12.14: A graphic representing the energy correction term.

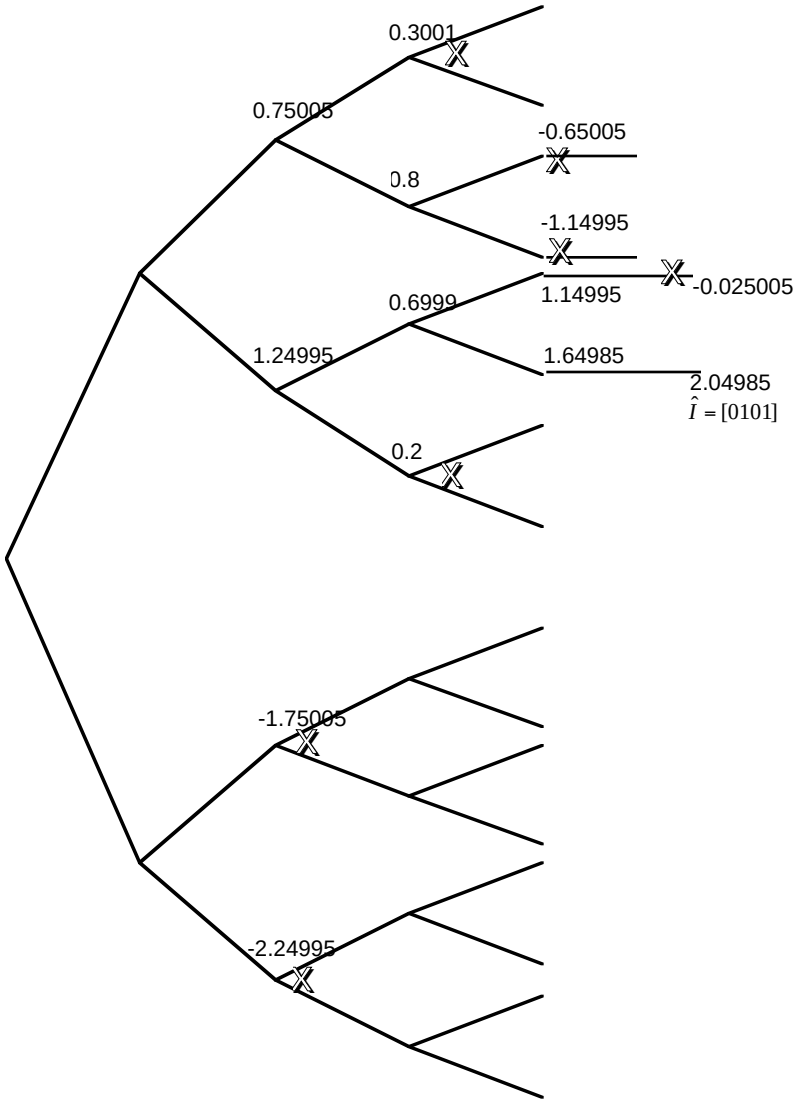


Figure 12.15: The full Ungerboeck computation for $N_u = 1$ and Example 12.13.

of having a complexity that is linear in the number of bit to be transmitted. The ACS structure of the Ungerboeck MLWD for stream modulations in frequency selective channels is important enough to have been implemented as a fundamental operation in digital processors.

12.2.7 MLWD Performance

The performance of a MLWD for stream modulation can again be assessed by looking at the union bound. This union bound is a function of the Euclidean distance spectrum. Due to the form of $\mathbf{G}$ there is significant structure and symmetry in the Euclidean distance spectrum that can be exploited in the distance computation. This chapter will not detail out the performance of MLWD for stream modulations as the following chapter will provide a more unified treatment of MLWD for a wider variety of situations and the performance will be explored using this more unified treatment.

12.3 Equalization

In practice there is often a desire to reduce the complexity of demodulation in frequency selective channels. Recall optimum demodulation for a binary stream modulation has a complexity that is $O\left(K_b 2^{N_u+1}\right)$. If the effective channel length is long, i.e., $N_u \gg 1$, then often complexity become prohibitive for low cost applications. Alternatively, if an M-ary modulation is employed on a frequency selective channel then the complexity becomes $O\left(K_b M^{N_u+1}\right)$. Consequently bandwidth efficient communication on frequency selective channels often has an optimum demodulator with a very high complexity. The engineering tradeoffs in cost often require sub-optimum solutions to be explored. In this section we will explore alternatives that have seen utility in practice for linear stream modulations. The architectures that are discussed here are directly analogous to the demodulator architectures that have been explored for OCDM on frequency selective channels as highlighted in Section 11.3.2.

Since in stream modulations the bit index is equivalent to a time index, it is useful to think about the processing as filtering operations. Recall the matched filter output is given as

$$\begin{aligned}\tilde{Q}(k) &= \sum_{l=1}^{K_b} D_z(l) V_{\tilde{u}}((k-l)T) + \tilde{N}(k) \\ &= E_b \sum_{l=1}^{K_b} D_z(l) g(k-l) + \tilde{N}(k) = E_b \sum_{l=-N_u}^{N_u} D_z(k-l) g(l) + \tilde{N}(k) \end{aligned} \tag{12.51}$$

where

$$\tilde{N}(k) = \int_0^{T_p+T_h} W_z(t) \tilde{u}^*(t-(k-1)T) dt \tag{12.52}$$

and $g(l)$ corresponds to elements of each row of the $\mathbf{G}$ matrix l positions from the diagonal. It is worth noting that given the formulation in this chapter that $g(l) = V_{\tilde{u}}(lT)/V_{\tilde{u}}(0)$ is the normalized autocorrelation function of the effective pulse shape so that $g(0) = 1$. Consequently by examining (12.51) it is apparent that $\tilde{Q}(k)$ can be viewed as being the result of putting a time series $D_z(k)$ through a known discrete time finite impulse response filter and observing the output in the presence of a colored noise. Denote the impulse response representing this filter as $g(k)$ and the transfer function as $G\left(e^{j2\pi f}\right) = \sum_k g(k) e^{-j2\pi f k}$. Equalization is typically viewed as a signal processing technique that provides reliable decisions in the presence of this frequency selective channel. Fig. 12.16 shows the model for equalization in frequency selective channels.

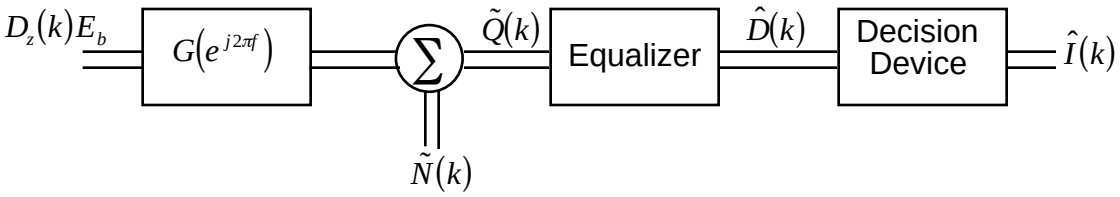


Figure 12.16: The discrete time model for stream modulation in a frequency selective channel.

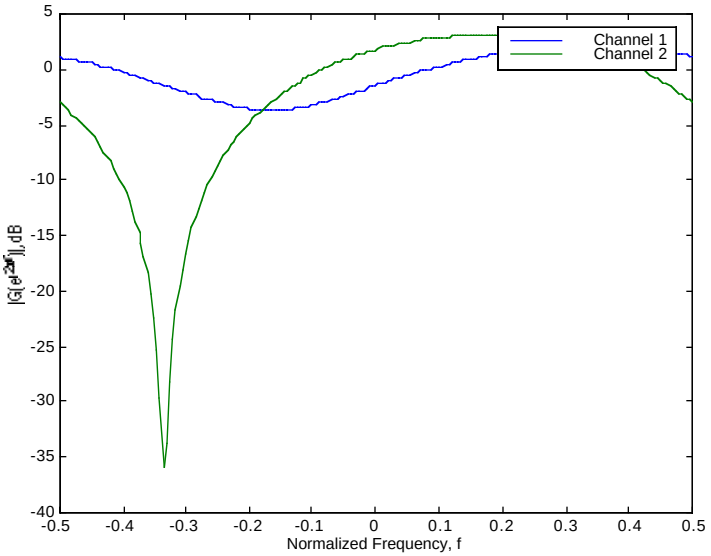


Figure 12.17: Discrete time equivalent channel transfer function.

Example 12.12: Consider the BPSK stream modulation and the channel given in Example 12.1, where $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = T$. The form of the channel given in (12.12) shows that

$$g(-1) = 0.4999\exp[-j\pi/3] \qquad g(0) = 1 \qquad g(1) = 0.4999\exp[j\pi/3]\,. \tag{12.53}$$

Consequently

$$\begin{aligned} G\left(e^{j2\pi f}\right) &= 0.4999\exp[j(2\pi f - \pi/3)] + 1 + 0.4999\exp[-j(2\pi f - \pi/3)] \\ &= 0.51\left(1 + 0.9802\exp[j(2\pi f - \pi/3)]\right)\left(1 + 0.9802\exp[-j(2\pi f - \pi/3)]\right). \end{aligned} \tag{12.54}$$

The magnitude of the discrete time equivalent channel for the two channels plotted in Fig. 11.2 is plotted in Fig. 12.17 (channel 1 is $\alpha_1 = 0.3\exp[j2\pi/3]$ and $\tau_1 = T$ and channel 2 is $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = T$). One of the discrete time channels shown in Fig. 12.17 has a deep notch in the frequency domain while the other is less frequency selective.

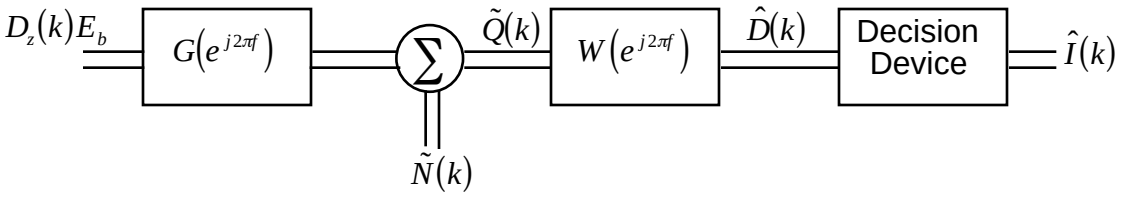


Figure 12.18: The linear equalizer for stream modulation.

12.3.1 Zero Forcing Equalizer

The first equalizers that were proposed were linear equalizers (discrete time linear filters) operating on the matched filter outputs. The form of the linear equalizer is shown in Fig. 12.18. A linear equalizer's transfer function will be denoted with $W\left(e^{j2\pi f}\right)$. The obvious first choice in equalizers would be one that restores Nyquist criterion for zero ISI [Luc65] i.e.,

$$W\left(e^{j2\pi f}\right) = \frac{1}{G\left(e^{j2\pi f}\right)}. \tag{12.55}$$

This equalizer is usually denoted the zero forcing (ZF) equalizer as it forces the intersymbol interference to zero. Since $G\left(e^{j2\pi f}\right)$ is finite impulse response the ZF solution, $W\left(e^{j2\pi f}\right)$, is an infinite impulse response filter but the impulse response decays at an exponential rate. Consequently the ZF equalizer can be well approximated with an finite impulse response. The traditional rule of thumb is that the linear zero forcing equalizer should be 3-5 times the length of the intersymbol interference, N_u. This implies that if $K_b \gg N_u$ then the complexity of the linear equalizer is $O(K_b)$ as desired.

This equalizer for stream modulation is directly analogous to the decorrelating detector introduce in Section 11.3.2. Envision K_b large, with k selected in the middle of the transmitted frame. Since for stream modulation each row of the $\mathbf{G}$ matrix is a shifted version of $g(l)$ or the impulse response of the equivalent filter for the stream modulation, the inverse filter impulse response corresponds to the values in k^{th} column of $\mathbf{G}^{-1}$. Again the utility of stream modulation is that symbols are modulated in time and hence the processing of the linear filter can use the same filter in a recursive fashion.

The resultant output of the zero forcing equalizer is

$$\hat{D}(k) = E_b D_z(k) + N_{zf}(k). \tag{12.56}$$

Consequently the final step in the demodulation is to compute a threshold test on the output of the equalizer, i.e.,

$$\Re\left[\hat{D}(k)\right] \underset{\hat{I}(k)=1}{\overset{\hat{I}(k)=0}{\gtrless}} 0. \tag{12.57}$$

Given this threshold test, the performance of the zero forcing equalizer for BPSK is trivially given as

$$P_B(E) = \frac{1}{2}\text{erfc}\left(\frac{E_b}{\sqrt{2\sigma_{zf}^2}}\right) \tag{12.58}$$

where $2\sigma_{zf}^2 = \text{var}\left(N_{zf}(k)\right)$. Defining the power spectrum of a stationary time series as $S_N\left(e^{j2\pi f}\right) = \sum_l R_N(k)e^{-j2\pi fk}$ where $R_N(k) = E\left[N(m)N^*(m-k)\right]$ [OW97] and using traditional linear systems

theory we have

$$\text{var}\left(N_{zf}(k)\right) = \int_{-0.5}^{0.5} S_{N_{zf}}\left(e^{j2\pi f}\right) df = \int_{-0.5}^{0.5} S_{\tilde{N}}\left(e^{j2\pi f}\right) \left|W\left(e^{j2\pi f}\right)\right|^2 df. \tag{12.59}$$

Noting that $S_{\tilde{N}}\left(e^{j2\pi f}\right) = N_0 E_b G\left(e^{j2\pi f}\right)$ (see Problem 12.2) gives

$$\text{var}\left(N_{zf}(k)\right) = N_0 E_b \int_{-0.5}^{0.5} W^*\left(e^{j2\pi f}\right) df = N_0 E_b w(0). \tag{12.60}$$

Consequently

$$P_B(E) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0 w(0)}}\right) \tag{12.61}$$

Since $w(0) \geq 1$ (see Problem 12.12) the performance of a zero forcing equalizer in a frequency selective channel is always degraded compared to optimum matched filter demodulation in the frequency flat channel. Any channel with a deep spectral null in $G\left(e^{j2\pi f}\right)$ will produce $w(0) \gg 1$ and a performance significantly degraded compared to the frequency flat case.

Example 12.13: Consider the discrete time channel given in Example 12.12, where $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = T$. The zero-forcing equalizer has the form

$$\begin{aligned} W\left(e^{j2\pi f}\right) &= \frac{1}{G\left(e^{j2\pi f}\right)} \\ &= \frac{1}{0.51\left(1+0.9802\exp[j(2\pi f - \pi/3)]\right)\left(1+0.9802\exp[-j(2\pi f - \pi/3)]\right)}. \end{aligned} \tag{12.62}$$

Partial fraction expansion gives

$$W\left(e^{j2\pi f}\right) = \frac{49.98}{\left(1+0.9802\exp[j(2\pi f - \pi/3)]\right)} - \frac{49.98 \times 0.9802\exp[-j(2\pi f - \pi/3)]}{\left(1+0.9802\exp[-j(2\pi f - \pi/3)]\right)}. \tag{12.63}$$

Recall that

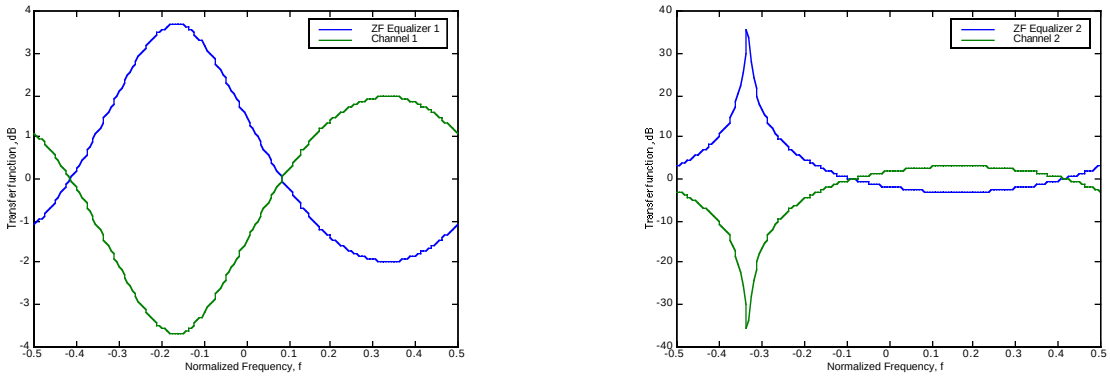
$$\frac{1}{1-x} = \sum_{l=0}^{\infty} x^l \tag{12.64}$$

so that

$$W\left(e^{j2\pi f}\right) = 49.98\sum_{l=0}^{\infty}\left(-0.9802\exp[j(2\pi f - \pi/3)]\right)^l + 49.98\sum_{l=1}^{\infty}\left(-0.9802\exp[-j(2\pi f - \pi/3)]\right)^l.$$

The loss in performance due to noise enhancement in this example is almost 17dB due to the deep spectral null in the discrete time equivalent channel. The magnitude of the zero forcing equalizer for the discrete time equivalent channels plotted in Fig. 12.17 are plotted in Fig. 12.19 (channel 1 is $\alpha_1 = 0.3\exp[j2\pi/3]$ and $\tau_1 = T$ and channel 2 is $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = T$). The channel inversion characteristic of the zero forcing equalizer is quite obvious from Fig. 12.19.

In conclusion, the concept of the ZF equalizer is simple as a linear filter is used to restore the orthogonality of the modulation and remove completely ISI. The demodulation complexity with a ZF equalizer is $O(K_b)$ and the demodulator does not need to know the SNR to be implemented. Two negative aspects to the implementation are that theoretically the ZF equalizer is an infinite impulse response filter and that if a deep null exists in the equivalent filter spectrum then a significant noise enhancement will occur by restoring the orthogonality condition.



a) Channel 1 b) Channel 2

Figure 12.19: The zero forcing equalizer transfer function.

12.3.2 Linear MMSE Equalizer

As with the OCDM demodulators introduced in Section 11.3.2, a linear MMSE equalizer can be proposed as an alternative to the ZF equalizer. The MMSE equalizer has the same structure as the zero forcing equalizer and will produce better performance. The problem can be reformulated to minimize the combined effects of residual ISI and the filtered noise. Defining the estimation error as

$$E(k) = E_b D_z(k) - \hat{D}(k) \tag{12.65}$$

then the goal is to find the filter, $W\left(e^{j2\pi f}\right)$ such that $\hat{D}(k) = E_b \sum_l w(l)\tilde{Q}(k-l)$ and that $E\left[|E(k)|^2\right]$ is minimized. A brief review of MMSE estimation and filtering is given in Appendix B. The linear MMSE filter is

$$W\left(e^{j2\pi f}\right) = E_b \frac{S_{D\tilde{Q}}\left(e^{j2\pi f}\right)}{S_{\tilde{Q}}\left(e^{j2\pi f}\right)}. \tag{12.66}$$

To get a more intuitive solution we again model the data symbols as a zero mean, unit variance white time series that is independent of the noise. This model for the data symbols and noise produces

$$R_{D\tilde{Q}}(m) = E\left[D_z(k)\tilde{Q}^*(k-m)\right] = E_b g^*(-m) = E_b g(m) \tag{12.67}$$

and

$$\begin{aligned} R_{\tilde{Q}}(m) = E\left[\tilde{Q}(k)\tilde{Q}^*(k-m)\right] &= E_b^2 \sum_l g(l)g^*(l-m) + R_{\tilde{N}}(m) \\ &= E_b^2 \sum_l g(l)g^*(l-m) + N_0 E_b g(m). \end{aligned} \tag{12.68}$$

Taking the Fourier transform and noting that $G\left(e^{j2\pi f}\right)$ is real gives

$$S_{D\tilde{Q}}\left(e^{j2\pi f}\right) = E_b G\left(e^{j2\pi f}\right) \qquad S_{\tilde{Q}}\left(e^{j2\pi f}\right) = E_b^2 G^2\left(e^{j2\pi f}\right) + N_0 E_b G\left(e^{j2\pi f}\right) \tag{12.69}$$

which results in the MMSE linear equalizer having the form

$$W\left(e^{j2\pi f}\right) = \frac{E_b}{E_b G\left(e^{j2\pi f}\right) + N_0} \tag{12.70}$$

It should be noted that as $N_0 \rightarrow 0$ then the MMSE linear equalizer converges to the zero forcing equalizer, i.e.,

$$W\left(e^{j2\pi f}\right) \rightarrow \frac{1}{G\left(e^{j2\pi f}\right)}. \tag{12.71}$$

The performance analysis with the MMSE linear equalizer is not as simple as the zero forcing equalizer. The output of the MMSE equalizer has the form

$$\hat{D}(k) = E_b \sum \tilde{g}(l) D_z(k-l) + N_{ME}(k) \tag{12.72}$$

where $\tilde{g}$ is due to the convolution of $g(l)$ with $w(l)$. Consequently the performance analysis must take into account the ISI and the Gaussian noise. Characterizing the noise is much like was done in the zero forcing algorithm, while characterizing the ISI is in general tedious. An MMSE linear equalizer provides better performance than the zero forcing equalizer but a deep spectral null in $G\left(e^{j2\pi f}\right)$ will still result in significant degradation. A disadvantage of the MMSE linear equalizer compared to the ZF equalizer is that the SNR must be known to compute the MMSE equalizer filter taps.

Example 12.14: Consider again the channel given in Example 12.12, where $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = T$. The MMSE equalizer has the form

$$\begin{aligned} W\left(e^{j2\pi f}\right) &= \frac{1}{G\left(e^{j2\pi f}\right) + \frac{N_0}{E_b}} \qquad (12.73) \\ &= \frac{1}{0.51\left(1 + 0.9802\exp\left[j(2\pi f - \pi/3)\right]\right)\left(1 + 0.9802\exp\left[-j(2\pi f - \pi/3)\right]\right) + \frac{N_0}{E_b}}. \end{aligned}$$

For example for $E_b/N_0 = 10$dB the resultant filter is given as

$$W\left(e^{j2\pi f}\right) = \frac{2.18}{\left(1 + 0.6414\exp\left[j(2\pi f - \pi/3)\right]\right)} - \frac{2.18 \times 0.6414\exp\left[-j(2\pi f - \pi/3)\right]}{\left(1 + 0.6414\exp\left[-j(2\pi f - \pi/3)\right]\right)}. \tag{12.74}$$

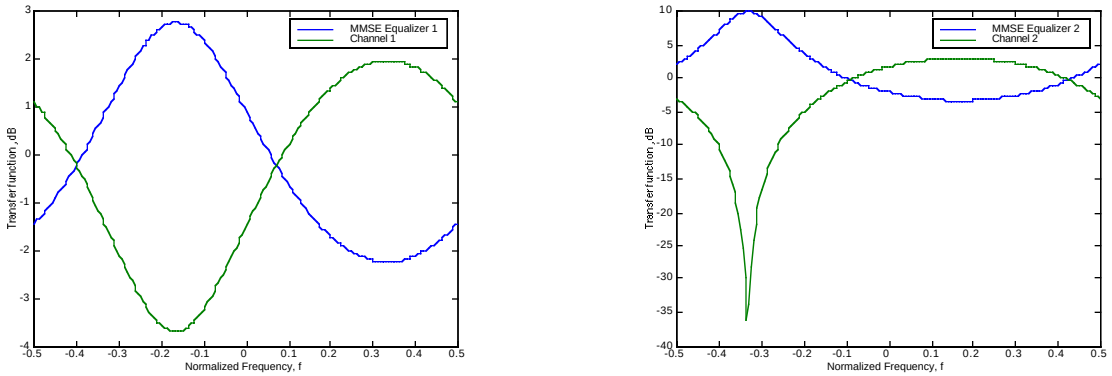
Recalling again that

$$\frac{1}{1-x} = \sum_{l=0}^{\infty} x^l \tag{12.75}$$

gives

$$W\left(e^{j2\pi f}\right) = 2.18\sum_{l=0}^{\infty}\left(-0.6414\exp\left[j(2\pi f - \pi/3)\right]\right)^l + 2.18\sum_{l=1}^{\infty}\left(-0.6414\exp\left[-j(2\pi f - \pi/3)\right]\right)^l.$$

It should be noted that since this filter is not trying to perfectly cancel all the ISI so the output of the equalizer will still have ISI that will corrupt the decision but the noise power will be less than the zero forcing equalizer output. The magnitude response of the MMSE equalizer for the discrete time equivalent channels plotted in Fig. 12.17 are plotted in Fig. 12.20 (channel 1 is $\alpha_1 = 0.3\exp[j2\pi/3]$ and $\tau_1 = T$ and channel 2 is $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = T$). The tradeoff between channel inversion and noise suppression of the MMSE equalizer is quite obvious from Fig. 12.20 as the gain at the frequencies where the channel has nulls is less pronounced than with the the zero forcing equalizer.



a) Channel 1 b) Channel 2

Figure 12.20: The MMSE equalizer transfer function. $E_b/N_0 = 10$ dB.

12.3.3 Decision Feedback Equalizer

The goal of this section is to present an analogous filtering structure for stream modulations in a frequency selective channel as the SIC for OCDM. It is desired to find a time recursive processor that consists of a linear filter with a decision directed interference removal and data reduction. The structure is denoted a decision feedback equalizer (DFE) [Aus67]. Since the goal in reduced complexity demodulation of stream modulations is to identify a time recursive demodulator, the linear filter should be targeted at detecting the earliest undemodulated bit. Detection and filtering of a time series by observations associated with the future of the observation process is known as anti-causal filtering or detection. For the discussion in this text we will concentrate on anti-causal MMSE filtering (see Appendix B) but a similar zero forcing like demodulator is possible.

The anti-causal MMSE filtering discussed here is a method to demodulate $D_z(k)$ based on the observations $\tilde{Q}(k+l),\ l \geq 0$ assuming that $D_z(i),\ i < k$ is perfectly known. The detection process in a DFE will start with $k = 1$ with no known symbols and recursively work through each $k \in \{1, \ldots, K_b\}$ with progressively more known symbols. For the sake of simplifying the notation let's denote the future observations with $\tilde{Q}(k+l) = \bar{Q}(l),\ l \geq 0$ and l indicates how far the observation is in the future. Note since $\tilde{Q}(k)$ is a stationary random process the second order moments of $\tilde{Q}$ and $\bar{Q}$ will be identical. As shown in Appendix B the anti-causal linear filter that minimizes the mean square error has the form

$$W_A\left(e^{j2\pi f}\right) = W_1\left(e^{j2\pi f}\right)\left[W_2\left(e^{j2\pi f}\right)\right]_- \tag{12.76}$$

where $W_1\left(e^{j2\pi f}\right)$ is an anti-causal whitening filter and $\left[W_2\left(e^{j2\pi f}\right)\right]_-$ is the remainder of the optimum unconstrained MMSE linear filter that is truncated to be anti-causal. Recall that using the idea of a spectral factorization gives the anti-causal whitening filter as

$$W_1\left(e^{j2\pi f}\right) = \frac{1}{\sqrt{\gamma_{\tilde{q}}} F_{\tilde{Q}}^{-}\left(e^{j2\pi f}\right)} \tag{12.77}$$

where is the anti-causal filter $F_{\tilde{Q}}^{-}\left(e^{j2\pi f}\right)$ obtained in the spectral factorization. Recall (see (12.69)) that the power spectrum that needs to be factored is given as

$$S_{\tilde{Q}}\left(e^{j2\pi f}\right) = E_b^2 G^2\left(e^{j2\pi f}\right) + N_0 E_b G\left(e^{j2\pi f}\right) = E_b G\left(e^{j2\pi f}\right)\left(E_b G\left(e^{j2\pi f}\right) + N_0\right) \tag{12.78}$$

Note that both $S_1\left(e^{j2\pi f}\right) = E_b G\left(e^{j2\pi f}\right)$ and $S_2\left(e^{j2\pi f}\right) = E_b G\left(e^{j2\pi f}\right) + N_0$ are valid power spectra[3]. Consequently $S_1\left(e^{j2\pi f}\right)$ and $S_2\left(e^{j2\pi f}\right)$ each have a spectral factorization so that

$$S_{\tilde{Q}}\left(e^{j2\pi f}\right) = \gamma_1 F_1^-\left(e^{j2\pi f}\right) F_1^+\left(e^{j2\pi f}\right) \gamma_2 F_2^-\left(e^{j2\pi f}\right) F_2^+\left(e^{j2\pi f}\right) \tag{12.79}$$

Using (12.79) it is apparent that

$$W_1\left(e^{j2\pi f}\right) = \frac{1}{\sqrt{\gamma_1} F_1^-\left(e^{j2\pi f}\right) \sqrt{\gamma_2} F_2^-\left(e^{j2\pi f}\right)}. \tag{12.80}$$

Here we see that the whitening filter is a product of two whitening filters for two different power spectra.

The second filter abstraction in the MMSE anti-causal linear filter has a similar simple form. Recall from Appendix B that

$$W_2\left(e^{j2\pi f}\right) = \frac{E_b S_{D\tilde{Q}}\left(e^{j2\pi f}\right)}{\sqrt{\gamma_{\tilde{q}}} F_{\tilde{Q}}^+\left(e^{j2\pi f}\right)} = \frac{E_b S_{D\tilde{Q}}\left(e^{j2\pi f}\right)}{\sqrt{\gamma_1} F_1^+\left(e^{j2\pi f}\right) \sqrt{\gamma_2} F_2^+\left(e^{j2\pi f}\right)} \tag{12.81}$$

Using (12.69), (12.81) can be re-expressed as

$$W_2\left(e^{j2\pi f}\right) = \frac{E_b^2 G\left(e^{j2\pi f}\right)}{\sqrt{\gamma_{\tilde{q}}} F_{\tilde{Q}}^+\left(e^{j2\pi f}\right)} = \frac{E_b \gamma_1 F_1^-\left(e^{j2\pi f}\right) F_1^+\left(e^{j2\pi f}\right)}{\sqrt{\gamma_1} F_1^+\left(e^{j2\pi f}\right) \sqrt{\gamma_2} F_2^+\left(e^{j2\pi f}\right)} = \frac{\sqrt{\gamma_1} F_1^-\left(e^{j2\pi f}\right)}{\sqrt{\gamma_2} F_2^+\left(e^{j2\pi f}\right)} \tag{12.82}$$

Since $\left(F_2^+\left(e^{j2\pi f}\right)\right)^{-1}$ is also a causal monic filter, the truncated anti-causal filter reduces to

$$\left[W_2\left(e^{j2\pi f}\right)\right]_- = \frac{E_b \sqrt{\gamma_1} F_1^-\left(e^{j2\pi f}\right)}{\sqrt{\gamma_2}}. \tag{12.83}$$

Consequently after a significant manipulation the optimum linear anti-causal filter has a fairly simple solution. Combining (12.80) and (12.83) in (12.76) gives

$$W_A\left(e^{j2\pi f}\right) = \frac{E_b}{\gamma_2 F_2^-\left(e^{j2\pi f}\right)}. \tag{12.84}$$

This simple solution produces a linear filter that can be identified based on spectral factorization of $S_2\left(e^{j2\pi f}\right) = E_b G\left(e^{j2\pi f}\right) + N_0$. The output of this linear filter is input into a decision device to produce a decision, i.e.,

$$\Re\left[\hat{D}(k)\right] \underset{\hat{I}(k)=1}{\overset{\hat{I}(k)=0}{\gtrless}} 0. \tag{12.85}$$

After obtaining this decision, the value of $\tilde{Q}(k)$ can be discarded and this decision can be remodulated and subtracted from all future observations

$$\bar{Q}^{(1)}(l) = \bar{Q}(l+1) - a\left(\hat{I}(k)\right) g(l+1) \qquad 0 \le N_u - 1. \tag{12.86}$$

$$\bar{Q}^{(1)}(l) = \bar{Q}(l+1) \qquad N_u \le l \le K_b - k \tag{12.87}$$

The important thing to note is that $\bar{Q}^{(1)}(l)$ and $\bar{Q}(l)$ have the exact same mathematical models hence a recursion for a DFE has been achieved.

[3]Recall that $E_b g(l) = V_{\tilde{u}}(lT)$.

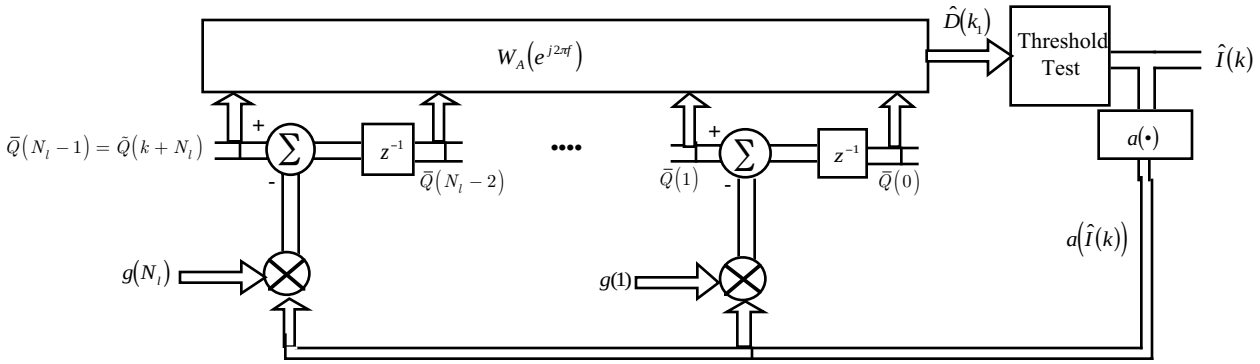


Figure 12.21: A block diagram of the DFE where interference is removed prior to the linear filter. The filter is initialized with $\bar{Q}(l-1)=\tilde{Q}(l)$ for $l \in \{1,\ldots,N_l\}$

Example 12.15: Consider again the channel given in Example 12.12, where $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = T$. The signal plus noise spectrum, $S_2\left(e^{j2\pi f}\right)$, is given as

$$\begin{aligned} S_2\left(e^{j2\pi f}\right) &= E_b G\left(e^{j2\pi f}\right) + N_0 \qquad (12.88)\\ &= \gamma_2 F_2^-\left(e^{j2\pi f}\right) F_2^+\left(e^{j2\pi f}\right) = E_b\gamma_m\left(1+\alpha_m^*\exp[j2\pi f]\right)\left(1+\alpha_m\exp[-j2\pi f]\right). \end{aligned}$$

It should be noted here that this channel only has one symbol worth of ISI. The feedforward anti-causal linear filter in this case in this channel is given as

$$W_A\left(e^{j2\pi f}\right) = \frac{1}{\gamma_m\left(1+\alpha_m^*\exp[j2\pi f]\right)}. \qquad (12.89)$$

For $E_b/N_0 = 10$dB this gives

$$W_A\left(e^{j2\pi f}\right) = 1.28\sum_{i=0}^{\infty}\left(-0.6414\exp[j(2\pi f-\pi/3)]\right)^i. \qquad (12.90)$$

The magnitude responses of the anti-causal MMSE equalizer for the discrete time equivalent channels plotted in Fig. 12.17 are plotted in Fig. 12.22 (channel 1 is $\alpha_1 = 0.3\exp[j2\pi/3]$ and $\tau_1 = T$ and channel 2 is $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = T$). This impulse response implies the estimator has the form

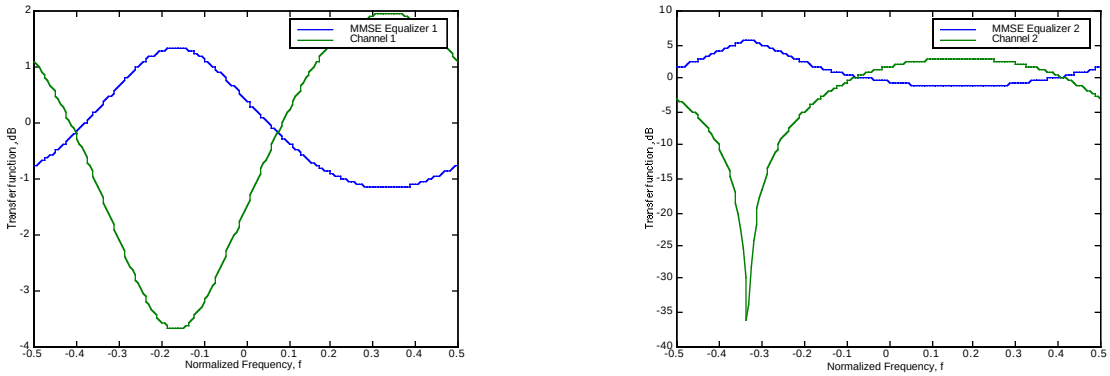
$$\hat{D}(k) = 1.28\bar{Q}(0) - 0.821\exp[-j\pi/3]\,\bar{Q}(1) + 0.5266\exp[-j2\pi/3]\,\bar{Q}(2) + \ldots \qquad (12.91)$$

For the first symbol this filter output is characterized by

$$\hat{D}(1) = 1.28\tilde{Q}(1) - 0.821\exp[-j\pi/3]\,\tilde{Q}(2) + 0.5266\exp[-j2\pi/3]\,\tilde{Q}(3) + \ldots$$

The decision are produced by a threshold test on $\Re\left[\hat{D}(k)\right]$. Because of only one symbol of ISI, the interference subtraction takes the form

$$\begin{aligned} \bar{Q}^{(1)}(0) &= \bar{Q}(1) - a\left(\hat{I}(k)\right)g(1) = \tilde{Q}(k+1) - a\left(\hat{I}(k)\right)g(1) \qquad (12.92)\\ \bar{Q}^{(1)}(l) &= \bar{Q}(l+1) = \tilde{Q}(k+l+1) \qquad l \geq 1. \qquad (12.93) \end{aligned}$$



a) Channel 1 b) Channel 2

Figure 12.22: The anti-causal MMSE equalizer transfer function.

While the discussion here was formulated assuming that all future observations can be used in the MMSE linear filter, usually in practice a finite length filter is implemented. Denoting this filter length as N_l a block diagram of the DFE is shown in Fig. 12.21. Note that often $N_l > N_u$ so that the interference cancellation that has to be implemented does not need to extend back to the beginning of the anti-causal filter. The DFE often provides significant performance improvement compared to a linear equalizer and that is because the decision feedback has removed all the causal ISI from the observations. Again it should be noted that this development assumed that all previous decisions are correct. This is not true in practice and the bit error probability performance of a DFE can be seriously degraded due to erroneous decisions being fedback.

The complexity of the DFE is $O\left((N_l + N_u)K_b\right)$. The anti-causal filter has a complexity of $O\left(N_l\right)$. In practice the feedforward filter is typically chosen such that $N_l < 10N_u$ where the exact value of N_u is chosen based on the kinds of channels that will be experienced in the application. The decision feedback (interference removal) has a complexity of $O\left(N_u\right)$. Each of these operations has to be done for each decoded bit resulting in the complexity being $O\left((N_l + N_u)K_b\right)$. Again if $K_b \gg N_u$ then the complexity of the DFE is $O\left(K_b\right)$ as desired.

A traditional method of implementing the DFE is not as shown in Fig. 12.21. This traditional method [Pro89] of implementing the DFE subtracts the ISI due to the demodulated bits at the output of the linear filter (as opposed to before the linear filter in Fig. 12.21). The block diagram of the traditional implementation of the DFE is shown in Fig. 12.23. The feedback filter should be such that the causal part of

$$E_b G\left(e^{j2\pi f}\right) W_A\left(e^{j2\pi f}\right) = \frac{\gamma_1 F_1^-\left(e^{j2\pi f}\right) F_1^+\left(e^{j2\pi f}\right)}{\gamma_2 F_2^-\left(e^{j2\pi f}\right)} \tag{12.94}$$

is cancelled. Using $S_2\left(e^{j2\pi f}\right) = S_1\left(e^{j2\pi f}\right) + N_0$ gives

$$\begin{aligned} E_b G\left(e^{j2\pi f}\right) W_A\left(e^{j2\pi f}\right) &= \frac{\gamma_2 F_2^-\left(e^{j2\pi f}\right) F_2^+\left(e^{j2\pi f}\right)}{\gamma_2 F_2^-\left(e^{j2\pi f}\right)} - \frac{N_0}{\gamma_2 F_2^-\left(e^{j2\pi f}\right)} \\ &= F_2^+\left(e^{j2\pi f}\right) - \frac{N_0}{\gamma_2 F_2^-\left(e^{j2\pi f}\right)}. \end{aligned} \tag{12.95}$$

It is clear in examining (12.95) that all the effects of the past bits can be removed by setting

$$B\left(e^{j2\pi f}\right) = F_2^+\left(e^{j2\pi f}\right) - 1. \tag{12.96}$$

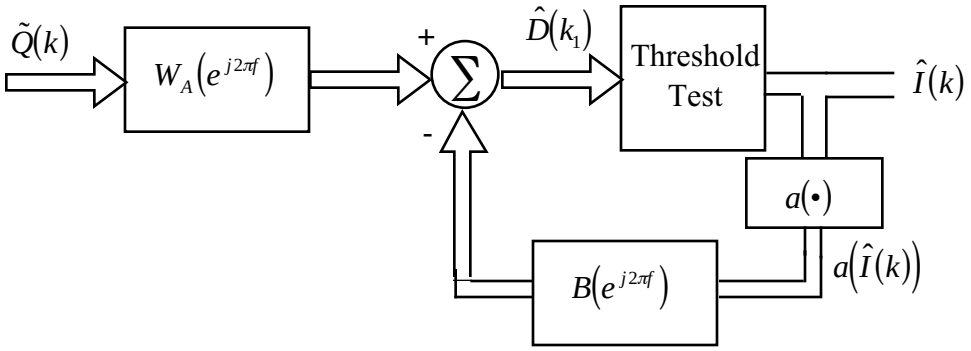


Figure 12.23: A block diagram of the DFE where interference is removed after the linear filter.

	Modulation Type		
	OCDM	OFDM	Stream
Optimum Demodulation	$O\left(2^{K_b}\right)$	$O\left(2^{K_b}\right)$	$O\left(K_b N_u 2^{N_u+1}\right)$
Suboptimal Demodulation	$O\left(K_b^2\right)$	$O\left(K_b\right)$	$O\left(K_b N_u\right)$

Table 12.1: A complexity comparison of the various orthogonal modulations on frequency selective fading channels.

The two implementations given in Fig. 12.21 and Fig. 12.23 have the same complexity and give the same outputs. This text has highlighted the form of the DFE given in Fig. 12.21 as it is more consistent with the SIC that was introduced for OCDM in a frequency selective channel. In addition the form of the equalizer in Fig. 12.21 has some advantages when the impulse response of the channel is sparse and this sparseness is used to achieve complexity reduction [FGF99].

12.4 Conclusion

This chapter examined stream modulations in frequency selective channels. It is shown that the optimum demodulator, as proposed by Ungerboeck, can be reformulated in a recursive processor that has a complexity of $O\left(K_b N_u 2^{N_u+1}\right)$. This complexity is so still too large for implementation so the suboptimum demodulators for OCDM were re-interpreted as time recursive processors and filters for stream modulations. These suboptimum demodulators have a complexity $O\left(K_b N_u\right)$. The complexity options available with the variety of orthogonal modulations is summarized in Table 12.1.

12.5 Homework Problems

Problem 12.1. For a binary stream modulation on a frequency selective channel assume that the effective pulse shape is characterized with

$$V_{\tilde{u}}(nT) = \begin{cases} E_b & n = 0 \\ 0.5E_b & n = \pm 1 \\ 0 & \text{elsewhere.} \end{cases} \tag{12.97}$$

Assume the corrupting noise is an AWGN with a one-sided spectral density of N_0. A MLWD is used to demodulate the data.

$I(l)=n_l$	$\tilde{T}_{n_l}(l)$	$\hat{\vec{I}}_{n_l}(l)$
0	5	$\vec{I}_1$
1	2	$\vec{I}_2$

Table 12.2: The Ungerboeck algorithm metrics at time l.

a) Give the MLWD in it's most general form for stream modulation.

b) The Ungerboeck implementation of the MLWD is a function of a parameter N_u. What is N_u in this case?

c) At a particular point in time, $2 \leq l \leq K_b - 1$, the Ungerboeck equalizer for this channel has a demodulation state characterized as in Table 12.2. Assume $E_b = 1$ and $\tilde{Q}(l+1) = 0.3$ find $\tilde{T}_{n_{l+1}}(l+1)$ and $\hat{\vec{I}}_{n_{l+1}}(l+1)$

Problem 12.2. Consider stream modulation on a frequency selective channel. Prove that since effective pulse shaped matched filter has the form $\tilde{Q}(k) = \tilde{S}(k) + \tilde{N}(k)$ where $\tilde{S}(k)$ is due to the filtering of $R_z(t)$ and $\tilde{N}(k)$ is due to the filtering of $W(t)$, then

$$R_{\tilde{N}}(m) = E\left[\tilde{N}(k)\tilde{N}^*(k-m)\right] = N_0 V_{\tilde{u}}(mT) = N_0 g(m) \tag{12.98}$$

and that

$$S_{\tilde{N}}\left(e^{j2\pi f}\right) = N_0 G\left(e^{j2\pi f}\right) \tag{12.99}$$

Problem 12.3. Consider BPSK stream modulation on a frequency selective channel with the following channel

$$\begin{aligned} V_{\tilde{u}}(nT) &= 1 \qquad n = 0 \\ &= 0.4 \qquad n \pm 1 \\ &= 0. \qquad \text{elsewhere.} \end{aligned} \tag{12.100}$$

with BPSK modulation.

a) Consider $K_b = 4$ and find $\mathbf{G}$.

b) When $\vec{\tilde{q}}_I = [\tilde{q}_I(1) \cdots \tilde{q}_I(4)] = [0.9 \ -0.3 \ 1.1 \ 1.2]$ find $\hat{I}(k),\ k \in \{1, \ldots, 4\}$ using Ungerboeck's MLWD. Clearly delineate each step in the recursion and the resulting partial ML metrics and surviving sequences computed and saved. Note since the modulation is only on the I-channel the quadrature component of the matched filter output does not effect the decision.

Problem 12.4. Consider BPSK stream modulation on a frequency selective channel with the following channel

$$\begin{aligned} V_{\tilde{u}}(nT) &= 1 \qquad n = 0 \\ &= 0.4 \qquad n \pm 1 \\ &= 0. \qquad \text{elsewhere.} \end{aligned} \tag{12.101}$$

with BPSK modulation.

a) Assume K_b is large and compute the zero forcing equalizer solution, $W\left(e^{j2\pi f}\right)$ and $w(k)$.

b) Assume K_b is large and compute the resulting bit error probability when using the zero forcing equalizer.

c) Consider $K_b = 4$ and find $\mathbf{G}$.

d) When $\vec{\tilde{q}}_I = [\tilde{q}_I(1) \cdots \tilde{q}_I(4)] = [0.9 \ -0.3 \ 1.1 \ 1.2]$ find $\hat{D}_z(k),\ k \in \{1, \ldots, 4\}$ using a zero forcing or decorrelating linear demodulator. Note since the modulation is only on the I-channel the quadrature component of the matched filter output does not effect the decision.

e) When $K_b = 21$ show that the middle row of $\mathbf{G}^{-1}$ approximates $w(k)$.

Problem 12.5. Consider BPSK stream modulation on a frequency selective channel with the following channel

$$\begin{aligned} V_{\tilde{u}}(nT) &= 1 \qquad n = 0 \\ &= 0.4 \qquad n \pm 1 \\ &= 0. \qquad \text{elsewhere.} \end{aligned} \tag{12.102}$$

with BPSK modulation. Assume that $E_b/N_0 = 10$dB.

a) Assume K_b is large and compute the MMSE equalizer solution, $W\left(e^{j2\pi f}\right)$ and $w(k)$.

b) Consider $K_b = 4$ and find $\mathbf{G}$ and the MMSE linear detector.

c) When $\vec{\tilde{q}}_I = [\tilde{q}_I(1) \cdots \tilde{q}_I(4)] = [0.9 \ -0.3 \ 1.1 \ 1.2]$ find $\hat{D}_z(k),\ k \in \{1, \ldots, 4\}$ using a MMSE linear demodulator. Note since the modulation is only on the I-channel the quadrature component of the matched filter output does not effect the decision.

d) When $K_b = 21$ show that the middle row of the MMSE equalizer approximates $w(k)$.

Problem 12.6. Consider BPSK stream modulation on a frequency selective channel with the following channel

$$\begin{aligned} V_{\tilde{u}}(nT) &= 1 \qquad n = 0 \\ &= 0.4 \qquad n \pm 1 \\ &= 0. \qquad \text{elsewhere.} \end{aligned} \tag{12.103}$$

with BPSK modulation. Assume that $E_b/N_0 = 10$dB.

a) Assume K_b is large and compute the anti-causal MMSE equalizer solution, $W\left(e^{j2\pi f}\right)$ and $w(k)$.

b) Consider $K_b = 4$ and find $\mathbf{G}$ and the MMSE linear detector.

c) When $\vec{\tilde{q}}_I = [\tilde{q}_I(1) \cdots \tilde{q}_I(4)] = [0.9 \ -0.3 \ 1.1 \ 1.2]$ find $\hat{D}_z(k),\ k \in \{1, \ldots, 4\}$ using a MMSE linear demodulator. Note since the modulation is only on the I-channel the quadrature component of the matched filter output does not effect the decision.

d) When $K_b = 21$ show that the first row of the MMSE equalizer approximates $w(k)$.

Problem 12.7. Recall that

$$w(0) = \int_{-0.5}^{0.5} \frac{1}{G\left(e^{j2\pi f}\right)} df \tag{12.104}$$

and

$$\int_{-0.5}^{0.5} G\left(e^{j2\pi f}\right) df = 1. \tag{12.105}$$

Use Schwarz's inequality to prove that $w(0) \geq 1$. This result implies that the ZF equalizer in a frequency selective channel always has worse performance than what would be achieved in a frequency flat channel.

Problem 12.8. If $\tilde{Q}(k) = \tilde{S}(k) + \tilde{N}(k)$ where $\tilde{S}(k)$ is due to the filtering of the signal, $R_z(t)$ by a filter matched to $\tilde{u}(t)$ and $\tilde{N}(k)$ is due to the filtering of the input noise, $W_z(t)$ by a filter matched to $\tilde{u}(t)$ then show

$$R_{\tilde{N}}(m) = N_0 V_{\tilde{u}}(-mT) = N_0 E_b g(-m). \tag{12.106}$$

Problem 12.9. Recall the channel given in Example 12.1 with $\alpha_1 = 0.3\exp[j2\pi/3]$ and $\tau_1 = 0.25T$. Assume a BPSK stream modulation is being used with

$$u(t) = \begin{cases} \sqrt{\dfrac{1}{T_u}} & 0 \leq t \leq T_u \\ 0 & \text{elsewhere} \end{cases} \tag{12.107}$$

where $T_u = T_p/4 = T$. Assume that $E_b/N_0 = 10$dB.

a) Compute the zero forcing linear equalizer impulse response.

b) Compute the MMSE linear equalizer impulse response.

c) Compute the anti-causal linear MMSE equalizer impulse response.

Problem 12.10. For a binary stream modulation on a frequency selective channel, assume that the effective pulse shape is characterized with

$$V_{\tilde{u}}(nT) = \begin{cases} E_b & n = 0 \\ 0.8E_b & n = \pm 1 \\ 0.2E_b & n = \pm 2 \\ 0 & \text{elsewhere.} \end{cases} \tag{12.108}$$

Assume the corrupting noise is an AWGN with a one-sided spectral density of N_0.

a) If $K_b = 5$ give the form for the matrix $\mathbf{G}$.

b) Give the form for $G\left(e^{j2\pi f}\right)$.

c) Assume a MLWD is used to demodulate the data. At a particular point in time, $2 \leq l \leq K_b - 1$, the Ungerboeck equalizer for this channel has a demodulation state characterized as in Table 12.3. Assume $E_b = 1$ and $\tilde{Q}(l+1) = 0.3$ find $\tilde{T}_{n_{l+1}}(l+1)$ and $\hat{\vec{I}}_{n_{l+1}}(l+1)$

$I(l) = n_l$	$\tilde{T}_{n_l}(l)$	$\hat{\vec{I}}_{n_l}(l)$
0	5	$\vec{I}_1$
1	2	$\vec{I}_2$
2	1	$\vec{I}_3$
3	3	$\vec{I}_4$

Table 12.3: The Ungerboeck algorithm metrics at time l.

Chapter 13

Orthogonal Modulations with Memory

13.1 Canonical Problems

Up to this point in this text we have introduced two general methods to communicate K_b bits on both frequency flat channels and frequency selective channels.

1. General M-ary modulations
    - The advantage of a general M-ary modulation is that it can achieve very good performance and arbitrary spectral efficiency.
    - The disadvantage of a general M-ary modulation is that without more structure the optimal demodulator has complexity $O(2^{K_b})$.
2. Orthogonal memoryless modulations (including stream modulations, OFDM, OCDM)
    - The advantage of orthogonal modulation is that the optimum demodulator has complexity $O(K_b)$ and a desired spectral efficiency can be achieved with a proper design of the modulation signals.
    - The disadvantage is that the performance is limited to that achievable with a single symbol transmission.
    - Frequency selective channels cause orthogonality to be lost. A variety of algorithms were introduced to address optimal and sub-optimal demodulation of memoryless orthogonal modulations in frequency selective channels. Performance on frequency selective channels is always lower bounded by the performance on frequency flat channels.

The goals for the remainder of the text will be to explore the remaining areas of tradeoff in performance, complexity, and spectral efficiency, i.e.,

- Improving over the performance or spectral characteristics of orthogonal modulations with a goal of maintaining a demodulation complexity that is $O(K_b)$.

Shannon has given us an upper bound on the performance that can be achieved for a given spectral efficiency (see Fig. 6.4) and there are a variety of ways to approach that performance with a demodulation complexity that is $O(K_b)$. The concept of orthogonal modulations with memory are what have allowed communication theory to approach the bounds provided by Claude Shannon in his ground breaking work on information theory [Sha48]. This chapter will be a brief introduction into the topic and an examination of the implications of adding memory to the performance and the spectral characteristics.

13.2 Orthogonal Modulations with Memory

This section addresses a class of communications problems where a linear orthogonal modulation is implemented with the modulation symbols having memory. The memory is normally included in the modulation to either improve the performance or change the spectral characteristics of memoryless linear orthogonal modulation. This type of modulation is referred to as modulation with memory (MWM) and incorporates most error control coding schemes [Wic95, LC04, BDMS91]. Our goal in this text is not to explore how to design these MWM but to understand the communication theory behind the performance, spectral efficiency, and the demodulation complexity. Design of MWM is often addressed in a course on error control coding. Suffice it to say here that a MWM adds memory to the modulation process with a goal of either changing the spectral characteristics of the transmitted signal or improving the resulting squared Euclidean distance spectrum.

This chapter first considers the special case of an orthogonal modulation with memory (OMWM) where one bit is sent with each symbol. For simplicity of discussion this section will again assume that the bits to be transmitted are equally likely and independent. The generalizations for correlated bits is possible but the added notational complexity is not worth the gain in generality. The resulting data symbols can then be transmitted using a linear orthogonal modulation. The data modulation symbols, $\tilde{D}_z(l)$, are due to the stream of information bits, $I(l), l = 1, K_b$. The tilde notation will be used to differentiate between modulations that have memory (tilde) and memoryless modulations (not tilde). For continuity with the previous discussions on orthogonal modulations this initial discussion considers exclusively modulations where the transmission bit rate is $R = 1$ bit per symbol. To keep a consistent normalization the mapping, $\tilde{D}_z(l) = a\left(J(l)\right)$, is selected such that $E\left[\left|\tilde{D}_z(l)\right|^2\right] = 1$. A more general formulation for linear stream modulations is considered later in this chapter.

A linear orthogonal modulation with memory as considered in this section consists of a finite state machine operating at the symbol rate where the output of the finite state machine is used as an input to an M_s-ary linear modulator. Fig. 13.1 shows the block diagram for a linear orthogonal modulation with memory. At each symbol time a new bit, $I(l)$ is input into the system and this produces a new constellation label, $J(l)$, and a new modulation state, $\sigma(l+1)$. Let N_s denotes the number of states in the modulation and the nonlinear equations governing the finite state machine are given as

$$\sigma(l+1) = g_1\left(\sigma(l), I(l)\right) \tag{13.1}$$

$$J(l) = g_2\left(\sigma(l), I(l)\right). \tag{13.2}$$

For notational purposes denote the set of all possible state values as Ω_σ. In general it is often desirable to have ν_c extra symbols transmitted to return the modulation to a common final state at the end of the transmission frame. The total length of the frame for the orthogonal modulation is denoted N_f, hence, in this case when $R = 1$ bit/symbol $N_f = K_b + \nu_c$, where ν_c is a code dependent constant. This return to a final common state is most often known as termination in the literature of MWM [Wic95, LC04, BDMS91]. Note that the effective rate is $R_{eff} = \frac{K_b}{K_b+\nu_c} < 1$ but if a large number of bits are transmitted then the loss in efficiency by including the termination becomes small.

A couple comments about OMWM are appropriate at this point. An OMWM can only provide a performance improvement or a change in spectral characteristics compared to a memoryless modulation if there is some redundancy to be exploited. An orthogonal modulation with memory has 2^{K_b} possible bit sequences as an input. The modulation symbols being mapped onto the orthogonal modulation have $M_s^{K_b+\nu_c}$ possible realizations. Hence if there is to be redundancy that the OMWM can exploit for large K_b then $M_s > 2$. The improved performance or spectral efficiency is achieved by picking the best 2^{K_b} possible transmitted symbol sequences out of the $M_s^{K_b+\nu_c}$ total possible transmitted symbol sequences. Secondly, the more states a OMWM has the more memory it contains. In general, more memory allows a designer to make better choices in the transmitted symbol sequences but at a cost of a

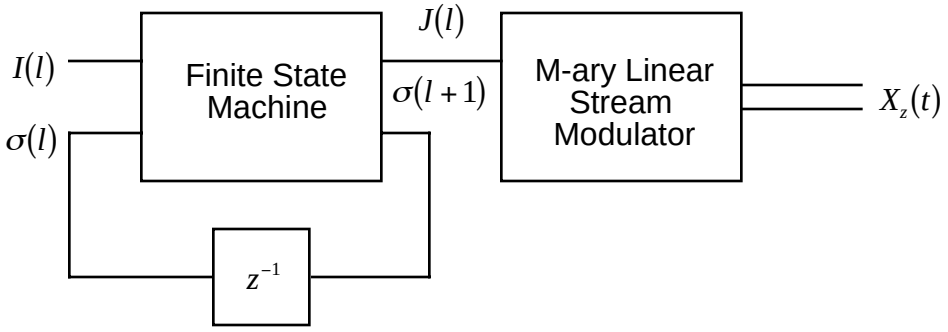


Figure 13.1: The block diagram for a modulation with memory with a transmission rate of one bit per symbol.

higher complexity demodulation. The optimum demodulation structure will be addressed next chapter. Finally, an OMWM is often represented with a directed graph [Wes01]. The vertices of the directed graph represent the state of the modulation at a particular time and the edges of the directed graph represent the allowable transitions between states. In the communications literature the directed graph is often referred to as a trellis. Because of this trellis representation OMWM are often referred to as trellis codes. This trellis representation of an OMWM is explored in the example considered in the next section.

13.2.1 MLWD for Orthogonal Modulations with Memory

Recall OMWM can be characterized with a finite state machine defined with

$$\begin{aligned} \sigma(l+1) &= g_1\left(\sigma(l), I(l)\right) && (13.3)\\ J(l) &= g_2\left(\sigma(l), I(l)\right). && (13.4) \end{aligned}$$

The equivalent modulation symbol is $\tilde{D}_z(l) = a\left(J(l)\right)$ with $a(\bullet)$ being the constellation mapping. Here we examine only the optimum word demodulation (MLWD) but the generalization to other optimum demodulators are possible, (e.g., MLBD). Other demodulators be explored in the sequel. The optimum word demodulator is still the MLWD for an orthogonal modulation except now the memory of the modulation must be accounted for computing the maximum likelihood metric. Due to the orthogonal modulation the MLWD has

$$\begin{aligned} \hat{\vec{I}} &= \arg \max_{i=0,M-1} T_i \\ &= \arg \max_{i=0,M-1} \sqrt{E_b} \sum_{k=1}^{N_f} \Re\left[\tilde{d}_i^*(k) Q(k)\right] - \frac{E_b}{2} \sum_{k=1}^{N_f} \left|\tilde{d}_i(k)\right|^2 \\ &= \arg \min_{i=0,M-1} \sum_{k=1}^{N_f} \left|Q(k) - \sqrt{E_b}\tilde{d}_i(k)\right|^2 \end{aligned} \quad (13.5)$$

where $Q(k)$ is the matched filter ouptut sample for an OMWM. Consequently the MLWD can be thought of as the possible transmitted modulation symbols (constrained by the modulation with memory) that is closest in Euclidean distance to the matched filter outputs that are observed ($Q(k)$) over the entire frame ($k = 1, N_f$). This minimum squared Euclidean distance decoder provides a nice analogy to the demodulation of memoryless modulations.

$\sigma(l+1) = g_1(I(l), \sigma(l))$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	1	2
2	3	4
3	1	2
4	3	4

$J(l) = g_2(I(l), \sigma(l))$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	0	2
2	3	1
3	2	0
4	1	3

$\tilde{D}_z(l) = a(J(l))$

$J(l)$	$\tilde{D}_z(l)$
0	$-3/\sqrt{5}$
1	$-1/\sqrt{5}$
2	$1/\sqrt{5}$
3	$3/\sqrt{5}$

Table 13.1: The finite state machine description of an example trellis code.

As with a general M-ary modulation the performance of the MLWD for an OMWM is determined by the squared Euclidean distance spectrum. Due to the use of an orthogonal modulation the squared Euclidean distance between any two words is given as

$$\int_{-\infty}^{\infty} |x_i(t) - x_j(t)|^2 \, dt = E_b \sum_{l=1}^{N_f} \left| \tilde{d}_i(l) - \tilde{d}_j(l) \right|^2. \tag{13.6}$$

Consequently the important characteristics in terms of performance of an OMWM is the squared Euclidean distance between the code sequences.

13.2.2 An Example OMWM Providing Better Performance

Consider a OMWM consisting of a four state trellis code (N_s=4) using 4PAM modulation (M_s=4) as proposed by Ho, Cavers and Varaldi [HCV93]. It should be noted here that this OMWM is a simple construction using a general combined modulation and coding technique proposed by Ungerboeck in [Ung82]. This example OMWM illustrates some important properties that will become apparent in the sequel. The modulation updates are given in Table 14.4. Here we have $\Omega_\sigma = \{1, 2, 3, 4\}$. Note that the modulation is normalized to have an average energy of unity. The trellis for this modulation is shown in Fig. 14.13. As mentioned before, the vertices of the trellis diagram represents the states at each time interval (hence for this code there are $N_s = 4$ vertices at each point in time) and the edges are the possible transitions (note for instance $\sigma(l) = 1$ can only transition to $\sigma(l+1) = 1$ and $\sigma(l+1) = 2$). There are a total of $2N_s = 8$ edges in this example. The edge labels indicate the input that causes the transition and the modulation symbol that is output for that input and state. For instance, when $\sigma(l) = 1$ and $I(l) = 0$ the output symbol is $\tilde{D}_z(l) = -3/\sqrt{5}$. This particular code requires at most two transitions to allow the modulation to transition from any state to any other state so this code has $\nu_c = 2$. For instance, if $\sigma(l) = 2$ then to get to $\sigma(l+2) = 1$ one would have to transition to $\sigma(l+1) = 3$ first. Alternatively if $\sigma(l) = 1$ then we could immediately transition to $\sigma(l+1) = 1$.

N_f trellis sections can be combined together to get a complete description of an OMWM. For instance for $K_b = 4$, Fig. 14.2 shows a trellis description for a modulation that has $\sigma(1) = 1$ and returns to $\sigma(7) = 1$. One can verify that there are 2^{K_b} paths through this trellis. To understand the performance of this particular OMWM all of the $16 \times 15/2 = 120$ terms in the squared Euclidean distance spectrum should be computed. Consider two of the paths in the trellis corresponding the words $\vec{I} = 0 = [0\ 0\ 0\ 0]$ and $\vec{I} = 1 = [1\ 0\ 0\ 0]$. These two words will produce modulation sequences of $\vec{\tilde{d}}_0 = [-3\sqrt{5}\ \ -3\sqrt{5}\ \ -3\sqrt{5}\ \ -3\sqrt{5}\ \ -3\sqrt{5}\ \ -3\sqrt{5}]$ and $\vec{\tilde{d}}_1 = [1\sqrt{5}\ 3\sqrt{5}\ 1\sqrt{5}\ \ -3\sqrt{5}\ \ -3\sqrt{5}\ \ -3\sqrt{5}]$. This results in a squared Euclidean distance of

$$\Delta_E(1, 0) = E_u(4\sqrt{5})^2 + E_u(6\sqrt{5})^2 + E_u(4\sqrt{5})^2 = 68E_u/5. \tag{13.7}$$

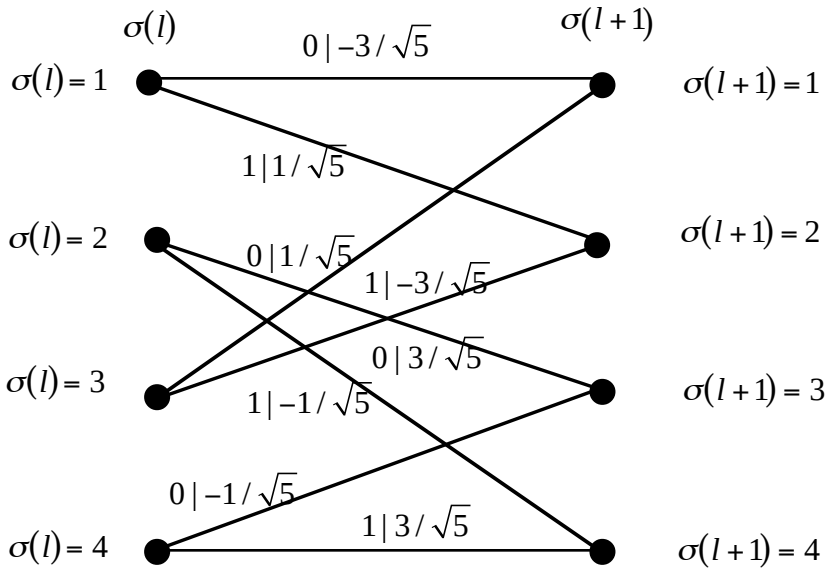


Figure 13.2: The trellis diagram of the Ho, Cavers, and Varaldi trellis code.

Hence at least these two paths produce a squared Euclidean distance better than the best memoryless modulation of the same rate, i.e., BPSK has $\Delta_E(i,j) = 4E_u$. The entire distance spectrum of this modulation is explored in the homework problems (see Problem 13.1) and the union bound to the probability of word is plotted in Fig. 13.4.

13.2.3 Discussion

This section introduced one example modulation to transmit one bit of information per orthogonal dimension. As a final point it is worth comparing the spectral efficiency performance of this example modulations with the upper bounds provided by information theory (see Section 6.3). As before we will denote reliable communication as being an error rate of 10^{-5}. This OMWM achieves reliable communications at E_b/N_0=7.8dB. The performance of this modulation versus the Shannon capacity is plotted in Fig. 13.5 and compared to the best orthogonal memoryless modulation with the same rate ($R = 1$). It should be noted that by considering a very simple modulation with memory (4 states) we have moved over 2dB closer to the achievable reliability predicted by Shannon. Using more states and a proper design of the modulation can improve the performance further. In fact the performance can be made arbitrarily close to the Shannon bound by increasing the complexity of the OMWM. The remaining items to be explored are the spectral characteristics of OMWM, a generalization of OMWM to an arbitrary rate, and the demodulation algorithm form and complexity. Note if we can find a demodulator that has a complexity that is $O(K_b)$ then we have succeeded in obtaining the desired characteristics (i.e., improved performance compared to memoryless orthogonal modulation while maintaining a complexity $O(K_b)$). The next chapter will examine the MLWD for an OMWM and show how to obtain the desired complexity. The remainder of this chapter will show that there is an equivalence between a memoryless stream modulation on a frequency selective channel and an OMWM, explore the spectral characteristics of stream modulation and generalize OMWM to an arbitrary rate.

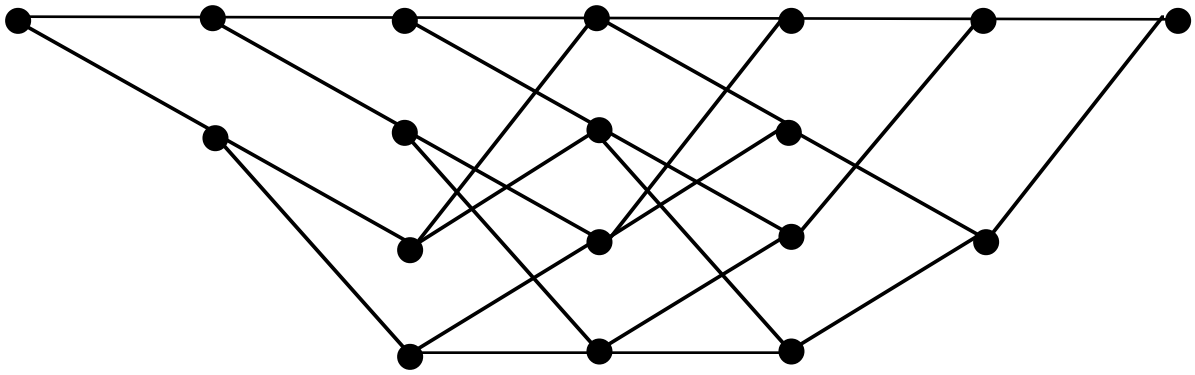


Figure 13.3: The trellis diagram of the Ho, Cavers Varaldi trellis code. $K_b = 4$.

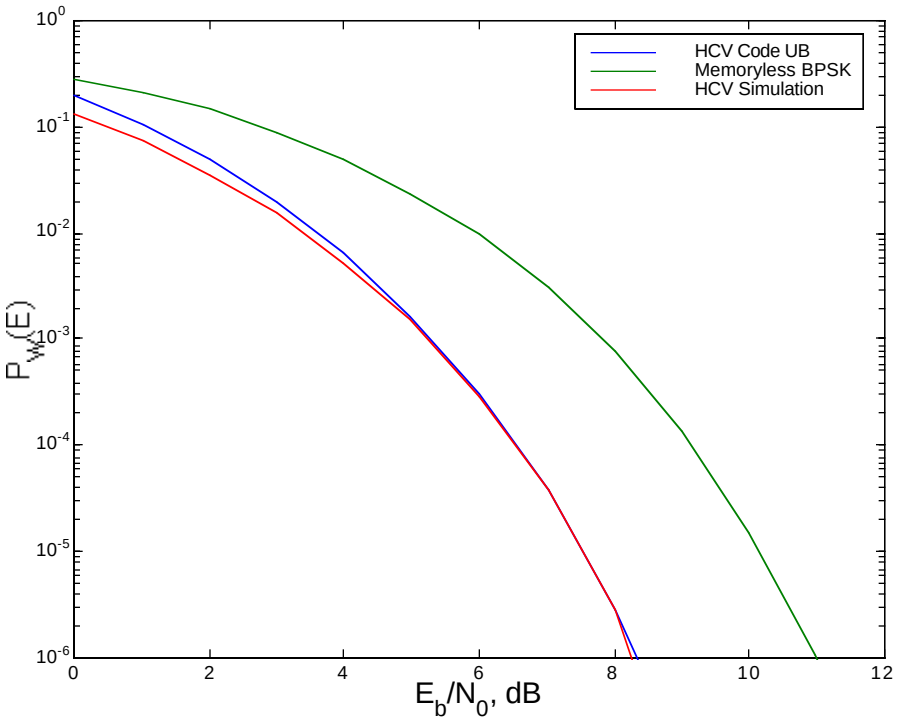


Figure 13.4: The union bound of the example OMWM and frame error rate of $K_b = 4$ memoryless orthogonally modulated bits using BPSK.

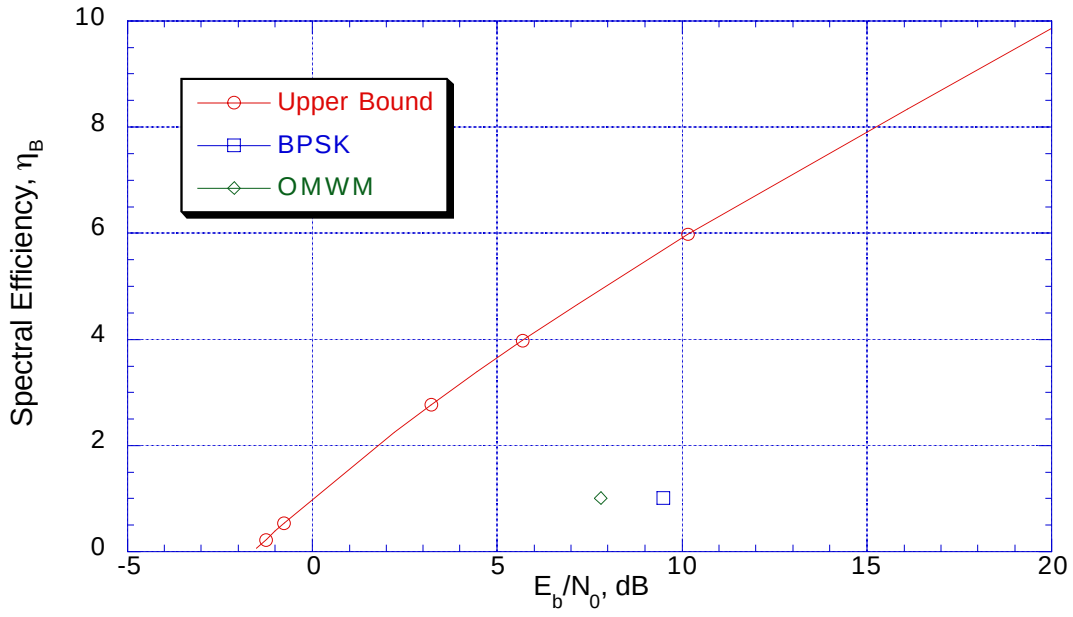


Figure 13.5: A comparison of performance of the example OMWM and Shannon's upper bound.

13.3 Spectral Characteristics of Stream OMWM

Orthogonal modulations with memory can both be used to achieve better performance or a modified transmitted spectrum compared to memoryless modulation. In general it is assumed in communications engineering that the information bits, $I(k)$, will be an independent (white), identically distributed sequence. In a stream modulation putting this white sequence into a OMWM will produce a modulation sequence that has memory. The memory in the modulation will produce a correlated time series $J(l),\ l = 1, N_c$ and this correlation will change the average transmitted signal energy spectrum per bit, $D_{x_z}(f)$. This section will explore this effect and how to characterize the transmitted energy spectrum from the modulation description.

13.3.1 An Example OMWM with a Modified Spectrum

This section consider an example of Alternate Mark Inversion (AMI) modulation. AMI modulation is an OMWM that is defined in Table 13.2. AMI codes and related codes are used in many telecommunications applications (e.g., T1 lines on coaxial cables) for reasons that will be apparent in the sequel. Consider a transmission with $K_b = 4$ ($N_f = 5$). This transmission has a trellis diagram shown in Fig. 13.6. In examining each of the paths of the trellis it is apparent that

$$\sum_{l=1}^{N_f} \tilde{d}_i(l) = 0. \tag{13.8}$$

This implies that if AMI modulation is used in a stream modulation then each transmission will not have a DC offset and that there will be a notch in the spectrum at $f = 0$. Precisely, if the data bits driving the trellis modulation are equally likely and independent then spectrum of the stream modulation will

$\sigma(l+1) = g_1\left(I(l), \sigma(l)\right)$

	$I(l)$	
State, $\sigma(l)$	0	1
1	1	2
2	1	2

$J(l) = g_2\left(I(l), \sigma(l)\right)$

	$I(l)$	
State, $\sigma(l)$	0	1
1	1	0
2	2	1

$\tilde{D}_z(l) = a(J(l))$

$J(l)$	$\tilde{D}_z(l)$
0	$-\sqrt{2}$
1	0
2	$\sqrt{2}$

Table 13.2: The finite state machine description of the AMI trellis code.

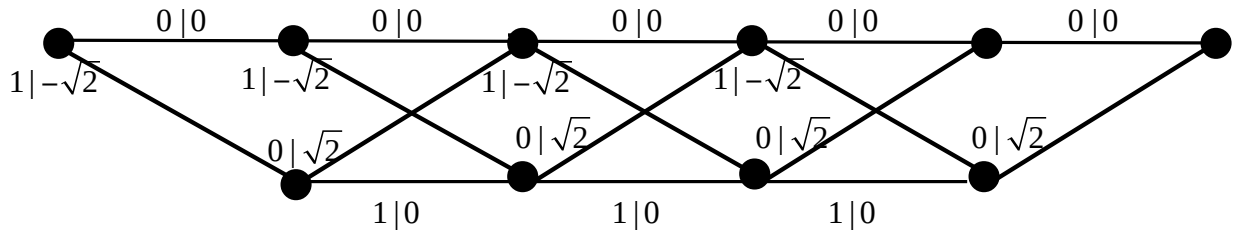


Figure 13.6: The trellis diagram for AMI with K_b=4.

be

$$\begin{aligned} D_{x_z}(f) &= \frac{E\left[G_{X_z}(f)\right]}{K_b} \qquad (13.9) \\ &= \frac{E\left[\left|\sum_l \tilde{D}_z(l)U(f)\exp[-j2\pi fT(l-1)]\right|^2\right]}{K_b} = \frac{\sum_{i=0}^{2^{K_b}-1}\left|\sum_l \tilde{d}_i(l)U(f)\exp[-j2\pi fT(l-1)]\right|^2}{2^{K_b}K_b}. \end{aligned}$$

For the example of AMI with $K_b = 4$ and stream modulation with

$$u_r(t) = \begin{cases} \sqrt{\frac{1}{T}} & 0 \le t \le T \\ 0 & \text{elsewhere} \end{cases} \qquad (13.10)$$

the average energy spectrum per bit is plotted in Fig. 13.7. This figure clearly shows the spectrum of an OMWM can be shaped (compare to Fig. 9.18) and this shaping ability is often why a modulation with memory is used in practice. The interesting aspect of AMI modulation is that it achieves about the same performance as BPSK modulation in addition to the spectral shaping characteristics (See Problem 13.14). This combination of spectral shaping without much loss in performance is the reason for AMI's adoption in many telecommunication applications.

13.3.2 Spectrum of OMWM for Large K_b

It is often of interest to evaluate the spectrum resulting from an OMWM in a stream modulation from a long frame of data (K_b large). It should be noted that the complexity of evaluating the average in (13.11) grows proportional to 2^{K_b}. Clearly for large frames this computation is increasingly impractical. Since there is a regular trellis structure to the modulation this can be exploited to compute the desired

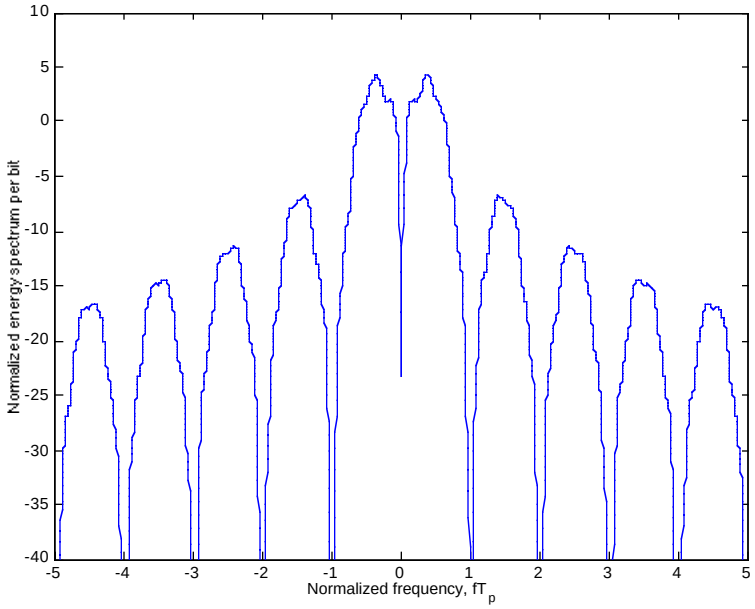


Figure 13.7: The spectrum of AMI used in a stream modulation with rectangular pulses. $K_b = 4$.

average energy spectrum per bit. To understand this structure we first expand

$$\begin{aligned} K_b D_{x_z}(f) &= E\left[\left|\sum_{l=1}^{N_f} \tilde{D}_z(l) U(f) \exp[-j2\pi fT(l-1)]\right|^2\right] \\ &= E_b \left|U(f)\right|^2 \sum_{l_1=1}^{N_f} \sum_{l_2=1}^{N_f} E\left[\tilde{D}_z(l_1)\tilde{D}_z^*(l_2)\right] \exp\left[-j2\pi f(l_1-l_2)T\right]. \end{aligned} \tag{13.11}$$

Since the trellis is the same every symbol in the frame and we assume the information bits are a white sequence, it is reasonable to assume that $\tilde{D}_z(l)$ is a stationary process. Since an OMWM typically starts from a fixed state and terminates to a fixed state this is not a true model as the distribution of the modulation at the frame boundaries will be different than the middle of the frame. If the frame is very long it is a good engineering approximation to ignore this non-stationarity at the frame boundaries. Also it is possible with certain OMWM modulation sequences can be produced in the middle of the frame that are not stationary time series but in general that will be the exception and not the rule. Consequently for a presentation that captures many of the important engineering applications the discussion will assume stationarity. The stationarity assumption implies that

$$E\left[\tilde{D}_z(l_1)\tilde{D}_z^*(l_2)\right] = R_{\tilde{D}}(l_1 - l_2). \tag{13.12}$$

Since all terms with the same time difference $l_1 - l_2$ will be constant the double sum can be rearranged to give

$$K_b D_{x_z}(f) = E_b \left|U(f)\right|^2 \sum_{m=-N_f+1}^{N_f-1} (N_f - |m|)\, R_{\tilde{D}}(m) \exp\left[-j2\pi fTm\right]. \tag{13.13}$$

Looking at the limit of a long frame size and noting that

$$\lim_{K_b\to\infty}\frac{N_f}{K_b}=1 \tag{13.14}$$

gives[1]

$$\lim_{K_b\to\infty}D_{x_z}(f)=E_b\left|U(f)\right|^2\sum_{m=-\infty}^{\infty}R_{\tilde{D}}(m)\exp\left[-j2\pi fTm\right]=E_b\left|U(f)\right|^2S_{\tilde{D}}\left(e^{j2\pi fT}\right). \tag{13.15}$$

where $S_{\tilde{D}}\left(e^{j2\pi fT}\right)$ is the power spectrum of the discrete random process $\tilde{D}_z(l)$ evaluated at a discrete frequency fT. It should be noted that $S_{\tilde{D}}\left(e^{j2\pi fT}\right)$ is a periodic function of f with a period of $1/T$. The transmitted energy spectrum per bit for stream modulation is seen to be a product of the pulse shape energy spectrum and the power spectral density of the discrete time modulation sequence. The form of the OMWM will determine the power spectrum of the discrete modulation sequence.

To compute $S_{\tilde{D}}\left(e^{j2\pi f}\right)$ the correlation function, $R_{\tilde{D}}(m)$, needs to be identified using the trellis structure of the OMWM. For a stationary random process the correlation function is given by

$$R_{\tilde{D}}(m)=\sum_{d_i\in\Omega_{\tilde{D}}}\sum_{d_j\in\Omega_{\tilde{D}}}d_id_j^*P_{\tilde{D}_z(l)\tilde{D}_z(l-m)}\left(d_i,d_j\right). \tag{13.16}$$

The key point in finding the correlation function is identifying an algorithm to compute the joint PMF of the modulation symbols at various time offsets from the trellis representation of the modulation. The modulation symbol at any point in time, $\tilde{D}_z(l)$, is determined entirely by specifying the edge of the trellis. The edges of the trellis can be enumerated by considering the modulation state at the same time, $\sigma(l)$, and the modulation state at the next time, $\sigma(l+1)$.

Definition 13.1 *An edge of the trellis at time l, denoted $S(l)$, is a pair of states $\{\sigma(l)=i,\sigma(l+1)=j\}$ such that $P\left(\sigma(l+1)=i\,|\sigma(l)=j\,\right)\neq 0$.*

For notational purposes we define the possible values that $S(l)$ can take as being the set Ω_s. The number of edges for a particular time near the middle of the frame for the $R=1$ OMWM is $\|\Omega_s\|=2N_s$. In general, the number of edges corresponding to a trellis transition is less than the total number of pairs $\{\sigma(l+1)=i,\sigma(l)=j\}$, $\|\Omega_\sigma\|^2=N_s^2$.

Property 13.1 *There is a functional mapping from $S(l)$ to $\tilde{D}_z(l)$.*

Consequently identifying $P_{S(l)S(l-m)}\left(s_i,s_j\right)$ will completely specify $P_{\tilde{D}_z(l)\tilde{D}_z(l-m)}\left(d_i,d_j\right)$.

Example 13.1: Consider AMI modulation again. There are $2N_s=4$ edges for AMI. The edges and the functional mapping to the modulation symbols for AMI are enumerated in Table 13.3

The needed joint probability, $P_{S(l)S(l-m)}\left(s_i,s_j\right)$, can be computed by forming some vectors and using linear algebra. First, the definition of conditional probability gives

$$P_{S(l)S(l-m)}\left(s_i,s_j\right)=P_{S(l)|S(l-m)}\left(s_i|s_j\right)P_{S(l-m)}\left(s_j\right). \tag{13.17}$$

Clearly from (13.17) the $P_{S(l-m)}\left(s_j\right)$ will be an important quantity to evaluate the spectrum of the OMWM. To have a concise notation define a vector (size $1\times 2N_s$) of the edge probabilities at time l

$$\vec{P}_{S(l)}=\left[P_{S(l)}\left(1\right)\;\;P_{S(l)}\left(2\right)\cdots P_{S(l)}\left(2N_s\right)\right] \tag{13.18}$$

[1]The readers will have to excuse the slightly inconsistent notation in this equation. This is the only place in the text where a discrete frequency variable and a continuous frequency variable both denoted with f are required in the same equation.

$s(l)$	$\sigma(l)$	$\sigma(l+1)$	$\tilde{D}_z(l)$
1	1	1	0
2	2	1	$-\sqrt{2}$
3	1	2	$\sqrt{2}$
4	2	2	0

Table 13.3: The edge enumeration for the AMI trellis code.

A one-step recursion of this probability can be obtained by using total probability and (13.17)

$$P_{S(l)}(s_i) = \sum_{n=1}^{2N_s} P_{S(l)S(l-1)}(s_i, n) = \sum_{n=1}^{2N_s} P_{S(l)|S(l1)}(s_i|n) P_{S(l-1)}(n) \tag{13.19}$$

This summation can be represented by a matrix equation

$$\vec{P}_s(l) = \vec{P}_s(l-1)\mathbf{S}_T \tag{13.20}$$

where $\mathbf{S}_T$ is often denoted a state transition matrix [Gal01] where

$$[\mathbf{S}_T]_{ij} = P_{S(l)|S(l-1)}(j|i). \tag{13.21}$$

The state transition matrix of an OMWM is easily identified by examining $g_1(I(k), \sigma(k))$. Examining $\vec{P}_s(l+1)$ it is clear that

$$\vec{P}_s(l+1) = \vec{P}_s(l)\mathbf{S}_T = \vec{P}_s(l-1)\mathbf{S}_T^2 \tag{13.22}$$

and by induction

$$\vec{P}_s(l) = \vec{P}_s(l-m)\mathbf{S}_T^m. \tag{13.23}$$

Examining (13.21) it is clear that

$$[\mathbf{S}_T^m]_{ij} = P_{S(l)|S(l-m)}(j|i). \tag{13.24}$$

Consequently half of the characterization necessary for the PMF in (13.17) is given by (13.24) and only a form for $P_{S(l)}(s_i)$ is needed for a complete evaluation of $R_{\tilde{D}}(m)$. Most often the stationary probability of each edge is equally likely, but in the case that an unusual chain is encountered the theory of Markov chains [Gal01] can be used to solve for the stationary probability distribution of the edges.

Example 13.2: Consider again the AMI modulation. The edge state transition matrix is

$$\mathbf{S}_T = \begin{bmatrix} 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0 & 0.5 \end{bmatrix} \tag{13.25}$$

and using a uniform edge probability, $P_{S(l-m)}(i) = 1/4,\ i = 1, 4$, gives

$$S_{\tilde{D}}\left(e^{j2\pi f}\right) = -0.5e^{-j2\pi f} + 1 - 0.5e^{j2\pi f} = 1 - \cos(2\pi f) \tag{13.26}$$

Equation (13.26) shows that AMI used as an OMWM in a stream modulation will always produce a notch at DC. This condition is evident in Fig. 13.7 where a short frame was considered ($K_b = 4$). This characteristic is advantageous in telecommunications systems as it prevents a large DC current from being driven over the coaxial or twisted pair cable prevalent in the telecommunications network.

13.4 Varying Transmission Rates with OMWM

The OMWM that has been presented so far has limited utility as it only allows for a rate of one information bit per modulated symbol. This is usually denoted one bit per channel utilization by communications engineers. OMWM of this type only allows a communication engineer to design a modulation that would move horizontally on the rate performance curve of Fig. 13.5. In certain situations it might be desired to lower the bit rate and achieve higher fidelity at a lower E_b/N_0 (e.g., deep space communications). Alternately it might be useful to achieve more than one bit per channel utilization. This is often the case when bandwidth is a scarce resource and the signal to noise ratio is relatively high (e.g., cable modems). This section will show how to generalize the OMWM to achieve these varying information transmission rates.

The key to this generalization is to enable multiple bit inputs and multiple symbol outputs in the OMWM. The general orthogonal modulation with memory, as considered in this section, consists of a finite state machine operating at an integer fraction of the symbol rate, $1/N_mT$. The K_b bits to be transmitted are broken up into blocks of K_s in length (a total of $N_b = K_b/K_s$ blocks per frame. At each symbol time a new set of K_s bits, $\vec{I}(m)$ is input into a finite state machine and this produces a new constellation label, $J(m)$, and a new modulation state, $\sigma(m+1)$. The constellation label, $J(m) = \vec{J}(m)$ of the finite state machine is used as an input to a modulator that produces an N_m symbol block of M_s-ary modulation symbols. Again to keep a consistent normalization the mapping, $\tilde{D}_z(l) = a\left(J_i(m)\right)$, is selected such that $E\left[\left|\tilde{D}_z(l)\right|^2\right] = R$. A total of N_b trellis transitions are needed to communicate the K_b bits. Fig. 13.8 shows the block diagram for a general linear orthogonal modulation with memory. N_s again denotes the number of states in the modulation and the nonlinear equations governing the updates are

$$\sigma(m+1) = g_1\left(\sigma(m), \vec{I}(m)\right) \tag{13.27}$$

$$\vec{J}(m) = g_2\left(\sigma(m), \vec{I}(m)\right). \tag{13.28}$$

Note $J(m) = 0, M_s^{N_m} - 1$ where each component of $\vec{J}(m)$ only take values $J_i(m) = 0, M_s - 1,\ i = 1, N_m$. The constellation label at time m and in position i will generate the modulation symbol at time $(m-1)N_m + i$ for the orthogonal modulation. In general it is again usually desirable to have ν_c extra symbols transmitted to return the modulation to a common final state at the end of the transmission frame. The total length of the frame for the orthogonal modulation is denoted N_f, hence, $N_f = N_bN_m + \nu_c$, where ν_c is a code dependent constant. Due to termination the effective rate is $R_{eff} = \frac{K_b}{N_bN_m+\nu_c} = \frac{K_bK_m}{K_bN_m+\nu_c}$ but if a large number of bits are transmitted then the rate becomes approximately $R = K_m/N_m$. If a rate less than one is desired then a communication engineer would choose $N_m > K_m$. If a higher transmission rate is desired then the communication engineer would choose $N_m < K_m$.

13.4.1 Spectrum of Stream Modulations with $N_m > 1$

When $N_m = 1$ the spectral characteristics of stream modulation can be computed as in Section 13.3, but if $N_m > 1$ the techniques of Section 13.3 need to be generalized. If K_b is small all the waveforms can be detailed out and the average energy spectrum can be easily computed in exactly the same way as previously detailed. This section will concentrate on evaluation of $D_{x_z}(f)$ given as

$$K_bD_{x_z}(f) = E_b\left|U(f)\right|^2 \sum_{l_1=1}^{N_f}\sum_{l_2=1}^{N_f} E\left[\tilde{D}_z(l_1)\tilde{D}_z^*(l_2)\right]\exp\left[-j2\pi f(l_1-l_2)T\right] \tag{13.29}$$

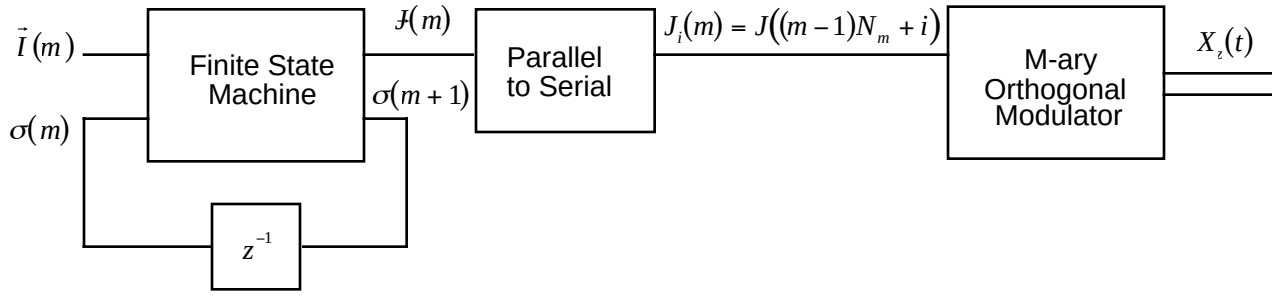


Figure 13.8: The block diagram for a general modulation with memory.

in the case of K_b large. Again the regular trellis structure of the modulation can be exploited. Again it is assumed that the information bits are a white sequence. Since the trellis transition is driven by a white process but the modulation symbols are produced N_m at a time $\tilde{D}_z(l)$ is a cyclo-stationary process [SW02]. The cyclo-stationarity assumption implies that

$$E\left[\tilde{D}_z(l_1)\tilde{D}_z^*(l_2)\right] = R_{\tilde{D}}(l_1 - l_2, l) \qquad l = 1, \ldots, N_m, \tag{13.30}$$

i.e., the correlation could be a function of the time in the between the transitions. The double sum can be rearranged in a similar way to that of Section 13.3.2 to give

$$K_b D_{x_z}(f) = E_b \left|U(f)\right|^2 \sum_{m=-N_f+1}^{N_f-1} \sum_{l=1}^{N_m} \left(N_b - |m| + g(m,l)\right) R_{\tilde{D}}(m,l) \exp\left[-j2\pi f T m\right]. \tag{13.31}$$

Again looking at the limit of a long frame size gives

$$\lim_{K_b \to \infty} D_{x_z}(f) = K_m E_b \left|U(f)\right|^2 \sum_{m=-\infty}^{\infty} \sum_{l=1}^{N_m} R_{\tilde{D}}(m,l) \exp\left[-j2\pi f T m\right]. \tag{13.32}$$

Realized that the function

$$\sum_{l=1}^{N_m} R_{\tilde{D}}(m,l) = \bar{R}_{\tilde{D}}(m) \tag{13.33}$$

now is analogous to the $R_{\tilde{D}}(m)$ of Section 13.3.2. Once this function is identified then the desired average energy spectrum can be computed as

$$\lim_{K_b \to \infty} D_{x_z}(f) = K_m E_b \left|U(f)\right|^2 \bar{S}_{\tilde{D}}\left(e^{j2\pi f T}\right). \tag{13.34}$$

where $\bar{S}_{\tilde{D}}\left(e^{j2\pi f T}\right)$ is the discrete Fourier transform of $\bar{R}_{\tilde{D}}(m)$. It should again be noted that $\bar{S}_{\tilde{D}}\left(e^{j2\pi f T}\right)$ is a periodic function of f with a period of $1/T$. The transmitted energy spectrum per bit for stream modulation is again seen to be a product of the pulse shape energy spectrum and the effective power spectral density of the discrete time modulation sequence. The form of the OMWM will determine the power spectrum of the discrete modulation sequence.

To compute $\bar{S}_{\tilde{D}}\left(e^{j2\pi f}\right)$ the correlation function, $R_{\tilde{D}}(m,l)$, needs to be identified using the trellis structure of the OMWM. The correlation function is given by

$$R_{\tilde{D}}(m,l) = \sum_{d_i \in \Omega_{\tilde{D}}} \sum_{d_j \in \Omega_{\tilde{D}}} d_i d_j^* P_{\tilde{D}_z((k-1)N_m+l)\tilde{D}_z((k-1)N_m+l-m)}(d_i, d_j). \tag{13.35}$$

$\sigma(m+1) = g_1\left(I(m), \sigma(m)\right)$

State, $\sigma(m)$	$I(m)$ = 0	$I(m)$ = 1
1	1	2
2	3	4
3	5	6
4	7	8
5	1	2
6	3	4
7	5	6
8	7	8

$\vec{J}(m) = g_2\left(I(m), \sigma(m)\right)$

State, $\sigma(m)$	$I(m)$ = 0	$I(m)$ = 1
1	[0 0]	[1 1]
2	[0 1]	[1 0]
3	[1 1]	[0 0]
4	[1 0]	[0 1]
5	[1 1]	[0 0]
6	[1 0]	[0 1]
7	[0 0]	[1 1]
8	[0 1]	[1 0]

Table 13.4: The finite state machine description of an example convolutional code.

The key point in finding the correlation function is identifying an algorithm to compute the joint PMF of the modulation symbols at various time offsets from the trellis representation of the modulation. Recall the modulation symbol at any point in time, $\tilde{D}_z(l)$, is determined entirely by specifying the edge of the trellis. The edges of the trellis can be enumerated as in Section 13.3.2.

Property 13.2 *There is a functional mapping from $S(k)$ to $\tilde{D}_z((k-1)N_m + l)$.*

The needed joint probability for the edges, $P_{S(l)S(l-m)}\left(s_i, s_j\right)$, can be computed using the techniques introduced in Section 13.3.2, i.e.,

$$[\mathbf{S}_T^m]_{ij} = P_{S(l)|S(l-m)}\left(j|i\right). \tag{13.36}$$

Consequently half of the characterization necessary for the joint PMF of two edges is given by (13.36) and only a form for $P_{S(l)}\left(s_i\right)$ is needed for a complete evaluation of $R_{\tilde{D}}(m,l)$.

The major difference in the computation of $R_{\tilde{D}}(m,l)$ is working out the number of trellis transitions as a function of l. If $m = N_m k + n$ for $n = 0, \ldots, N_m - 1$ then there will need to be k trellis transitions modeled for $l > n$ and $k+1$ trellis transitions for $l \leq n$. Once the number of transitions are computed then $R_{\tilde{D}}(m,l)$ can be computed using the techniques of Section 13.3.2 and the symbol mappings for each edge. Once $R_{\tilde{D}}(m,l)$ is computed then $\bar{R}_{\tilde{D}}(m)$ can be evaluated. A discrete Fourier transform leads to the complete characterization of the spectrum of the OMWM. A detailed example will be considered in the sequel.

13.4.2 Example $R < 1$: Convolutional Codes

Consider an OMWM consisting of an eight state convolutional code [LC04] (N_s=8) using BPSK modulation (M_s=2) that has $R = 1/2$ ($K_m = 1$ and $N_m = 2$). The modulation updates are given in Table 13.4. Since this code only sends 1 bit every two symbols the spectral efficiency of the modulation ($\eta_B \approx 1/2$ for this OMWM) will be reduced compared to BPSK, for example. Note that the modulation is normalized to have an average energy of $K_m = 1$ per trellis transition, i.e.,

$$\tilde{D}_z((m-1)N_m + i) = \frac{1}{\sqrt{2}}\left(-1\right)^{J_i(m)} \tag{13.37}$$

The trellis for this modulation is shown in Fig. 13.9. There are a total of $2N_s = 16$ edges in this example. The edge labels indicate the input that causes the transition and the modulation symbol that is output for that input and state. This particular code requires at most three transitions to allow the

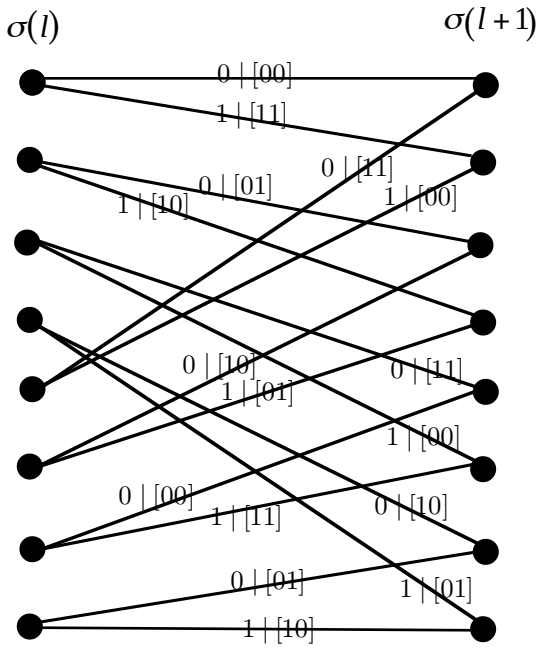


Figure 13.9: The trellis diagram of the example convolutional code.

modulation to transition from any state to any other state so this code has $\nu_c = 3$. Again N_f trellis sections can be combined together to get a complete description of an OMWM. For instance for $K_b = 4$, Fig. 13.10 shows a trellis description for a modulation that has $\sigma(1) = 1$ and returns to $\sigma(8) = 1$. One can verify that there are 2^{K_b} paths through this trellis. To understand the performance of this particular OMWM squared Euclidean distance spectrum should be computed. The entire distance spectrum of this modulation is explored in the homework problems (see Problem 13.16) and the union bound to the probability of word error is plotted in Fig. 13.11. The important thing to notice is that the performance is markedly improved compared to a modulation that does not use memory. For instance if we compare performance at $P_W(E) = 10^{-5}$ then this example convolutional code is about 2dB better that the 16FSK, which is the memoryless modulation we considered in Chapter 8 with roughly the same spectral efficiency.

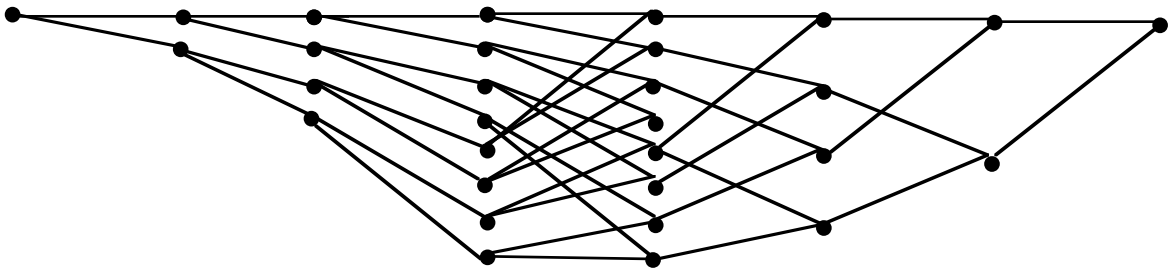


Figure 13.10: The trellis diagram of the example convolutional code. $K_b = 4$.

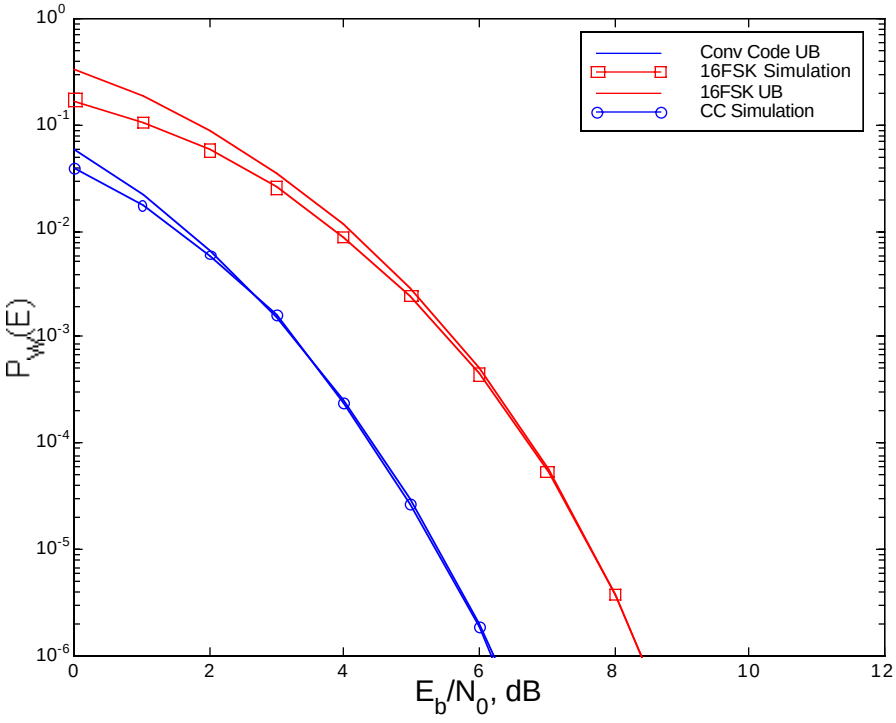


Figure 13.11: The union bound of the example convolutional code OMWM and frame error rate of $K_b = 4$ memoryless orthogonally modulated bits using 16FSK.

$\sigma(m+1) = g_1\left(\vec{I}(m), \sigma(m)\right)$

State, $\sigma(m)$	$\vec{I}(m)$			
	0 0	1 0	0 1	1 1
1	1	1	2	2
2	3	3	4	4
3	1	1	2	2
4	3	3	4	4

$J(m) = g_2\left(\vec{I}(m), \sigma(m)\right)$

State, $\sigma(m)$	$\vec{I}(m)$			
	0 0	1 0	0 1	1 1
1	0	4	2	6
2	1	5	3	7
3	2	6	0	4
4	3	7	1	5

Table 13.5: The finite state machine description of an example trellis coded modulation.

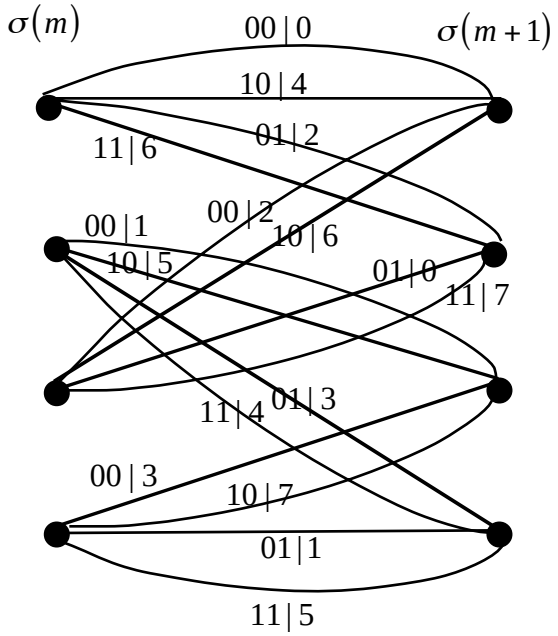


Figure 13.12: The trellis diagram of the example trellis coded modulation.

13.4.3 Example $R > 1$: Trellis Codes

Consider an OMWM consisting of a four state trellis coded modulation [Ung82] (N_s=4) using 8PSK modulation (M_s=8) that has $R = 2$ ($K_m = 2$ and $N_m = 1$). The modulation updates are given in Table 13.5. Note that the modulation is normalized to have an average energy of $R = 2$ per trellis transition, i.e.,

$$\tilde{D}_z(m) = \sqrt{2}\exp\left[\frac{j\pi(2J(m)+1)}{8}\right] \tag{13.38}$$

The trellis for this modulation is shown in Fig. 13.9. There are a total of $2^{K_m}N_s = 16$ edges in this example and there are parallel edges in the trellis. Parallel edges connect the same states but are associated with different input information bits and output modulations symbols. This particular code requires at most two transitions to allow the modulation to transition from any state to any other state so this code has $\nu_c = 2$.

Again N_f trellis sections can be combined together to get a complete description of an OMWM. For instance for $K_b = 4$, Fig. 13.13 shows a trellis description for a modulation that has $\sigma(1) = 1$ and

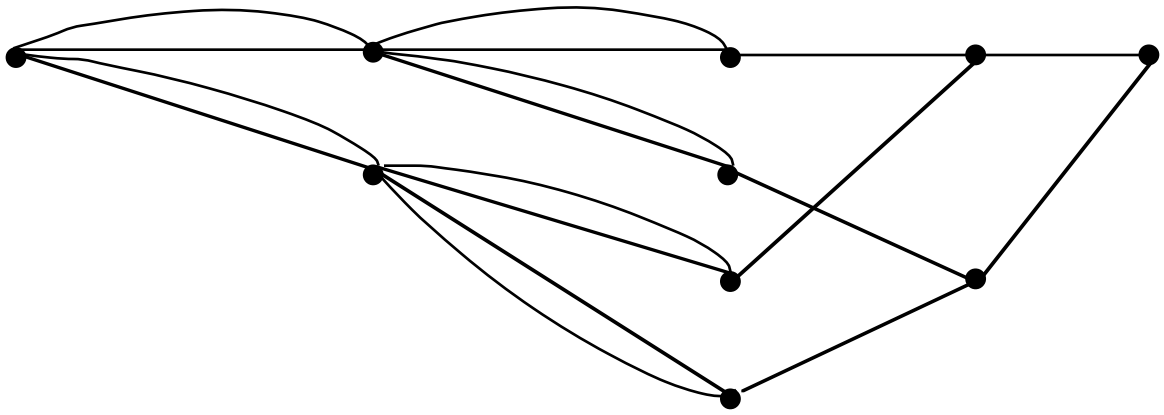


Figure 13.13: The trellis diagram of the example trellis coded modualtion. $K_b = 4$.

returns to $\sigma(5) = 1$. One can verify that there are 2^{K_b} paths through this trellis. It is interesting to note that due to the parallel transitions there is more than on way to terminate the trellis for each of the states (4 possibilities in this example). These extra paths during termination would allow you to transmit $K_b = 6$ bits in the same time ($N_f = 4$) as $K_b = 4$ bits without much loss in performance. To understand the performance of this particular OMWM, the squared Euclidean distance spectrum should be computed. The entire distance spectrum of this modulation is explored in the homework problems (see Problem 13.17) and the union bound to the probability of word is plotted in Fig. 13.14. The important thing to notice is that again the performance is markedly improved compared to a modulation that does not use memory. For instance if we compare performance at $P_W(E) = 10^{-5}$ then this example trellis coded modulation is about 3dB better than QPSK, which is the memoryless modulation we considered in Chapter 8 with the same spectral efficiency.

13.4.4 Example for Spectral Shaping: Miller Code

Consider an OMWM consisting of a four state coded modulation [Mil63] (N_s=4) using BPSK modulation (M_s=2) that has $R = 1/2$ ($K_m = 1$ and $N_m = 2$) known as the Miller code. The modulation updates are given in Table 13.6. Again the modulation is normalized to have an average energy of $R = 1$ per trellis transition, i.e.,

$$\tilde{D}_z((m-1)N_m + i) = \frac{1}{\sqrt{2}}(-1)^{J_i(m)} \tag{13.39}$$

The trellis for this modulation is shown in Fig. 13.15. There are a total of $2N_s = 8$ edges in this example. The edge labels indicate the input that causes the transition and the modulation symbol that is output for that input and state. A Miller code is typically not terminated as it is not used to achieve a desired performance (squared Euclidean distance) and it is not possible to return to a common state in an identical number of transitions from each state. To understand the performance of this particular OMWM, the squared Euclidean distance spectrum again should be computed and this distance spectrum is explored in the homework problems (see Problem 13.19).

Since the Miller code was not conceived to improve the distance properties but to change the time and spectral characteristics of an orthogonal modulation, it is of interest to explore these properties. An interesting property of the Miller code is that at most there are 4 symbols in a row that are the same polarity. Codes of this type that limit the number of symbols between a transition are often referred

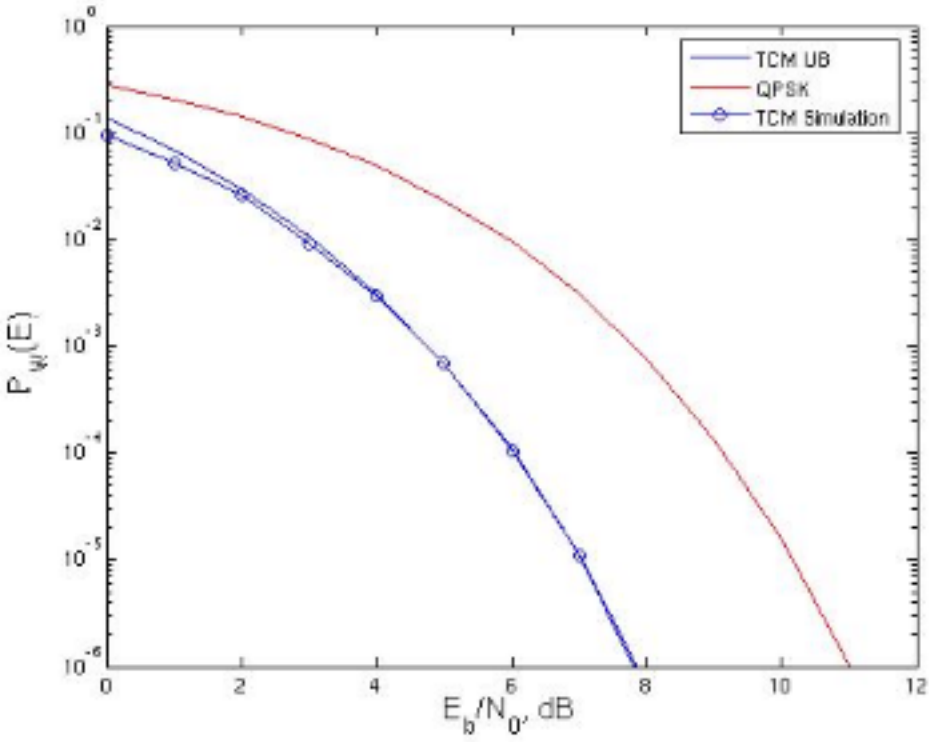


Figure 13.14: The union bound and simulated performance of the example $R = 2$ OMWM and frame error rate of QPSK. $K_b = 4$.

$\sigma(m+1) = g_1\left(I(m), \sigma(m)\right)$

	$I(m)$	
State, $\sigma(m)$	0	1
1	2	3
2	1	4
3	2	4
4	1	3

$\vec{J}(m) = g_2\left(I(m), \sigma(m)\right)$

	$I(m)$	
State, $\sigma(m)$	0	1
1	[0 1]	[0 0]
2	[1 0]	[1 1]
3	[0 1]	[1 1]
4	[1 0]	[0 0]

Table 13.6: The finite state machine description of the Miller code.

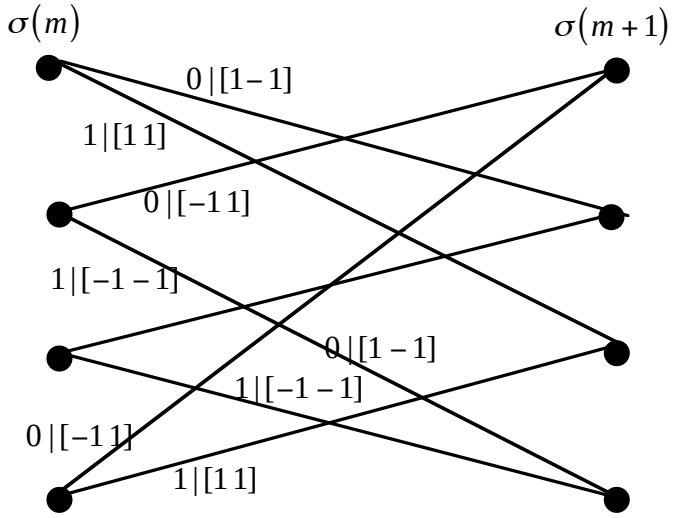


Figure 13.15: The trellis diagram of the Miller code.

to as run length limited (RLL) codes. This type of code is important in magnetic recording due to the physics of magnetic recording. A more important characteristic of the Miller code is that the memory of the code shapes the transmitted spectrum. This shaped spectrum again has advantages in magnetic recording and playback. This spectrum can be explored using the results of Section 13.4.1.

Finding the joint probability distribution of the edges is the important step in finding the transmitted spectrum. The edge enumeration of the Miller code is given in Table 13.7. Recall the edge probability update to compute the needed modulation symbol correlation is given as

$$\vec{P}_s(l) = \vec{P}_s(l-1)\mathbf{S}_T \tag{13.40}$$

where by examining Table 13.7 we have

$$\mathbf{S}_T = \begin{bmatrix} 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 \end{bmatrix}. \tag{13.41}$$

This edge transition probability can be used to compute $R_{\tilde{D}}(m,l)$. For example if $\vec{\tilde{D}}(l)$ is defined as the vector of symbol outputs at time $(m-1)N_m + l$ for all the different edges $S(m)$ then we have

$$R_{\tilde{D}}(1,1) = \sum_{j=1}^{2N_s}\sum_{i=1}^{2N_s}\left[\vec{\tilde{D}}(2)\right]_j\left[\vec{\tilde{D}}(1)\right]_i^*\left[\vec{P}_s(l)\right]_i\left[\vec{P}_s(l)\mathbf{S}_T^0\right]_j = 0 \tag{13.42}$$

and

$$R_{\tilde{D}}(1,2) = \sum_{j=1}^{2N_s}\sum_{i=1}^{2N_s}\left[\vec{\tilde{D}}(1)\right]_j\left[\vec{\tilde{D}}(2)\right]_i^*\left[\vec{P}_s(l)\right]_i\left[\vec{P}_s(l)\mathbf{S}_T^1\right]_j = 0.25. \tag{13.43}$$

$s(m)$	$\sigma(m)$	$\sigma(m+1)$	$\vec{\tilde{D}}_z(m)$
1	2	1	[1 0]
2	4	1	[1 0]
3	1	2	[0 1]
4	3	2	[0 1]
5	1	3	[0 0]
6	4	3	[0 0]
7	2	4	[1 1]
8	3	4	[1 1]

Table 13.7: The edge enumeration for the Miller code.

This leads to $\bar{R}_{\tilde{D}}(1) = 0.25$. The average correlation, $\bar{R}_{\tilde{D}}(m)$, can be computed numerically using results like (13.42) and (13.43) for all m and the DFT can be taken of $\bar{R}_{\tilde{D}}(m)$ to give the final result, $\bar{S}_{\tilde{D}}\left(e^{j2\pi fT}\right)$. The plot of modulation symbol average power spectrum is shown in Fig. 13.16. Having a spectrum of the modulation concentrated at higher frequencies in magnetic recording is important since lower frequencies of the modulation tend to interfere with the servo mechanism of the magnetic read/write head. The Miller code is another example of how the OMWM can be designed to achieve a desired spectral characteristic.

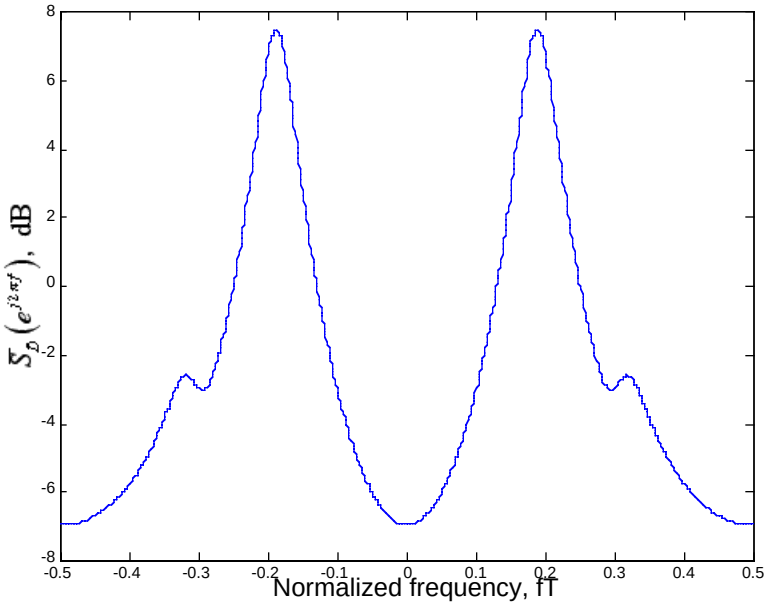
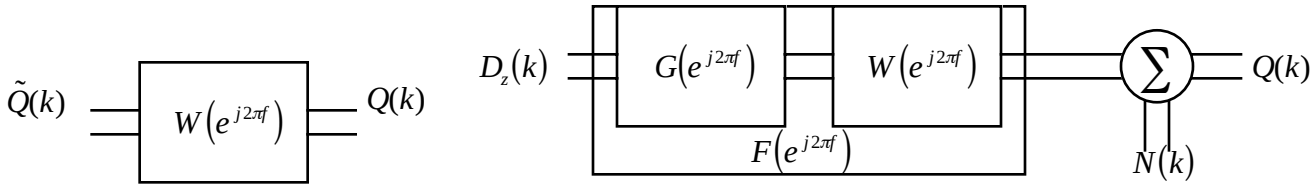


Figure 13.16: The effective power spectrum of the modulation symbols for Miller coded modulation.

13.5 Forney Equivalence

Linear stream modulation on frequency selective channels that was introduced in Chapter 11 and OMWM that was introduced in this chapter can be viewed in a unified framework. This equivalence of



a) Transformation. b) Signal model.

Figure 13.17: The linear transformation that turns a stream modulation on a frequency selective channel into a orthogonal modulation with memory.

stream modulation on a frequency selective channel and an OMWM is herein denoted Forney equivalence [For72]. In this section we want to clearly distinguish between memoryless stream modulations on a frequency selective channel ($\tilde{Q}(k)$ and $D_z(l)$) and the OMWM ($Q(k)$ and $\tilde{D}_z(l)$). The discussion will again focus on the case of a binary stream modulation (using BPSK) but the generalization is straightforward. Recall that the optimum demodulator for memoryless stream modulation, $D_z(l)$, on a frequency selective channel is

$$\begin{aligned}
\hat{\vec{I}} &= \arg \max_{i=0,M-1} \tilde{T}_i \\
&= \arg \max_{i=0,M-1} \sum_{k=1}^{K_b} \Re\left[d_i^*(k)\tilde{Q}(k)\right] - \frac{E_i}{2} \\
&= \arg \max_{i=0,M-1} \sqrt{E_b} \sum_{k=1}^{K_b} \Re\left[d_i^*(k)\tilde{Q}(k)\right] - \frac{1}{2}\sum_{l_1=1}^{K_b}\sum_{l_2=1}^{K_b} d_i(l_1)d_i^*(l_2)V_{\tilde{u}}((l_2-l_1)T)
\end{aligned} \tag{13.44}$$

where $\tilde{Q}(k)$ is the output sample of the match filter to the distorted pulse for the k^{th} transmitted symbol. In contrast, the orthogonal modulation with memory has an optimum demodulator of

$$\begin{aligned}
\hat{\vec{I}} &= \arg \max_{i=0,M-1} T_i. \\
&= \arg \max_{i=0,M-1} \sqrt{E_b} \sum_{k=1}^{N_f} \Re\left[\tilde{d}_i^*(k)Q(k)\right] - \frac{E_b}{2}\sum_{k=1}^{N_f}\left|\tilde{d}_i(k)\right|^2
\end{aligned} \tag{13.45}$$

where again N_f is the number of transmitted symbols from the modulation with memory. The goal of this section is to show that a simple linear filter, $W\left(e^{j2\pi f}\right)$ transforms the demodulator for the frequency selective channel into an equivalent OMWM. This transformation is shown in Fig. 13.17-a). Recall that stream modulation on a frequency selective channel is characterized by the discrete time linear filter $G\left(e^{j2\pi f}\right)$. The concatenation of the filter $G\left(e^{j2\pi f}\right)$ with the filter $W\left(e^{j2\pi f}\right)$ results in a filter $F\left(e^{j2\pi f}\right)$. $F\left(e^{j2\pi f}\right)$ is known as the equivalent discrete time white noise filter model of the channel. The diagram of the whitened signal model is shown in Fig. 13.17-b).

The form of the simple linear filter can be found using the results that were developed in the derivation of the DFE in Chapter 11. Recall that

$$\tilde{Q}(k) = \sum_{l=1}^{K_b} D_z(l)g(k-l) + \tilde{N}(k) \tag{13.46}$$

where the noise is colored with $R_{\tilde{N}}(m) = E_b N_0 g(m)$. The spectral factorization for the PSD of this noise is given as

$$S_{\tilde{N}}\left(e^{j2\pi f}\right) = N_0 E_b G\left(e^{j2\pi f}\right) = N_0 S_1\left(e^{j2\pi f}\right) = N_0\gamma_1 F_1^+\left(e^{j2\pi f}\right)F_1^-\left(e^{j2\pi f}\right). \tag{13.47}$$

Property 13.3 Forney Equivalence *[For72] Choosing* $W\left(e^{j2\pi f}\right) = \left(\sqrt{\gamma_1}F_1^-\left(e^{j2\pi f}\right)\right)^{-1}$ *provides the desired transformation where* $\tilde{D}_z(l) = \sum_{m=0}^{N_u} D_z(l-m)f_1^+(m)$ *where* $f_1^+(n)$ *is the impulse response of the filter having a frequency response* $\sqrt{\gamma_1}F_1^+\left(e^{j2\pi f}\right)$.

It should be noted here that $W\left(e^{j2\pi f}\right)$ is an anti-causal filter and $\sqrt{\gamma_1}F_1^+\left(e^{j2\pi f}\right)$ is a causal filter. Also in the spectral factorization $\left(f_1^+(n)\right)^* = f_1^-(-n)$ and $f_1^+(n)$ is a filter of length N_u as defined in Chapter 11. Since K_b bits are transmitted the equivalent OMWM will have $N_f = K_b + N_u$.

Proof: The goal is to show that (13.44) and (13.45) become identical when the transforming filter is chosen as $W\left(e^{j2\pi f}\right) = \left(\sqrt{\gamma_1}F_1^-\left(e^{j2\pi f}\right)\right)^{-1}$. The physical structure that is being considered is given in Fig. 13.18. Choosing $W\left(e^{j2\pi f}\right) = \left(\sqrt{\gamma_1}F_1^-\left(e^{j2\pi f}\right)\right)^{-1}$ and denoting the filter length to be $N_u + 1$ implies that

$$\tilde{Q}(k) = \sum_{n=0}^{N_u} f_1^-(-n)Q(k+n) = \sum_{l=1}^{K_b+N_u} f_1^-(k-l)Q(l) = \sum_{l=1}^{N_f} f_1^-(k-l)Q(l) \qquad k = 1, K_b. \tag{13.48}$$

The matched filter to the transmitted sequence (the first term of (13.45)) is then given as

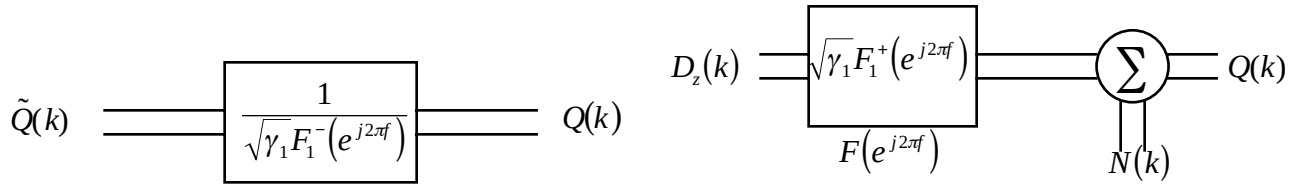
$$\begin{aligned}
\Re\left[\sum_{k=1}^{N_f} \tilde{d}_i^*(k)Q(k)\right] &= \Re\left[\sum_{k=1}^{N_f}\sum_{m=1}^{K_b} d_i^*(m)\left(f_1^+(k-m)\right)^* Q(k)\right] \\
&= \Re\left[\sum_{m=1}^{K_b} d_i^*(m)\sum_{k=1}^{K_b+N_u}\left(f_1^+(k-m)\right)^* Q(k)\right] \\
&= \Re\left[\sum_{m=1}^{K_b} d_i^*(m)\sum_{k=1}^{K_b+N_u} f_1^-(m-k)Q(k)\right] \\
&= \Re\left[\sum_{m=1}^{K_b} d_i^*(m)\tilde{Q}(m)\right]
\end{aligned} \tag{13.49}$$

Consequently the first terms of (13.44) and (13.45) (the matched filter terms) are equivalent when we have set $\tilde{D}_z(l) = \sum_{m=0}^{N_u} D_z(l-m)f_1^+(m)$. Additionally

$$\begin{aligned}
\sum_{k=1}^{N_f}\left|\tilde{d}_i(k)\right|^2 &= \sum_{k=1}^{N_f}\left|\sum_{m=0}^{N_u} d_i(k-m)f(m)\right|^2 \\
&= \sum_{k=1}^{N_f}\left(\sum_{l=1}^{K_b} d_i(l)f_1^+(k-l)\right)\left(\sum_{m=1}^{K_b} d_i^*(m)\left(f_1^+(k-m)\right)^*\right) \\
&= \sum_{l=1}^{K_b}\sum_{m=1}^{K_b} d_i(l)d_i^*(m)\sum_{k=1}^{N_f} f_1^+(k-l)\left(f_1^+(k-m)\right)^* \\
&= \sum_{l=1}^{K_b}\sum_{m=1}^{K_b} d_i(l)d_i^*(m)\sum_{n=0}^{N_u} f_1^+(n)f_1^-((m-l)-n).
\end{aligned} \tag{13.50}$$

The properties of the spectral factorization theorem in Appendix B then produces

$$\sum_{n=0}^{N_u} f_1^+(n)f_1^-((m-l)-n) = E_b g(m-l) = E_b V_{\tilde{u}}((m-l)T). \tag{13.51}$$



a) Transformation. b) Signal model.

Figure 13.18: The linear transformation that turns a stream modulation on a frequency selective channel into a orthogonal modulation with memory.

Using this in (13.50) gives

$$\sum_{k=1}^{N_c} \left| \tilde{d}_i(k) \right|^2 = \sum_{l=1}^{K_b} \sum_{m=1}^{K_b} d_i(l) d_i^*(m) V_{\tilde{u}}((m-l)T). \tag{13.52}$$

Equation (13.52) shows that the second terms (the energy correction terms) of (13.44) and (13.45) are equivalent when we have set $\tilde{D}_z(l) = \sum_{m=0}^{N_u} D_z(l-m) f_1^+(m)$ and this implies that optimum demodulation of memoryless stream modulations on a frequency selective channel can be made to be equivalent to optimum demodulation of orthogonal modulations with memory by the given transformation. □

The discrete time equivalent whiten filter model is now given in Fig. 13.18. This model also is causal. The equivalent modulation with memory is given as

$$\tilde{D}_z(l) = \sum_{m=0}^{N_u} D_z(l-m) f_1^+(m). \tag{13.53}$$

Since this modulation with memory is a function of the latest $N_u + 1$ information bits it is apparent that the modulation state is given as $\sigma(k) = [I(k-1), \cdots, I(k-N_u)]$ and consequently there are 2^{N_u} states in the modulation with memory.

$\sigma(l+1) = g_1(I(l), \sigma(l))$

	$I(l)$	
$\sigma(l)$	0	1
1	1	2
2	1	2

$\tilde{D}_z(l) = a(g_2(I(l), \sigma(l)))$

	$I(l)$	
$\sigma(l)$	0	1
1	$\sqrt{0.51}\left(1+0.98e^{j\pi/3}\right)$	$\sqrt{0.51}\left(-1+0.98e^{j\pi/3}\right)$
2	$\sqrt{0.51}\left(1-0.98e^{j\pi/3}\right)$	$\sqrt{0.51}\left(-1-0.98e^{j\pi/3}\right)$

Table 13.8: The finite state machine description of the equivalent orthogonal modulation with memory.

Example 13.3: Consider again the channel given in Example 11.1, where $\alpha_1 = 0.7\exp[j\pi/3]$ and $\tau_1 = T$. The signal spectrum, $S_1\left(e^{j2\pi f}\right)$, is given as

$$\begin{aligned} S_1\left(e^{j2\pi f}\right) &= E_b G\left(e^{j2\pi f}\right) = \gamma_1 F_1^-\left(e^{j2\pi f}\right) F_1^+\left(e^{j2\pi f}\right) \\ &= E_b 0.51\left(1+0.98\exp[j(2\pi f-\pi/3)]\right)\left(1+0.98\exp[-j(2\pi f-\pi/3)]\right). \end{aligned} \tag{13.54}$$

It should be noted here that this channel only has one symbol worth of ISI. The anti-causal whitening filter is given as

$$W\left(e^{j2\pi f}\right) = \frac{1.4}{\sqrt{E_b}\left(1+0.98\exp[j(2\pi f-\pi/3)]\right)} = 1.4\sum_{i=0}^{\infty}\left(-0.98\exp[j(2\pi f-\pi/3)]\right)^i. \tag{13.55}$$

It should be noted that the anti-causal whitening filter has an infinite impulse response. The whitening filter outputs will be of the form

$$\begin{aligned} Q(k) &= 1.4\tilde{Q}(k) - 1.372\exp[-j\pi/3]\,\tilde{Q}(k+1) + 1.344\exp[-j2\pi/3]\,\tilde{Q}(k+2) - \ldots \\ &= \sqrt{E_b}\left(\sqrt{0.51}D_z(k) + \sqrt{0.51}0.98\exp[j\pi/3]\,D_z(k-1)\right) + N(k) \end{aligned} \tag{13.56}$$

When BPSK modulation is used on this channel the equivalent modulation with memory that exists at the output of the anti-causal whitening filter is given in Table 13.8. It should be noted that the binary modulation (BPSK) has been transformed into a 4-ary modulation with memory with a constellation that is channel dependent. The Euclidean distance characteristics of this equivalent modulation and the consequently performance of stream modulation on this channel are explored in the homework (see Problem 13.21)

The important point that results from Forney equivalence is that from this point forward OMWM and stream modulations on frequency selective channels can be treated as a single entity. The optimum demodulation structures and the performance analysis will all have the same form. This unified look at low complexity demodulation and at find the union bound to the performance will be addressed in the following chapter. There is a unified nature in the tools needed by a communication engineer and this Forney equivalence is further evidence of this unified theme. Once a communication engineer acquires the feel for this unified framework they can comfortably address problems across a broad range of topics.

13.6 Conclusions

This chapter introduced the idea of orthogonal modulations with memory (OMWM). OMWM are used in practice to get better performance than orthogonal modulations or to change the spectral characteristics of orthogonal modulations while still maintaining a demodulation complexity that is $O(K_b)$. The demodulation complexity of OMWM was not addressed in this chapter but will be addressed in the next

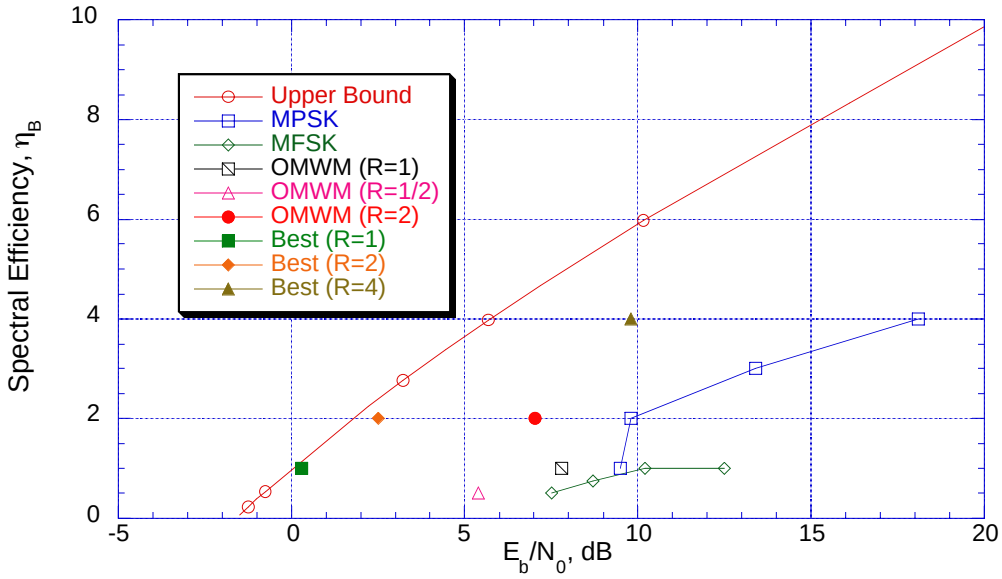


Figure 13.19: A comparison between the modulations considered in this book.

chapter in detail. OMWM can be implemented in a wide variety of rates and complexities. This variety of rates and complexity was illustrated with a handful of simple examples. For example Fig. 13.19 is a plot of the achieved spectral efficiency for the memoryless modulations and the OMWM that were considered in this book. Clearly these simple examples of OMWM have moved the achieved performance closer to that predicted by Shannon than were achieved by memoryless orthogonal modulations. These simple examples, while illustrative, do not really reflect the sophistication that is currently the state of the art in digital communications. Also included in Fig. 13.19 are the operating points of several of the best OMWM that have appeared in the literature. The discussion of these modulations and the methods of demodulation requires a level of sophistication that is not appropriate for where this text is at in the development of digital communication theory, so we will not go into more details. Interested readers should take a course or read a book on modern error correcting codes [Wic95, LC04, BDMS91]. The important point is that Shannon's upperbound is one that is achievable in modern communications. The next couple chapters will show that OMWM can have a demodulation complexity that is $O(K_b)$ as desired. Finally this chapter proved the equivalence of OMWM and stream modulations on frequency selective channels so that both of these situations can be viewed with a common framework. The use of this framework will begin with the discussion in the next chapter.

13.7 Homework Problems

Problem 13.1. Consider a $R = 1$ OMWM consisting of a four state trellis code (N_s=4) using 4PAM modulation (M_s=4) as proposed by Ho, Cavers and Varaldi [HCV93] and detailed in Section 13.2.2. The modulation is defined as

$\sigma(l+1) = g_1\left(I(l), \sigma(l)\right)$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	1	2
2	3	4
3	1	2
4	3	4

$J(l) = g_2\left(I(l), \sigma(l)\right)$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	0	2
2	3	1
3	2	0
4	1	3

$\tilde{D}_z(l) = a(J(l))$

$J(l)$	$D_z(l)$
0	$-3/\sqrt{5}$
1	$-1/\sqrt{5}$
2	$1/\sqrt{5}$
3	$3/\sqrt{5}$

a) For $K_b = 4$ and the termination back to $\sigma(7) = 1$ as discussed in the text ($N_f = 6$) find the squared Euclidean distance spectrum of the code and plot the union bound to the probability of word error.

b) Draw two transmitted paths and the corresponding error paths through the trellis associated with the minimum squared Euclidean distance error event.

c) For $K_b = 4$ and no trellis termination (i.e., no return to a common final state, $N_f = K_b$) find the squared Euclidean distance spectrum of the code and plot the union bound to the probability of word error.

d) Draw two transmitted paths and the corresponding error paths through the trellis associated with the minimum squared Euclidean distance error event.

Problem 13.2. The following is a commonly used $R = 1$ OMWM with four states ($N_s = 4$) and QPSK modulation ($M_s = 4$):

$\sigma(l+1) = g_1\left(I(l), \sigma(l)\right)$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	1	2
2	3	4
3	1	2
4	3	4

$\tilde{D}_z(l) = g_2\left(I(l), \sigma(l)\right)$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	d_0	d_2
2	d_1	d_3
3	d_2	d_0
4	d_3	d_1

where $d_0 = 1$, $d_1 = j$, $d_2 = -1$, and $d_3 = -j$.

a) Starting in $\sigma(1) = 1$ with $\vec{I} = [1\ 1\ 0\ 1]$ give the values of $I(5)$ and $I(6)$ such that $\sigma(7) = 1$ (for $N_f = 6$). Draw the trellis and show the path through the trellis that corresponds to this transmitted word.

b) Find the squared Euclidean distance spectrum of this code for $K_b = 4$ with termination to $\sigma(7) = 1$.

c) With a transmitted waveform produced by $\vec{I} = [1\ 1\ 0\ 1]$ enumerate out the three unique smallest non-zero squared Euclidean distances to the transmitted waveform in the conditional union bound and show a corresponding path through the trellis.

d) Is this code better or worse in terms of resulting performance than the HCV code used as an example in the text?

Problem 13.3. Consider the $R = 1$ OMWM given by

$\sigma(l+1) = g_1\left(I(l), \sigma(l)\right)$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	1	2
2	3	4
3	1	2
4	3	4

$\tilde{D}_z(l) = g_2\left(I(l), \sigma(l)\right)$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	d_0	d_2
2	d_1	d_3
3	d_2	d_0
4	d_3	d_1

Design the symbol mapping $\tilde{D}_z(l)$ (i.e., give values of d_0, d_1, d_2, d_3) to give better performance than the code presented in the previous problem while leaving all constellation points on the unit circle. Assume the code is terminated back to state 1. One example is sufficient to demonstrate you understand the concepts and there is no need to find the optimum mapping to get full credit.

Problem 13.4. A concept that is useful in the analysis of coded modulations is geometric uniformity [For91]. A code is said to be geometrically uniform if the conditional squared Euclidean distance spectrum for each of the possible transmitted codes words is identical. Since the conditional squared Euclidean distance spectrum of a code is identical the union bound can be computed by only considering the conditional squared Euclidean distance spectrum for one possible path in the trellis. The concept of geometric uniformity allows one to reduce the terms needed to be computed for the union bound from $(M-1)M$ to $M-1$ where again $M = 2^{K_b}$. Consider the following two trellis codes

$\sigma(l+1) = g_1\left(I(l), \sigma(l)\right)$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	1	2
2	3	4
3	1	2
4	3	4

$J(l) = g_2\left(I(l), \sigma(l)\right)$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	0	2
2	3	1
3	2	0
4	1	3

$\tilde{D}_z(l) = a(J(l))$

$J(l)$	$\tilde{D}_z^{(1)}(l)$	$\tilde{D}_z^{(2)}(l)$
0	$-3/\sqrt{5}$	1
1	$-1/\sqrt{5}$	j
2	$1/\sqrt{5}$	-1
3	$3/\sqrt{5}$	$-j$

Consider the case of K_b=4 with termination back to $\sigma(7) = 1$, are either of these codes geometrically uniform?

Problem 13.5. Consider the following $R = 1$ OMWM with seven states ($N_s = 7$) and 4-PAM modulation ($M_s = 4$):

$\sigma(l+1) = g_1\left(I(l), \sigma(l)\right)$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	2	4
2	3	5
3	4	6
4	1	7
5	2	4
6	3	5
7	4	6

$J(l) = g_2\left(I(l), \sigma(l)\right)$

State, $\sigma(l)$	$I(l)$ = 0	$I(l)$ = 1
1	1	0
2	1	0
3	1	0
4	3	0
5	3	2
6	3	2
7	3	2

$\tilde{D}_z(l) = a(J(l))$

$J(l)$	$\tilde{D}_z(l)$
0	$-3/\sqrt{5}$
1	$-1/\sqrt{5}$
2	$1/\sqrt{5}$
3	$3/\sqrt{5}$

a) Starting in $\sigma(1) = 4$ and that we want to return to $\sigma = 4$, show that if K_b is odd then this is possible in one transition and if K_b is even then this is possible in two transitions.

b) For $K_b = 3$ find the squared Euclidean distance spectrum and plot the union bound. Will this modulation perform better or worse than BPSK?

c) For $K_b = 3$ when the modulation is used in linear stream modulation, show that

$$\int_{-\infty}^{\infty} x_i(t)dt = 0. \tag{13.57}$$

d) For $K_b = 3$ when the modulation is used in linear stream modulation, find and plot $D_{x_z}(f)$? How do you see the characteristic of c) manifested in this plot?

Problem 13.6. Consider the channel impulse response

$$h_z(t) = h_0\delta(t) + h_1\delta(t - 0.25T) \tag{13.58}$$

and a BPSK stream modulation with a square root raised cosine pulse, $u(t)$, with excess bandwidth 50% and $T_u = 6T$. Compute the causal whitening filter $W\left(e^{j2\pi f}\right)$ for $h_0 = 0.8$ and $h_1 = 0.6$ and the impulse response and the equivalent modulation with memory.

a) How many states will be in the modulation with memory?

b) Often the optimal demodulator complexity is too high and some approximations need to be made. By making an engineering judgment shorten either the whitening filter or impulse response that defines the modulation with memory that is decoded (note this will result in a mismatch between the channel output and what is assumed for demodulation). How will your choices effect the performance and complexity of the demodulator?

Problem 13.7. Recall a memoryless binary linear stream modulation has the form

$$X_z(t) = \sum_{l=1}^{K_b} D_z(l)\tilde{u}(t - (l-1)T) \tag{13.59}$$

where $D_z(l) = a(I(l))$ is the modulation symbol and $\tilde{u}(t)$ is the pulse shape. For the case where BPSK modulation is used and

$$V_{\tilde{u}}(nT) = \begin{cases} 1 & n = 0 \\ 0.5 & n = \pm 1 \\ 0 & \text{elsewhere.} \end{cases} \tag{13.60}$$

The equivalent OMWM for an orthogonal stream modulation in a frequency selective channel as discussed in text has the form given in Fig. 13.20 where $\tilde{Q}(k)$ is the output of the filter matched to $\tilde{u}(t)$ and $Q(k) = \tilde{D}_z(k) + N(k)$ where $N(k)$ is a white noise and

$$\tilde{D}_z(k) = \sum_{l=0}^{N_u} D_z(k-m)f_1^+(m) \tag{13.61}$$

is the equivalent modulation with memory.

a) Identify $f_1^+(m),\ m \in \{0, \ldots, N_u\}$.

b) Detail the equivalent modulation with memory. Provide either the trellis diagram or the functions $g_1(I(l), \sigma(l))$ and $g_2(I(l), \sigma(l))$.

c) If $K_b = 4$ and $\vec{I} = [0\ 0\ 1\ 1]$ give $\tilde{D}_z(l)$ $l \in \{1, \ldots, K_b + N_u\}$

d) Compute the conditional union bound to the probability of word error given that $\vec{I} = [0\ 0\ 1\ 1]$ was the transmitted word.

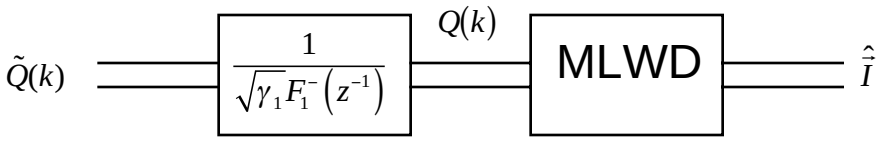


Figure 13.20: The whitening filter approach to optimal demodulation of non-Nyquist linear modulations.

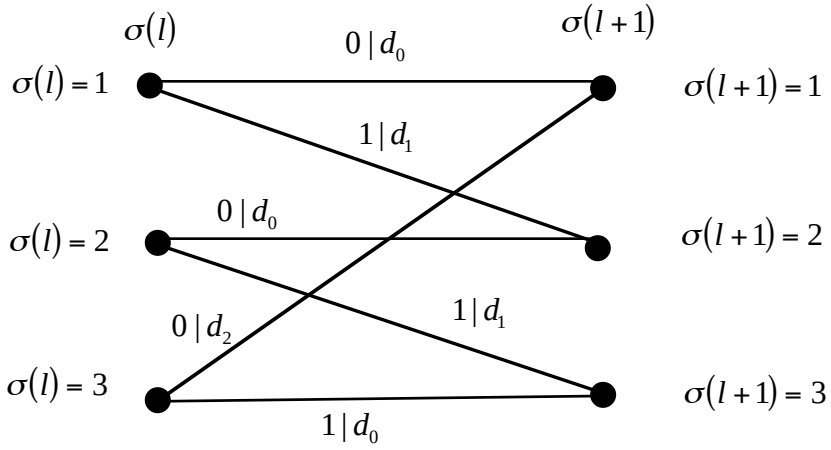


Figure 13.21: A trellis diagram for a modulation with memory.

e) Given the result in d) how does the word error performance of a system where Nyquist criterion can be maintained compare to the word error performance of one where Nyquist criterion cannot be maintained when optimal demodulation is implemented.

Problem 13.8. In Fig. 13.20, recall that $\tilde{Q}(k) = \tilde{S}(k) + \tilde{N}(k)$ and that $Q(k) = S(k) + N(k)$. Show that

$$R_N(m) = E\left[N(k)N^*(k-m)\right] = N_0\delta_m \tag{13.62}$$

or that the filter obtained from a spectral factorization of $V_{\tilde{u}}(mT)$ is also a filter that whitens the noise.

Problem 13.9. An OMWM is defined by the trellis in Fig. 13.21. Assume that $d_0 = 1$, $d_1 = \exp[j2\pi/3]$, $d_2 = \exp[-j2\pi/3]$ and $\sigma(1) = 1$.

a) For an input arbitrary $K_b = 3$ bit sequence and if the modulation has to stay as defined by the trellis how many symbols will be needed to guaranteed a return to $\sigma = 1$.

b) Will this MWM give better performance than memoryless BPSK modulation when $K_b = 3$ when terminated as in part a)? Plot the union bound

c) Is this OMWM geometrically uniform (see Problem 13.4 for the definition of geometrically uniform)?

d) It is possible to define a one symbol termination (i.e., $N_f = K_b+1$), which might not be as defined by the trellis, that would result in a minimum distance that is the same as if the modulation had stayed on the trellis and had $N_f > K_b + 1$. Give what the final symbol should be as a function of $\sigma(K_b + 1)$ to achieve this goal. Plot a union bound and compare it with the performance when $N_f > K_b + 1$ and the modulation stays on the trellis.

Problem 13.10. Your company, Dakota Instruments, is building a communication system that uses a BPSK stream modulation. A particular client wants you to use it on a frequency selective channel. A colleague of yours has characterized the channel and she provided you an unnormalized discrete time equivalent white noise channel model given as

$$\sqrt{\gamma_1}F_1^+(e^{j2\pi f}) = 1 - e^{-j2\pi f} + 0.25e^{-j4\pi f} \tag{13.63}$$

You will be transmitting a large number of bits (K_b is large).

a) Find the energy of the equivalent pulse shape, $\tilde{u}(t)$.

b) Assume no noise is present and for a k in the middle of the frame, give all possible values of $Q(k)$. How many states will be in the equivalent modulation with memory?

c) Assume no noise is present and for a k in the middle of the frame, give all possible values of $\tilde{Q}(k)$.

Problem 13.11. If AMI modulation is used as an OMWM in an OFDM system, what characteristic would it produce? Would you consider AMI useful in any way as an OMWM for an OFDM system. *Hint: AMI puts a DC notch in a stream modulation's spectrum.*

Problem 13.12. Consider an OMWM consisting of a four state trellis code (N_s=4) using 4PAM modulation (M_s=4) as proposed by Ho, Cavers and Varaldi [HCV93] and detailed in Section 13.2.2. Assume this modulation is used in a stream modulation with K_b large and

$$u_r(t) = \begin{cases} \sqrt{\dfrac{1}{T}} & 0 \le t \le T \\ 0 & \text{elsewhere.} \end{cases} \tag{13.64}$$

a) Enumerate out the edges, $s(l)$, and the mappings to $\tilde{D}_z(l)$.

b) Identify the edge transition matrix, S_T.

c) Compute the average energy spectrum per bit, $D_{X_z}(f)$, of the resultant stream modulation.

Problem 13.13. Consider a OMWM consisting of a two state trellis code (N_s=2) using 3PAM modulation (M_s=3) denoted as AMI modulation in the text. The modulation is defined as

$\sigma(l+1) = g_1\left(I(l), \sigma(l)\right)$

	$I(l)$	
State, $\sigma(l)$	0	1
1	1	2
2	1	2

$J(l) = g_2\left(I(l), \sigma(l)\right)$

	$I(l)$	
State, $\sigma(l)$	0	1
1	1	0
2	2	1

$\tilde{D}_z(l) = a(J(l))$

$J(l)$	$D_z(l)$
0	-2
1	0
2	2

a) For $K_b = 4$ and the termination back to $\sigma(6) = 1$ as discussed in the text ($N_f = 5$) find the squared Euclidean distance spectrum of the code and plot the union bound to the probability of word error.

b) Draw two pairs of transmitted paths and the corresponding error paths through the trellis associated with the minimum squared Euclidean distance error event.

c) For $K_b = 4$ and no trellis termination (i.e., no return to a common final state, $N_f = K_b$) find the squared Euclidean distance spectrum of the code and plot the union bound to the probability of word error.

d) Draw two pairs of transmitted paths and the corresponding error paths through the trellis associated with the minimum squared Euclidean distance error event.

Problem 13.14. In Problem 13.5 a OMWM is presented that at least for small K_b was shown to produce a notch at DC in the transmitted spectrum. Using the results of Section 13.3.2 find the average energy spectrum per bit for K_b large.

Problem 13.15. Consider a OMWM consisting of a eight state convolutional code (N_s=8) using BPSK modulation (M_s=2) and detailed in Section 13.4.2. The modulation is defined in Table 13.4.

a) For $K_b = 4$ and the termination back to $\sigma(8) = 1$ as discussed in the text ($N_f = 14$) find the squared Euclidean distance spectrum of the code and plot the union bound to the probability of word error.

b) Draw two transmitted paths and the corresponding error paths through the trellis associated with the minimum squared Euclidean distance error event.

c) Is this modulation geometrically uniform? (See Problem 13.4)

Problem 13.16. Consider an OMWM consisting of an eight state convolutional code as detailed in Table 13.4 and discussed in Section 13.4.2. Assume this modulation is used in a stream modulation with K_b large and

$$u_r(t) = \begin{cases} \sqrt{\frac{1}{T}} & 0 \le t \le T \\ 0 & \text{elsewhere.} \end{cases} \tag{13.65}$$

a) Enumerate out the edges, $s(l)$, and the mappings to $\tilde{D}_z(l)$.

b) Identify the edge transition matrix, S_T.

c) Compute the average energy spectrum per bit, $D_{X_z}(f)$, of the resultant stream modulation.

Problem 13.17. Consider an OMWM consisting of a four state trellis coded modulation (N_s=4) using 8PSK modulation (M_s=8) and detailed in Section 13.4.3. The modulation is defined in Table 13.5.

a) For $K_b = 4$ and the termination back to $\sigma(7) = 1$ as discussed in the text ($N_f = 6$) find the squared Euclidean distance spectrum of the code and plot the union bound to the probability of word error.

b) Draw two transmitted paths and the corresponding error paths through the trellis associated with the minimum squared Euclidean distance error event.

c) Is this modulation geometrically uniform? (See Problem 13.4)

Problem 13.18. When the information bits are all equally likely and independent, prove that if

$$E\left[D_z\left((m-1)N_m + j\right)|\sigma(m) = i\right] = 0 \qquad j \in \{1, \ldots, N_m\}, i \in \{1, \ldots, N_s\} \tag{13.66}$$

then the transmitted energy spectrum per bit of an OMWM using stream modulation is

$$D_{x_z}(f) = E_b \left|U(f)\right|^2. \tag{13.67}$$

In other words the spectrum is unchanged by the OMWM compared to a standard linear stream modulation [BDMS91, Big86]. Which of the example OMWM in this chapter satisfy this condition?

Problem 13.19. Consider an OMWM consisting of a four state (N_s=4) OMWM using BPSK modulation (M_s=2) and detailed in Section 13.4.4. The modulation is defined in Table 13.6.

a) For $K_b = 4$ and no termination ($N_f = 8$) find the squared Euclidean distance spectrum of the code and plot the union bound to the probability of word error.

b) Draw two transmitted paths and the corresponding error paths through the trellis associated with the minimum squared Euclidean distance error event.

c) Is this modulation geometrically uniform? (See Problem 13.4)

Problem 13.20. Consider the situation in Example 13.3 where a linear stream modulation is used on a frequency selective channel. If QPSK modulation, $\Omega_d = \{1+j, 1-j, -1+j, -1-j\}$, is used (in place of BPSK) to send two bits per symbol find the Forney equivalent OMWM. Specify how many states are in the OMWM and the effective symbol constellation points.

Problem 13.21. Consider the situation in Example 13.3 where a linear stream modulation (BPSK) is used on a frequency selective channel. Find the union bound for $K_b = 4$ to the performance of the MLWD for this example. Plot this performance and compare to the performance of BPSK on a frequency flat channel. Is the Forney equivalent OMWM geometrically uniform? (See Problem 13.4)

Problem 13.22. In the example that considered an $R = 2$ Ungerboeck trellis code in Section 13.4.3 it was stated that an extra two bits can be transmitted during the termination stages.

a) For the example detailed in Section 13.4.3 show the $N_f = 4$ trellis where $K_b = 6$ and $\sigma(5) = 1$.

b) Has the minimum squared Euclidean distance changed by adding these two extra bits?

c) Compute the union bound of this new case. Plot this union bound along with the union bound of the $K_b = 4$ case. Justify the difference.

Problem 13.23. Forney equivalence was shown in the text for a memoryless stream modulation on a frequency selective channel. A similar equivalent modulation can be formed if a stream MWM is used on a frequency selective channel. The trellis structure of the equivalent modulation is a concatenation of the trells for the OMWM and the trellis resulting from the frequency selective channel. The resultant trellis is often denoted the supertrellis in the literature. For the frequency selective channel given in Example 13.3 and the $R = 1$ modulation detailed in Section 13.2.2

a) Show the equivalent modulation with memory has $N_s = 8$.

b) Identify all the elements of the equivalent constellation?

c) Give tables with $\sigma(k+1) = g_1\left(I(k), \sigma(k)\right)$ and $J(k) = g_2\left(I(k), \sigma(k)\right)$ for the concatenated equivalent OMWM.

13.8 Projects

Not completed this edition

Chapter 14

Maximum Likelihood Sequence Demodulation

Up to this point in this text we have introduced three general methods to communicate K_b bits on both frequency flat channels and frequency selective channels.

1. General M-ary modulations
   - The advantage of a general M-ary modulation is that it can achieve very good performance and arbitrary spectral efficiency.
   - The disadvantage of a general M-ary modulation is that without more structure the optimal demodulator has complexity $O(2^{K_b})$.
2. Orthogonal memoryless modulations (including stream modulations, OFDM, OCDM)
   - The advantage of orthogonal modulation is that the optimum demodulator has complexity $O(K_b)$ and a desired spectral efficiency can be achieved with a proper design of the modulation signals.
   - The disadvantage is that the performance is limited to that achievable with a single symbol transmission.
   - Frequency selective channels cause orthogonality to be lost. A variety of algorithms were introduced to address optimal and sub-optimal demodulation of memoryless orthogonal modulations in frequency selective channels. Performance on frequency selective channels is always lower bounded by the performance on frequency flat channels.
3. Orthogonal Modulations with Memory
   - The advantage of orthogonal modulations with memory (OMWM) is that a desired spectral efficiency and a performance arbitrarily close to Shannon's limit can be achieved with a proper design of the modulation signals.

The goal for the next two chapters will be to explore the remaining areas of tradeoff in performance, complexity, and spectral efficiency, i.e.,

- Showing that OMWM can be demodulated with a complexity that is $O(K_b)$. Since OMWM can achieve performance close to the Shannon bound, achieving a complexity $O(K_b)$ for OMWM demodulation is the "Holy Grail" of communication theory and this book can claim to have completed an introduction to communication theory satisfactorily.

This chapter will be a brief introduction into the optimum word demodulation of OMWM and will derive a key algorithm that a communication engineer should understand; the Viterbi Algorithm (VA) [Vit67, For73].

The trellis structure of an OMWM will allow the complexity of optimum demodulation to be reduced from $O(2^{K_b})$, as would be required for a general modulation, to $O(K_b)$. This linear demodulation complexity with the significantly improved performance or a desired spectral characteristics are the goals for which communications engineers strive. The formulation and understanding of the form of the low complexity optimum demodulator for OMWM is one of the most important tools for a communication engineer. This chapter explores the MLWD structures and the next chapter explores the MAPBD structures. The development in these two chapters explores the case of OMWM when $R = 1$ in detail and the generalization to other R will be explore in the homework.

These optimum demodulation algorithms can best be understood by introducing a graphical representation of the modulation with memory. A good reference that explores this idea is [Fre98]. A graphical model of OMWM is shown in Fig. 14.1. In a modulation with memory there are random variable that represent this modulation and the observations associated with the demodulation of this modulation with memory. These random variables are associated with the circles in Fig. 14.1. Many of the random variables in the modulation with memory are a result of functions of other random variables in the modulation with memory. These functions are represented by letters. In our discussion deterministic functions will be represented by $g(\bullet)$ and random functions (those containing auxiliary randoms variables) will be denoted by $f(\bullet)$. The equations governing the modulation finite state machine are examples of the $g(\bullet)$ type functions and mapping from constellation selection signal, $J(m)$, to the matched filter output, $Q(k)$, is an example of the $f(\bullet)$ type function where the auxiliary random variable in this case is the channel noise. The goal in optimum demodulator algorithms is to use the observations, $Q(k)$, to make optimum estimates about the state of the OMWM, typically the values of $I(m)$. By examining graphical models for digital communication a great deal of commonality can be exploited in examining optimal demodulation structures for OMWM.

It is apparent from examining Fig. 14.1 that an OMWM only has local connections/relations between the random variables in the demodulation. This local connectivity that is defined by the finite state machine of the OMWM is what enables a simplified MLWD. Finding MLWD algorithm that has a complexity $O(K_b)$ and deriving analysis methods also of reasonable complexity that predict the performance of MLWD is the focus of this chapter

14.1 Maximum Likelihood Sequence Demodulation

A brute force demodulation of an OMWM would still have a complexity of $O(2^{K_b})$. Recall the MLWD for an OMWM is given as

$$\begin{aligned}
\hat{\vec{I}} &= \arg \max_{i \in \{0,\ldots,M-1\}} T_i \qquad (14.1)\\
&= \arg \max_{i \in \{0,\ldots,M-1\}} \sqrt{E_b} \sum_{k=1}^{N_f} \Re\left[\tilde{d}_i^*(k) Q(k)\right] - \frac{E_b}{2} \sum_{k=1}^{N_f} \left|\tilde{d}_i(k)\right|^2 \\
&= \arg \min_{i \in \{0,\ldots,M-1\}} \sum_{k=1}^{N_f} \left|Q(k) - \sqrt{E_b}\tilde{d}_i(k)\right|^2 \\
&= \arg \max_{i \in \{0,\ldots,M-1\}} P_{\vec{I}}\left(i \left| \vec{Q}\right.\right)
\end{aligned}$$

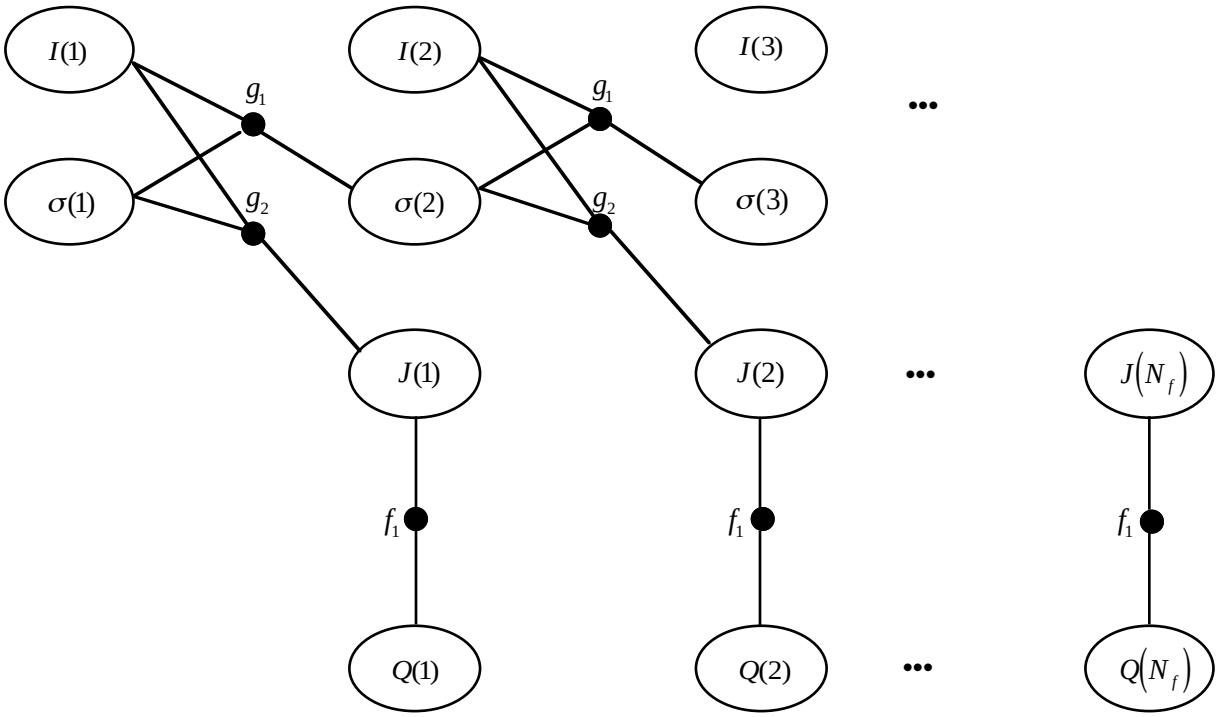


Figure 14.1: A graphical model of an OMWM and demodulator sufficient statistics for the case $R = 1$.

where $Q(k)$ is the matched filter ouptut sample for an OMWM. Further understanding of this demodulator can be gained by restating the optimum demodulator to give

$$\hat{\vec{I}} = \arg \min_{i \in \{0,\ldots,M-1\}} \Delta_E \left(\vec{Q}, \sqrt{E_b} \vec{\tilde{d}}_i \right) \tag{14.2}$$

where $\Delta_E \left(\vec{Q}, \sqrt{E_b} \vec{\tilde{d}}_i \right)$ is the squared Euclidean distance between the vector of matched filter outputs, $\vec{Q}$, and the vector of possible noiseless signals that could be produced by the OMWM, $\sqrt{E_b} \vec{\tilde{d}}_i$. Evaluating this squared Euclidean distance over all possible 2^{K_b} possible signals produced by the OMWM without considering the trellis structure would require complexity $O(2^{K_b})$. The discussion of computational complexity in this text is intended for clarity of understanding. The complexity measure of an algorithm will be measured in operations (e.g., additions, multiplications and compares) and this counting is not always a fair comparison in all implementations. There are also many ways to improve the complexity of optimum demodulation algorithms and not all will be discussed in this text book since the goal is a clarity of understanding of the algorithms.

Example 14.1: Consider again the 4-PAM code of Section 13.2.2 with $K_b = 4$ and $N_f = 6$. The trellis diagram for this transmission is shown in Fig. 14.2. The computations needed to be done to implement the MLWD can be summarized with

1. Compute the $N_f = 6$ matched filter outputs, $Q(k)$, $k \in \{1, \ldots, N_f\}$. This particular computation is dependent on the choice for the orthogonal modulation so we will denote this computation as taking N_{QA} additions and N_{QM} multiplications.

2. For each postulated transmitted sequence, $i \in \{0, M-1\}$, the MLWD must form

$$T_i = \sum_{k=1}^{N_f} T_i(k) = -\Delta_E \left(\vec{Q}, \sqrt{E_b}\vec{\tilde{d}}_i \right). \tag{14.3}$$

   This requires for each time instance, k, that

$$T_i(k) = -\left| Q(k) - \sqrt{E_b}\tilde{d}_i(k) \right|^2 \tag{14.4}$$

   be computed for each $i \in \{0, M-1\}$. This symbol–wise squared Euclidean distance metric requires 3 real additions and 2 real multiplications. Consequently for each possible transmitted sequence, $i \in \left\{0, 2^{K_b} - 1\right\}$ a total of $6 + 6 \times 3 = 24$ additions and $2 \times 6 = 12$ multiplies are needed to compute the maximum likelihood metric for the sequence.

To compute all the maximum likelihood metrics leads to a complexity of $16 \times 24 + 6 \times N_{QA} = 384 + 6 \times N_{QA}$ additions and $16 \times 12 + 6 \times N_{QM} = 192 + 6 \times N_{QM}$ multiplies. Finally finding the largest likelihood (smallest Euclidean square distance) requires a total of 15 comparison operations. The important point to note here is that all computations (adds, multiplies, and compares) all have a multiplicative factor that is proportional to $M = 2^{K_b}$.

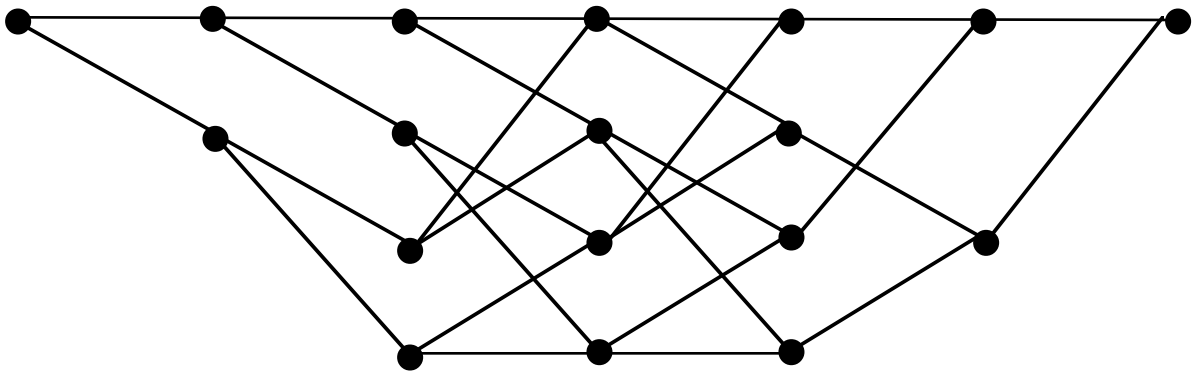


Figure 14.2: A trellis diagram for a modulation with memory having K_b=4, $N_f = 6$, and N_s=4.

The simplest form of the MLWD for the OMWM was first found by Viterbi [Vit67] and has been denoted maximum likelihood sequence demodulation (MLSD) or more often the Viterbi Algorithm (VA). Forney recognized this common structure in the stream modulation in a frequency selective channel [For72]. An interesting tutorial article on the Viterbi algorithm is given in [For73]. It should be noted that the Viterbi algorithm is so common in modern communication systems that most processors that

are used in communication applications often have special purpose hardware accelerators to specifically implement certain important operations common in the Viterbi algorithm. Also Viterbi algorithm is a special case of a more general tool called dynamic programming [Bel57]. It is well accepted that understanding the Viterbi algorithm is a fundamental step in a communication system engineer's educational process. Because of this reason this text provides both a heuristic overview and a formal derivation in the sequel.

14.1.1 The Viterbi Algorithm: Enabling Observations

The basic idea in MLWD is that given the trellis structure of an OMWM the maximum likelihood demodulated word correspnds to the path through the trellis that has the smallest Euclidean square distance. The reduced complexity form for MLSD was predicated on two observations:

1. The computations associated with the edges of the trellis are common to many sequences.

2. The 2^{K_b} competing paths through the trellis often have a great deal of commonality.

The number of edges will determine the number of Euclidean square distance computations that have to be performed on the MLWD. Using the notation from Chapter 13 we can denote the edges at time l with $S(l)$ (see Definition 13.1). There is a mapping from edges to symbols transmitted, hence if we identify the number of edges in the trellis we have upper bounded the number of computations of the form

$$T_i(k) = -\left|Q(k) - \sqrt{E_b}\tilde{d}_i(k)\right|^2 \tag{14.5}$$

that must be performed. Again for simplicity of discussion we will consider only the $R = 1$ OMWM and the generalization is straightforward. The number of edges in a trellis for a $R = 1$ OMWM is

$$K_e = 2N_sK_b - K_t \tag{14.6}$$

where K_t is an integer depending on the particular trellis. Consequently the number of squared Euclidean distance computations that is needed in the MLWD only grows linearly with the number of bits to be transmitted, K_b.

Example 14.2: Returning to the 4 state OMWM that is described with the trellis in Fig. 14.2. There is a total of 28 edges in this case ($K_b = 4$, $N_s = 4$, and $K_t = 4$). All of the needed squared Euclidean distance computations can be accomplished with $28 \times 3 = 84$ additions and $28 \times 2 = 56$ multiplications. These 28 edge associated Euclidean square distances need to be added up for each of the possible $2^{K_b} = 16$ possible paths of length $N_f = 6$ through the trellis. The total complexity of this form of the MLWD that only computes one squared Euclidean distance per edge is $84 + 16 \times 6 + 6 \times N_{QA} = 180 + 6 \times N_{QA}$ additions and $56 + 6 \times N_{QM}$ multiplications. Compare this with the results from Example 14.1 and a significant savings is observed by only making each edge computation once.

Using the common trellis paths followed by many data sequences can help reduce the amount of processing further. Note this idea was explored in the Ungerboeck MLWD for stream modulation in the frequency selective channel that was introduced in Section 12.2. The key idea here is that at any point if two sequences have a common future path through the trellis then there is no loss in eliminating the one with the lower likelihood metric accumulated up to that time.

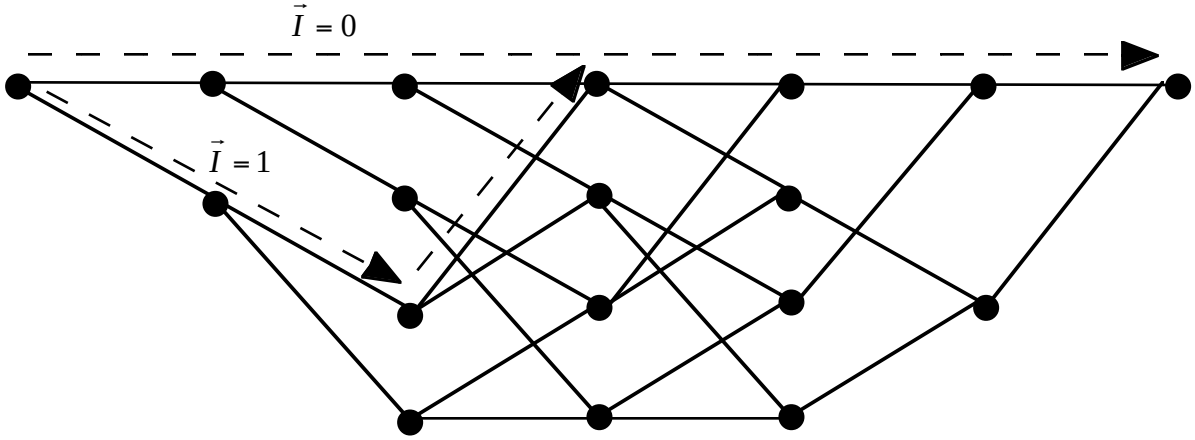


Figure 14.3: The two paths through the trellis denoted by $\vec{I} = 0$ and $\vec{I} = 1$.

Example 14.3: Consider again the 4-PAM code of Section 13.2.2 with $K_b = 4$ and $N_f = 6$. Two sequences in this possible OMWM are denoted

$$\vec{I} = 0 = [0\ 0\ 0\ 0] \qquad \vec{I} = 1 = [1\ 0\ 0\ 0]\,. \tag{14.7}$$

These two input words produce

$$\vec{\tilde{d}}_0 = \left[-3/\sqrt{5}\ \ -3/\sqrt{5}\ \ -3/\sqrt{5}\ \ -3/\sqrt{5}\ \ -3/\sqrt{5}\ \ -3/\sqrt{5}\right] \tag{14.8}$$

and

$$\vec{\tilde{d}}_1 = \left[1/\sqrt{5}\ 3/\sqrt{5}\ 1/\sqrt{5}\ \ -3/\sqrt{5}\ \ -3/\sqrt{5}\ \ -3/\sqrt{5}\right]. \tag{14.9}$$

The ML metric for these two words can be divided into parts corresponding to

$$T_0 = -\sum_{k=1}^{3}\left|Q(k) + \frac{3\sqrt{E_b}}{\sqrt{5}}\right|^2 - \sum_{k=4}^{6}\left|Q(k) + \frac{3\sqrt{E_b}}{\sqrt{5}}\right|^2 \tag{14.10}$$

and

$$T_1 = -\left|Q(1) - \frac{\sqrt{E_b}}{\sqrt{5}}\right|^2 - \left|Q(2) - \frac{3\sqrt{E_b}}{\sqrt{5}}\right|^2 - \left|Q(3) - \frac{\sqrt{E_b}}{\sqrt{5}}\right|^2 - \sum_{k=4}^{6}\left|Q(k) + \frac{3\sqrt{E_b}}{\sqrt{5}}\right|^2. \tag{14.11}$$

The first three of the terms in the six term summation are different for each of the sequences. The last three terms of the summation are identical and hence will make no difference in whether $T_0 > T_1$. If we examine the trellis diagram in Fig 14.3 it is pretty clear that the path followed through the trellis for each of these information words differs only in the first three trellis sections and is the same in the last three trellis sections. Consequently a choice between whether $\vec{I} = 0$ or $\vec{I} = 1$ will have a larger ML metric (smaller squared Euclidean distance) can be made after the third matched filter output is computed.

Decisions between sequences can be made anytime future paths of the competing sequences would be

the same. This implies at each time in the trellis only N_s paths need to be maintained (one for each state). Since the paths leaving a state will have a common future, only the path with the largest ML metric coming into a state can end up being the ML path. Consequently at each state with multiple incoming paths there will be a sequence with the largest ML metric. Only this metric and the path through the trellis corresponding to that "winning" metric need be saved for future processing. Since only $N_s \ll 2^{K_b}$ paths need to be maintained at each state and the decision process is distributed across the whole frame the complexity can be significantly reduced. The decision processing now has a complexity linear in the number of bits transmitted (i.e., $N_s K_b$) versus the exponential complexity (2^{K_b}) that would be needed if the modulation did not have the trellis structure.

Example 14.4: Consider again the 4-PAM code of Section 13.2.2 with $K_b = 4$ and $N_f = 6$. Points where paths join in the trellis are represented by numbers in the trellis diagram of Fig. 14.4. The first time decision are made (points 1-4) eliminates half of the possible paths in the trellis (8 paths will remain). The second time when decisions are made (points 5-8) will eliminate half of the remaining paths (4 will remain). The third time (points 9-10) eliminate all but two and the final decision can be made at point 11.

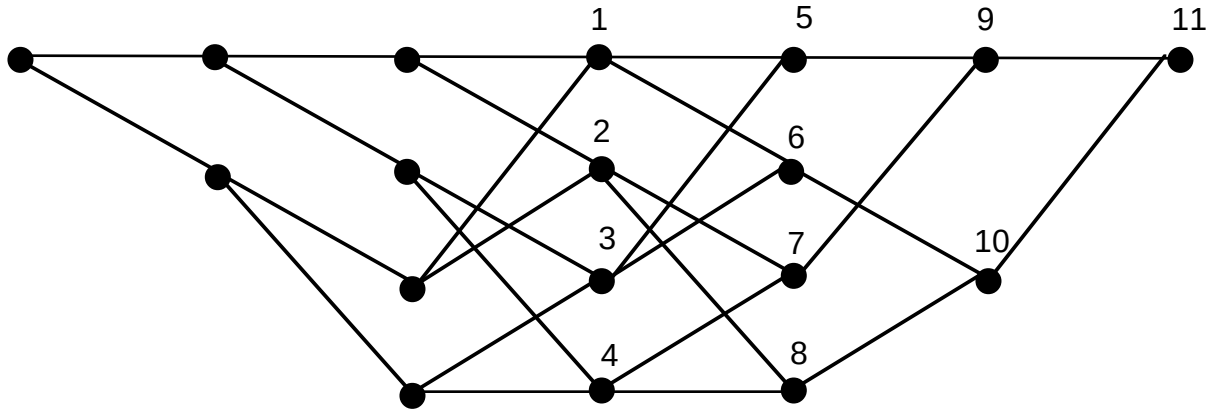


Figure 14.4: An enumeration of the decision nodes.

The Viterbi algorithm takes advantage of these two characteristics to produce a decoding complexity that is $O(K_b)$. Since all transmitted signals are defined via paths on the trellis the Euclidean squared distance associated with each edge only need to be computed once. Also since all codewords are defined on a trellis the common future of paths on a trellis can be used to eliminate a significant number of paths from possible contention as the maximum likelihood metric. The next subsection provides a formal derivation of the Viterbi algorithm to complement the more heuristic discussion in this subsection.

14.1.2 The Viterbi Algorithm: Formal Derivation

First a couple of definitions will aid the derivation.

Definition 14.1 *For any time series, $X(k)$ from time 1, $X(1)$, up to time k, $X(k)$ shall be denoted as $\vec{X}_p(k)$.*

Definition 14.2 *The set of all sequences, $\vec{I}_p(k-1)$ that end up in state $\sigma(k) = i_s$ is denoted $\Omega_{i_s}(k)$.*

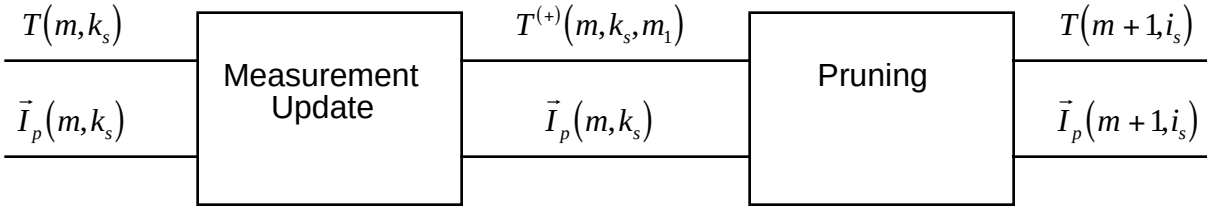


Figure 14.5: The block diagram of the MLSD recursion.

The Viterbi algorithm is a time recursive algorithm that computes the maximum likelihood path through a trellis that describes an OMWM. The one step recursion is an time update on the following two statistics

1. The best accumulated ML decision metric at time m for each state $\sigma(m) \in \{1, \ldots, N_s\}$. Denote

$$T(m, k_s) = \max_{i \in \Omega_{k_s}(m)} - \sum_{k=1}^{m-1} \left| Q(k) - \sqrt{E_b}\tilde{d}_i(k) \right|^2 \qquad k_s \in \{1, \ldots, N_s\} \tag{14.12}$$

2. The data sequence that achieves the best accumulated ML decision metric at time m for each state $\sigma(m) \in \{1, \ldots, N_s\}$. Here denote

$$\hat{\vec{I}}_p(m, k_s) = \arg \max_{i \in \Omega_{k_s}(m)} - \sum_{k=1}^{m-1} \left| Q(k) - \sqrt{E_b}\tilde{d}_i(k) \right|^2 \qquad k_s \in \{1, \ldots, N_s\} \tag{14.13}$$

Consequently the Viterbi algorithm takes inputs of $T(m, k_s)$ and $\hat{\vec{I}}_p(m, k_s)$ for $k_s \in \{1, \ldots, N_s\}$ and provides a way to compute $T(m+1, i_s)$ and $\hat{\vec{I}}_p(m+1, i_s)$ for $i_s \in \{1, \ldots, N_s\}$. The block diagram of this recursive algorithm is shown in Fig. 14.5.

The desired one-step outputs are

$$T(m+1, i_s) = \max_{i \in \Omega_{i_s}(m+1)} - \sum_{k=1}^{m} \left| Q(k) - \sqrt{E_b}\tilde{d}_i(k) \right|^2 \qquad i_s \in \{1, \ldots, N_s\} \tag{14.14}$$

$$\hat{\vec{I}}_p(m, i_s) = \arg \max_{i \in \Omega_{i_s}(m)} - \sum_{k=1}^{m-1} \left| Q(k) - \sqrt{E_b}\tilde{d}_i(k) \right|^2 \qquad i_s \in \{1, \ldots, N_s\} \tag{14.15}$$

An important notational formalism that facilitates a simple derivation will be a method to indicate valid state transitions (enumerate the edges in the trellis) and this formalism is captured in the following definition

Definition 14.3 *The function $g_3(i_s, k_s, m_s), i_s \in \{1, \ldots, N_s\}, k_s \in \{1, \ldots, N_s\}, m_s = 0, 1$ with a range of $\{0, 1\}$ will indicate which transitions in the trellis are allowable. Specifically $g_3(i_s, k_s, m) = 1$ if $\sigma(m) = k_s$ and $I(m) = m_s$ implies that $\sigma(m+1) = i_s$.*

Example 14.5: Consider again the 4-PAM code of Section 13.2.2. The enumeration of $g_3(\bullet,\bullet,\bullet)$ is given in the following table

$g_3(\sigma(m+1),\sigma(m),I(m)))$

		$I(m)$	
$\sigma(m)$	$\sigma(m+1)$	0	1
1	1	1	0
1	2	0	1
1	3	0	0
1	4	0	0
2	1	0	0
2	2	0	0
2	3	1	0
2	4	0	1
3	1	1	0
3	2	0	1
3	3	0	0
3	4	0	0
4	1	0	0
4	2	0	0
4	3	1	0
4	4	0	1

The function g_3 for this code indicates, for instance, that $\sigma(m+1)=1$ can only be arrived at from $\sigma(m)=1$ and $\sigma(m)=3$ and only for $I(m)=0$.

More specifically $\Omega_{i_s}(m+1)$ is the set of all sequences $\vec{I}_p(m)$ such that $g_3(i_s,\bullet,\bullet)=1$, i.e.,

$$\Omega_{i_s}(m+1)=\left\{\vec{I}_p(k):g_3\left(i_s,k_s,m_1\right)=1\right\}\qquad k_s\in\{1,\ldots,N_s\},m_1=0,1 \tag{14.16}$$

Now (14.14) can be restated as

$$\begin{aligned} T\left(m+1,i_s\right) &= \max_{i\in\Omega_{i_s}(m+1)} -\sum_{k=1}^{m}\left|Q(k)-\sqrt{E_b}\tilde{d}_i(k)\right|^2 \qquad i_s\in\{1,\ldots,N_s\} \\ &= \max_{\substack{k_s\in\{1,\ldots,N_s\}\\ i\in\Omega_{k_s}(m)\\ m_1=0,1}} \frac{-\left|Q(m)-\sqrt{E_b}\tilde{d}_{(k_s,m_1)}\right|^2}{g_3(i_s,k_s,m_1)} - \sum_{k=1}^{m-1}\left|Q(k)-\sqrt{E_b}\tilde{d}_i(k)\right|^2 \qquad i_s\in\{1,\ldots,N_s\}. \end{aligned} \tag{14.17}$$

It should be noted that dividing the term corresponding to the most recent possible edge in the summation by $g_3(\bullet)$ only ensures that the maximization is done over valid paths in the trellis. Noting that the second term of (14.17) is not a function of $I(m)=m_1$ and that the first term is only a function of

$\sigma(m) = k_s$ and $I(m) = m_1$, enables the following simplification

$$\begin{aligned} T(m+1, i_s) &= \max_{\substack{k_s \in \{1,\dots,N_s\} \\ i \in \Omega_{k_s}(m) \\ m_1 = 0,1}} \frac{-\left|Q(m) - \sqrt{E_b}\tilde{d}_{(k_s,m_1)}\right|^2}{g_3(i_s, k_s, m_1)} + \max_{i \in \Omega_{k_s}(m)} -\sum_{k=1}^{m-1} \left|Q(k) - \sqrt{E_b}\tilde{d}_i(k)\right|^2 \quad (14.18) \\ &\qquad i_s \in \{1, \dots, N_s\}. \\ &= \max_{\substack{k_s \in \Omega_\sigma \\ m_1 = 0,1}} \left(\frac{-\left|Q(m) - \sqrt{E_b}\tilde{d}_{(k_s,m_1)}\right|^2}{g_3(i_s, k_s, m_1)} + T(m, k_s) \right) \qquad i_s \in \{1, \dots, N_s\}. \quad (14.19) \\ &= \max_{\substack{k_s \in \Omega_\sigma \\ m_1 = 0,1}} T^{(+)}(m, k_s, m_1) \end{aligned}$$

The ML metric update for each state considers each edge that connects to the state and looks at the sum of the metric from the state at the previous time and the edge metric that connects the states. Similarly the winning sequence is updated with

$$\hat{\vec{I}}_p(m+1, i_s) = \arg \max_{\substack{k_s \in \{1,\dots,N_s\} \\ i \in \Omega_{k_s}(m) \\ m_1 = 0,1}} \frac{-\left|Q(m) - \sqrt{E_b}\tilde{d}_{(k_s,m_1)}\right|^2}{g_3(i_s, i, m_1)} + \max_{i \in \Omega_{k_s}(m)} -\sum_{k=1}^{m-1} \left|Q(k) - \sqrt{E_b}\tilde{d}_i(k)\right|^2. \quad (14.20)$$

Example 14.6: Consider again the 4-PAM code of Section 13.2.2 and $\sigma(m+1) = 1$. There are two edges that connect to $\sigma(m+1) = 1$ from the states $\sigma(m) = 1$ and $\sigma(m+1) = 3$. Consequently the accumulated ML metric at time $m+1$ is given as

$$\begin{aligned} T(m+1, 1) &= \max\{Z_1, Z_2\} \quad (14.21) \\ &= \max\left\{ -\left|Q(m) + \frac{3\sqrt{E_b}}{\sqrt{5}}\right|^2 + T(m,1), -\left|Q(m) - \frac{\sqrt{E_b}}{\sqrt{5}}\right|^2 + T(m,3) \right\}. \end{aligned}$$

If $Z_1 > Z_2$ then

$$\hat{\vec{I}}_p(m+1, 1) = [\hat{\vec{I}}_p(m, 1) \;\; 0] \quad (14.22)$$

otherwise

$$\hat{\vec{I}}_p(m+1, 1) = [\hat{\vec{I}}_p(m, 3) \;\; 0]. \quad (14.23)$$

The Viterbi algorithm is a relatively simple way to compute the MLWD output. The essential idea is that only the best ML metric has to be kept at each time and at each state. The time recursive update at each state only has to consider the previous states that are edge connected to this state and find the winning metric coming into the state. The total metric coming into the state is the sum of the previous state ML metric and the branch metric. The winning bit sequence at each state is then the bit value associated with the edge of the ML path concatenated with the sequence from the previous state on the ML path. The mechanization of this process requires an add (to generate the updated metric), a compare (looking to see which metric is larger) and a select (selecting the winning sequence) so engineers often refer to add compare select (ACS) processing units when discussing the Viterbi algorithm.

14.1.3 Initialization and Termination

Intialization and termination of the Viterbi algorithm (VA) are critical to achieving optimum demodulation performance. Initialization is simply achieved by setting

$$T(1,1) = 0 \qquad T(1,k_s) = -\infty \qquad k_s \in \{2,\ldots,N_s\} \tag{14.24}$$

and operating the algorithm as in steady-state when the modulation starts off with $\sigma(1) = 1$. A similar initialization would occur when the modulation starts off in a different state. This initialization works because zero is the additive identity and having metrics of the other states set to $-\infty$ ensure that these states and their associated updates never impact the demodulation. Note this initialization procedure actually produces more computations than absolutely necessary but the algorithm is more consistent and easier to implement with this initialization.

Termination can also be done in a similar simple way. Recall that after K_b trellis updates in the VA the following metrics are obtained

$$T\left(K_b+1,i_s\right) = \max_{i\in\Omega_{i_s}(K_b+1)} -\sum_{k=1}^{m}\left|Q(k)-\sqrt{E_b}\tilde{d}_i(k)\right|^2 \qquad i_s\in\{1,\ldots,N_s\} \tag{14.25}$$

$$\hat{\vec{I}}_p\left(K_b+1,i_s\right) = \arg\max_{i\in\Omega_{i_s}(K_b+1)} -\sum_{k=1}^{m}\left|Q(k)-\sqrt{E_b}\tilde{d}_i(k)\right|^2 \qquad i_s\in\{1,\ldots,N_s\}\,. \tag{14.26}$$

At this point there are only N_s surviving sequences that could possibly be the ML word decision and the associated partial maximum likelihood metrics. The surviving sequences are also now fully K_b in length and do not need to be further extended. The final decision can be given as

$$\hat{\vec{I}} = \arg\max_{i_s\in\{1,\ldots,N_s\}}\max_{i\in\Omega_{i_s}(K_b+1)} -\sum_{k=1}^{m}\left|Q(k)-\sqrt{E_b}\tilde{d}_i(k)\right|^2 - \sum_{k=K_b+1}^{N_f}\left|Q(k)-\sqrt{E_b}\tilde{d}_{i_s}(k)\right|^2 \tag{14.27}$$

where $\tilde{d}_{i_s}(k), k\in\{K_b+1,\ldots,N_f\}$ is the transmitted symbols that would be sent in terminating the trellis when the modulation starts the termination process at $\sigma(K_b+1) = i_s$. Hence termination only has to add the terms associated with the squared Euclidean distance between the matched filter outputs and the termination symbols for each state over the last ν_c symbols to the state metrics and find the largest of the resulting metrics, i.e.,

$$\hat{\vec{I}} = \arg\max_{i_s\in\{1,\ldots,N_s\}} T(K_b+1,i_s) - \sum_{k=K_b+1}^{N_f}\left|Q(k)-\sqrt{E_b}\tilde{d}_{i_s}(k)\right|^2. \tag{14.28}$$

Again a more computationally efficient method than shown in (14.28) is possible by eliminating half the sequences every trellis stage during the termination but the method shown in (14.28) is conceptually most straightforward.

14.2 Performance of MLWD for OMWM

The error performance of a OMWM can again be bounded by using the union bound. To this end the word error probability is

$$P_W(E) = P\left(\hat{\vec{I}}\neq\vec{I}\right) = \sum_{j=0}^{M-1} P\left(\hat{\vec{I}}\neq j\,\middle|\,\vec{I}=j\right)P\left(\vec{I}=j\right). \tag{14.29}$$

In most situation it is assumed that $P\left(\vec{I}=i\right)=1/M=1/2^{K_b}$. Recall here j corresponds to a path through the trellis that defines the OMWM. Using the ideas from Chapter 8 the conditional word error probability is given as

$$P\left(\hat{\vec{I}}\neq j\,\middle|\,\vec{I}=j\right)=P\left(\bigcup_{\substack{i=0\\ j\neq i}}^{M-1}\left\{T_{i|j}>T_{j|j}\right\}\right)\leq\sum_{\substack{i=0\\ j\neq i}}^{M-1}P\left(T_{i|j}>T_{j|j}\right) \tag{14.30}$$

where again $T_{i|j}$ is the likelihood metric for the i^{th} signal conditioned on the event the j^{th} signal was transmitted. While, in general, a closed form expression for this conditional error probability is not easily found, the union bound, first explored in Chapter 8, is again applicable and an important tool for understanding the performance of an OMWM. It should also be noted that in engineering practice, word errors are often denoted frame errors. Consequently the goal of this section is to bound the word or frame error probability performance of the MLWD. This bound will give insights into the performance of OMWM. The approach to this problem will

1. Introduce the idea of a simple error event. The idea of a simple error event will reduce the number of sequences that need to be considered in a conditional union bound.
2. Using simple error events, the conditional union bound can be computed via a modified trellis
3. Using the insight from the conditional union bound computation, the complete union bound can be computed via an error state diagram.

This three step approach to performance analysis of OMWM introduces many tools that are useful to a communication engineer.

14.2.1 Simple Error Events

This section is focused on eliminating some codeword sequences from the conditional union bound. A transmitted codeword, $\vec{I}=j$, defines a path through the trellis. There are many ways to define this path but one way that uniquely defines a path is to define the set of states that that this path traverses, $\sigma_j(k),\ k\in\{1,\ldots,N_f\}$. Any codeword, $\vec{I}=i$, considered in the conditional union bound can also be specified by the states it traverses during the frame, $\sigma_i(k),\ k\in\{1,\ldots,N_f\}$. The codewords that can be removed from the conditional union bound are removed due to the trellis structure of the modulation. This is accomplished by using the idea of a simple error events.

Definition 14.4 *A* **simple error event** *is an error event that begins at time k and has length L_e which implies*

$$\begin{aligned}\sigma_j(k) &= \sigma_i(k) \qquad I_i(k)\neq I_j(k)\\ \sigma_j(k+m) &\neq \sigma_i(k+m) \qquad m=1,L_e-1\\ \sigma_j(k+L_e) &= \sigma_i(k+L_e)\end{aligned} \tag{14.31}$$

For a simple error event the likelihood metric of the possible incorrect codeword and the likelihood metric of the transmitted codeword are only different for the L_e symbols of the error event. Specifically if $\vec{I}=j$ is the transmitted codeword and if the possible incorrect codeword $\hat{\vec{I}}=i$ is composed of one simple error event at time k and of length L_e then

$$\begin{aligned}T_{j|j} &= T_{j|j}^{(-)}(k)+T_{j|j}(k,k+L_e)+T_{j|j}^{(+)}(k+L_e)\\ T_i &= T_{j|j}^{(-)}(k)+T_{i|j}(k,k+L_e)+T_{j|j}^{(+)}(k+L_e)\end{aligned} \tag{14.32}$$

where $T_{j|j}^{(-)}(k)$ is the accumulated likelihood metric for word $\vec{I}=j$ from time 1 to $k-1$, $T_{j|j}^{(+)}(k+L_e)$ is the accumulated likelihood metric for word $\vec{I}=j$ from time $k+L_e+1$ to N_f, $T_{j|j}(k,k+L_e)$ is the accumulated likelihood metric for word $\vec{I}=j$ from time k to $k+L_e$. Consequently

$$\left\{T_{i|j} > T_{j|j}\right\} = \left\{T_{i|j}(k,k+L_e) > T_{j|j}(k,k+L_e)\right\}. \tag{14.33}$$

The only part of the ML metric that will cause a conditional decoding error is accumulated during the course of the simple error event.

Definition 14.5 *Error events composed of two or more simple error events are denoted compound error events.*

Property 14.1 *Any word error $\hat{\vec{I}}=i$ that is a compound error event for $\vec{I}=j$ does not need to be included in the union bound given in (14.30).*

Proof: This can be proven for a codeword with two simple error events and a simple inductive argument leads to it being true for any number of simple error events. Assume that $\hat{\vec{I}}=i$ consists of two simple error events: one starting at k_1 of length L_{e1} and one starting at k_2 of length L_{e2}. Equivalently

$$\begin{aligned} T_{j|j} &= T_{j|j}^{(-)}(k_1) + T_{j|j}(k_1,k_1+L_{e1}) + T_{j|j}(k_1+L_{e1}+1,k_2-1) + T_{j|j}(k_2,k_2+L_{e2}) + T_{j|j}^{(+)}(k_2+L_{e2}) \\ T_{i|j} &= T_{j|j}^{(-)}(k_1) + T_{i|j}(k_1,k_1+L_{e1}) + T_{j|j}(k_1+L_{e1}+1,k_2-1) + T_{i|j}(k_2,k_2+L_{e2}) + T_{j|j}^{(+)}(k_2+L_{e2}) \end{aligned}$$

If $T_{i|j} > T_{j|j}$ then either $T_{i|j}(k_1,k_1+L_{e1}) > T_{j|j}(k_1,k_1+L_{e1})$ or $T_{i|j}(k_2,k_2+L_{e2}) > T_{j|j}(k_2,k_2+L_{e2})$. Consequently

$$\left\{T_{i|j} > T_{j|j}\right\} \subset \left\{T_{i|j}(k_1,k_1+L_{e1}) > T_{j|j}(k_1,k_1+L_{e1})\right\} \bigcup \left\{T_{i|j}(k_2,k_2+L_{e2}) > T_{j|j}(k_2,k_2+L_{e2})\right\}$$

and the event does not need to be included in the union bound due to Property 8.1 as both of the composite simple error events are already included in the union bound. □

Using only simple error events now gives a tighter form to the conditional union bound given as

$$P\left(\bigcup_{\substack{i=0\\ j\neq i}}^{M-1}\{T_i > T_j\}\,\middle|\,\vec{I}=j\right) \leq \sum_{k=1}^{K_b}\sum_{n\in\Omega_e(k,j)} P\left(T_{k,n,j} > T_j\,\middle|\,\vec{I}=j\right) < \sum_{\substack{i=0\\ j\neq i}}^{M-1} P\left(T_i > T_j\,\middle|\,\vec{I}=j\right) \tag{14.34}$$

where $T_{k,n,j}$ denotes the likelihood metric for n^{th} simple error event starting at time k for the transmitted word j. Note $\Omega_e(k,j)$ is the set of simple error events starting at time k for transmitted codeword $\vec{I}=j$ and $N_e(k,j)$ denotes the cardinality of this set or the number of simple error events that start at time k. The tighter bound here is due to the fact that any error event that is composed of multiple simple error events has been taken out of the brute force union bound. Hence the bound in (14.34) only enumerates the simple error events. Again the PWEP is given as

$$P\left(T_i > T_j\,\middle|\,\vec{I}=j\right) = P\left(T_{i|j} > T_{j|j}\right) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\Delta_E(i,j)}{4N_0}}\right) \tag{14.35}$$

where again

$$\Delta_E(i,j) = \int_{-\infty}^{\infty} |x_j(t) - x_i(t)|^2\,dt \tag{14.36}$$

is the square Euclidean distance between $x_i(t)$ and $x_j(t)$. Consequently the union bound to the conditional error probability can be computed by enumerating all possible simple error events for a transmitted word and summing the PWEPs associated with these simple error events, i.e.,

$$P\left(\hat{\vec{I}} \neq j \,\middle|\, \vec{I} = j\right) \leq \sum_{k=1}^{K_b} \sum_{n \in \Omega_e(k,j)} \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\Delta_E(\{k,n\},j)}{4N_0}}\right) \tag{14.37}$$

where $\Delta_E(\{k,n\},j)$ is the squared Euclidean distance between the transmitted signal, $x_j(t)$, and the signal for the n^{th} simple error event starting at time k for $\vec{I} = j$, $x_{\{k,n,j\}}(t)$. For simplicity of notation we denote $\Omega_e(j)$ the set of all simple error events and $\Omega_e(k,j)$ the set of all simple error events starting at time k. The conditional union bound is then completely characterized by the enumeration of the first error events and the associated squared Euclidean distance. It should be noted that the number of compound error events grows significantly with the frame size. For small frame lengths the tightening of the union bound is not significant but it becomes more significant with larger K_b.

Example 14.7: Consider a linear modulation with K_b=4, $N_f = 6$, and N_s=4 whose trellis diagram is given in Fig. 14.2. Note this trellis has $\sigma(1) = 1$ and $\sigma(7)$=1. Assume the top transition corresponds to $I(k)$=0 and the bottom transition corresponds to $I(k)$=1. If the transmitted word is $\vec{I} = 0$ then the simple sequences are easily enumerated:

$$\begin{aligned} k &= 1 \quad \Omega_e(1,0) = \{1,3,5,7,11,13,15\} \\ k &= 2 \quad \Omega_e(2,0) = \{2,6,10,14\} \\ k &= 3 \quad \Omega_e(3,0) = \{4,12\} \\ k &= 4 \quad \Omega_e(4,0) = \{8\} \end{aligned} \tag{14.38}$$

In this simple example only one word, $\vec{I} = 9$, of the 15 possible error sequences is composed of more than one simple sequence. This particular word takes a path through the trellis that leaves the transmitted state at $k = 1$ and returns to the transmitted state at $k = 3$ and immediately leaves the transmitted state at $k = 4$ and again rejoining the transmitted state at $k = N_f + 1 = 7$.

14.2.2 The Modified Trellis

The enumeration of the simple error events for a modulation with memory corresponds to an enumeration of the possible paths through a modified trellis. Note the error events that are excluded from the union bound are paths through the trellis that have left the transmitted path and rejoined it at a future point more than one time. Consequently if the trellis for the modulation with memory is modified to eliminate the possibility of an error event having more than one simple error event then we have a way to enumerate all simple error events. This is accomplished by adding an additional state per trellis stage. This state represents the termination of a simple error event at that stage or time. Since once a simple error event completes it no longer has to be considered, each of these augmented states have no outputs or are "absorbing". Once this modified trellis is constructed the necessary simple error events for the union bound correspond directly to all paths through the modified trellis.

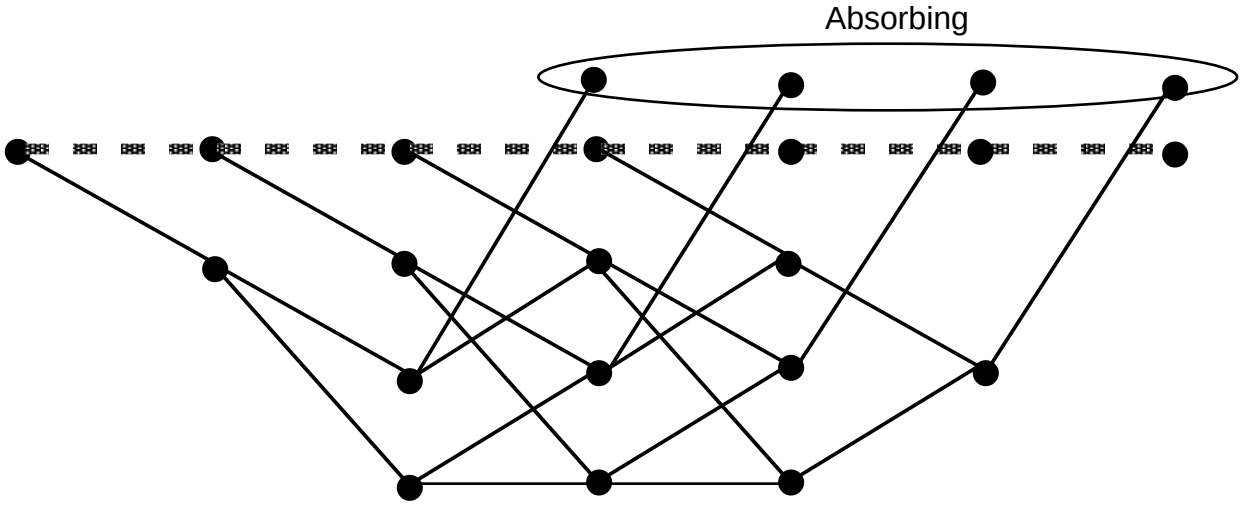


Figure 14.6: A trellis diagram for a modulation with memory having K_b=4, $N_c = 6$, and N_s=4 used for enumerating all the simple error events corresponding to $\vec{I} = 0$.

Example 14.8: Consider a linear modulation with K_b=4, $N_f = 6$, and N_s=4 whose trellis diagram is given in Fig. 14.2. The modified trellis to be used for the computation of the union bound for $\vec{I} = 0$ is given in Fig. 14.6. Note the dashed path corresponds to the transmitted signal while all the paths that terminate in the "absorbing" state correspond to the simple error events. If the transmitted word is $\vec{I} = 0$ then the simple sequences that terminate at each time are easily enumerated:

$$\begin{aligned} k &= 4 \quad \{1\} \\ k &= 5 \quad \{2,3\} \\ k &= 6 \quad \{4,5,6,7\} \\ k &= 7 \quad \{8,10,11,12,13,14,15\} \end{aligned} \tag{14.39}$$

The modified trellis enumerates all the simple error events and the union bound can be computed by identifying the Euclidean square distance of all the simple error events. Recall that for an OMWM that

$$\Delta_E(i,j) = E_b \sum_{l=1}^{N_f} \left| \tilde{d}_i(l) - \tilde{d}_j(l) \right|^2 = \sum_{l=1}^{N_f} \Delta_E(i,j,l). \tag{14.40}$$

Consequently the Euclidean square distance is additive over each time interval of the frame. In fact, each edge not associated with the edge that the transmitted signal uses has an associated edge squared Euclidean distance compared to the transmitted signal. Since an edge is entirely determined by the two states it connects, we will use interchangeably the notation

$$\Delta_E(\vec{I} = i, \vec{I} = j, l) = \Delta_\sigma(\sigma_i(l), \sigma_i(l+1), \sigma_j(l), \sigma_j(l+1)) \tag{14.41}$$

to indicate the edge Euclidean square distance. For notational consistency if an edge does not exist between either $\sigma_i(l)$ and $\sigma_i(l+1)$ or $\sigma_j(l)$ and $\sigma_j(l+1)$ then we set

$$\Delta_\sigma(\sigma_i(l), \sigma_i(l+1), \sigma_j(l), \sigma_j(l+1)) = \infty. \tag{14.42}$$

The total squared Euclidean distance for each path in the modified trellis is the sum of the edge squared Euclidean distance. Consequently for reasonably short frames the modified trellis diagram can function to both enumerate the simple error events and the associated Euclidean square distance.

Example 14.9: Consider again the 4-PAM code of Section 13.2.2 and assume the transmitted edge at time l is given as $\sigma_j(l) = 1$ and $\sigma_j(l+1) = 1$. The edge Euclidean distance is then given as

$\sigma(l)$	$\sigma(l+1)$	$\Delta_\sigma(\sigma_i(l), \sigma_i(l+1), 1, 1)$
1	1	0
1	2	16/5
1	3	∞
1	4	∞
2	1	∞
2	2	∞
2	3	36/5
2	4	4/5
3	1	16/5
3	2	0
3	3	∞
3	4	∞
4	1	∞
4	2	∞
4	3	4/5
4	4	36/5

The combination of the modfied trellis and the edge distance enumeration allows the conditional union bound to be computed. The modified trellis allows a communication engineer to enumerate all of the possible simple error events for the transmitted word. The squared Euclidean distance of each of the simple error events can be computed from the edge distance enumeration. Once the squared Euclidean distance is known for each simple error events the conditional union bound can be computed.

Example 14.10: Consider again the 4-PAM OMWM of Section 13.2.2 and assume $\vec{I} = 0$. The enumerated simple error events have the following sets of squared Euclidean distance

$$\Delta_E(1,0) = \frac{16}{5} + \frac{36}{5} + \frac{16}{5} \qquad \Delta_E(2,0) = 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5}$$

$$\Delta_E(3,0) = \frac{16}{5} + \frac{4}{5} + \frac{4}{5} + \frac{16}{5} \qquad \Delta_E(4,0) = 0 + 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5}$$

$$\Delta_E(5,0) = \frac{16}{5} + \frac{36}{5} + 0 + \frac{36}{5} + \frac{16}{5} \qquad \Delta_E(6,0) = 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5}$$

$$\Delta_E(7,0) = 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5} \qquad \Delta_E(8,0) = 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5}$$

$$\Delta_E(10,0) = 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5} \qquad \Delta_E(11,0) = 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5}$$

$$\Delta_E(12,0) = 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5} \qquad \Delta_E(13,0) = 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5}$$

$$\Delta_E(14,0) = 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5} \qquad \Delta_E(15,0) = 0 + \frac{16}{5} + \frac{36}{5} + \frac{16}{5}$$

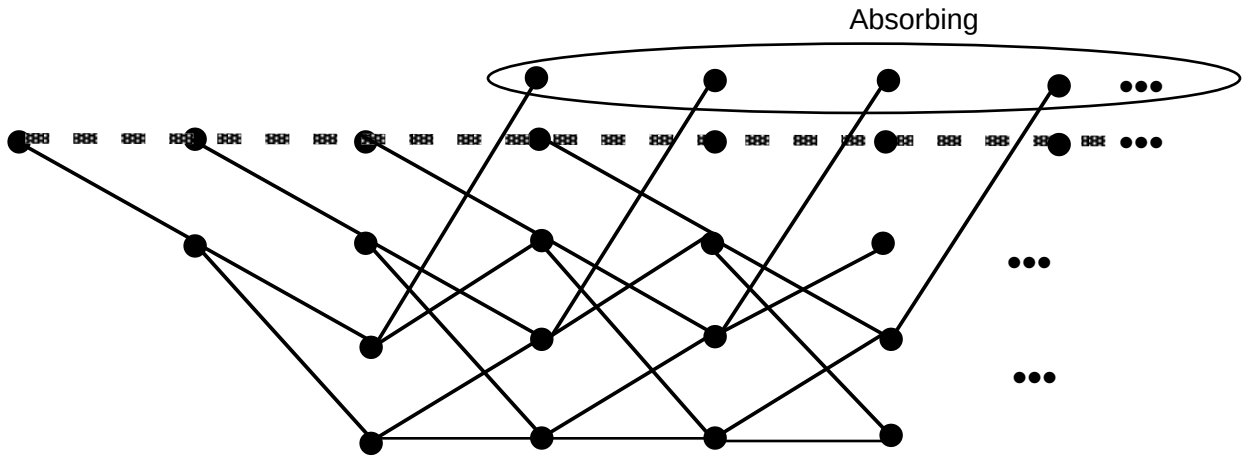


Figure 14.7: A trellis diagram for a modulation with memory having N_s=4 used for enumerating all the simple error events corresponding to $\vec{I} = 0$.

14.3 Approximations for Large Frame Size

In many communications applications the frame or word size is large. Large frames can make the enumeration of the simple error event tedious. Fortunately the large frame provides a more uniform structure to the OMWM and this structure can lead to simplified union bounds. In the case of a large frame size the trellis used to enumerate the simple error events begins to look more uniform and the trellis termination effects do not have such a big effect. As an example consider a modulation with memory having $N_s = 4$ with a very large K_b where $\vec{I} = 0$. The trellis for enumerating the simple error events is shown in Fig. 14.7. In examining the trellis in Fig. 14.7 it is apparent that the longer the trellis becomes in time the more $\Omega_e(k, 0)$ begins to have the same form for each k (except for the time shift). Consequently this has led to a simpler form for the analysis of the word or frame error probability. Assuming an infinite length trellis one only need to enumerate the simple error events starting at a single time k. The simple error events for a single time k are often denoted the first error events in the literature. The enumeration of the simple error events in the union bound is greatly simplified by only having to consider error events starting at a single time. While at first glance assuming an infinite trellis size seems to produce an infinite number of first error events, the recursive structure embedded in the first error events allows some straightforward computations. The union bound now can be approximated (and strictly upper bounded) as

$$P\left(\hat{\vec{I}} \neq j \,\middle|\, \vec{I} = j\right) \leq K_b \sum_{n \in \Omega_e(1,j)} \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\Delta_E(\{n, j\}, j)}{4N_0}}\right). \tag{14.43}$$

where $\Delta_E(\{n, j\}, j)$ is the squared Euclidean distance between the transmitted signal, $x_j(t)$, and the signal for the n^{th} first error event for $\vec{I} = i$, which corresponds to the signal denoted $x_{\{n,j\}}(t)$.

When enumerating out first error events the trellis code is best viewed as a finite state machine. This finite state machine has a number of similiarities to a finite state Markov chain [Gal01] and several tool from that area will be used to derive some solutions for the infinite frame case. Recall the tools of Markov chains were used in Section 13.3.2, as well, when considering the spectrum of a long frame of data from an OMWM. This discussion will only consider the case of the transmitted codeword being $\vec{I} = 0$ to simplify the ideas and the sequel will generalize the ideas to consider all possible transmitted

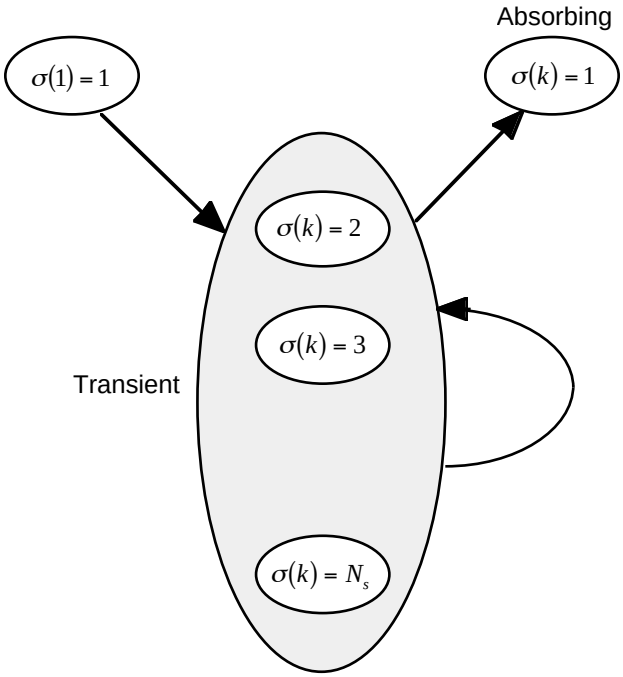


Figure 14.8: A state diagram for a modulation with memory used for enumerating all the first error events corresponding to $\vec{I} = 0$.

codewords. It will be assumed here in the discussion that $\vec{I} = 0$ has $\sigma(k) = 1\ \forall k$. The state transition diagram for enumerating the first error events for $\vec{I} = 0$ is given in Fig. 14.8. Essentially the first error event must diverge from $\sigma(2) = 1$. This forces a transition into the "transient" states, where these transient states represent $\sigma_i(k) \neq 1, k > 1$. A first error event then moves among the transient states until the error event is completed. This completion is represented as a transition to the absorbing state that represents $\sigma_i(k) = 1$. All paths through the state diagram in Fig. 14.8 represent all the possible first error events of the modulation with memory. The state diagram will prove useful for numerically evaluating the union bound to the frame or word error rate.

Example 14.11: Consider again the modulation with memory having N_s=4 and $\vec{I} = 0$. Fig. 14.9 shows the state transition diagram for enumerating the first error events for this modulation corresponding to $\vec{I} = 0$.

There is a roadblock at this point in that there is an infinite number of first error events that would arise by considering the chain in Fig. 14.8. Consequently there is no way to enumerate the Euclidean square distance like discussed in the previous section. While the Euclidean square distance has an additive characteristic in the chain, i.e.,

$$\Delta_E(i,0) = E_b \sum_{l=1}^{N_f} \left| \tilde{d}_i(l) - \tilde{d}_0(l) \right|^2 = \sum_{l=1}^{N_f} \Delta_E(i,0,l). \tag{14.44}$$

most results for characterizing finite state machine outputs exist when the characterisitcs of interest

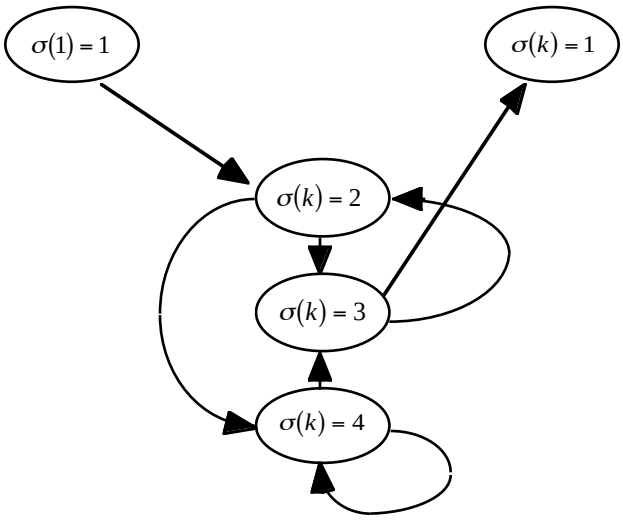


Figure 14.9: A state diagram for a modulation with memory with N_s=4 used for enumerating all the first error events corresponding to $\vec{I} = 0$.

have a multiplicative form in the chain [Gal01, MZ60]. Consequently progress could be made if

$$\frac{1}{2}\text{erfc}\left(\sqrt{\frac{\Delta_E(i,0)}{4N_0}}\right) = \prod_{l=1}^{L_e(i)} P\left(i,0,l\right) \tag{14.45}$$

where $L_e(i)$ is the length of the error event for codeword $\vec{I} = i$. While (14.45) is not valid, the consideration of (14.44) does give us hope when we note

$$\exp\left[C\Delta_E(i,0)\right] = \prod_{l=1}^{L_e(i)} \exp\left[C\Delta_E(i,0,l)\right] \tag{14.46}$$

where C is an arbitrary constant. Using the idea in (14.46) it is possible to make progress on two fronts in the analysis

1. An upper bound (the Chernoff bound [Che52]) to the conditional union bound can be formed by noting that (see (7.28))

$$\frac{1}{2}\text{erfc}\left(\sqrt{\frac{\Delta_E(i,0)}{4N_0}}\right) \leq \frac{1}{2}\exp\left[\frac{-\Delta_E(i,0)}{4N_0}\right] = \frac{1}{2}\prod_{l=1}^{L_e(i)} \exp\left[\frac{\Delta_E(i,0,l)}{4N_0}\right] \tag{14.47}$$

2. The true conditional union bound can be numerically calculated by using Craig's form [Cra91, SD98] for the erfc $(\bullet)$ (see Problem 9.17), i.e.,

$$\frac{1}{2}\text{erfc}\left(\sqrt{\frac{\Delta_E(i,0)}{4N_0}}\right) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \exp\left[\frac{-\Delta_E(i,0)}{4N_0\cos^2\left(n_p\right)}\right] dn_p = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \prod_{l=1}^{L_e(i)} \exp\left[\frac{-\Delta_E(i,0,l)}{4N_0\cos^2\left(n_p\right)}\right] dn_p \tag{14.48}$$

The Chernoff bound to the performance for OMWM is considered by many texts [For72, VO79, Pro89, BDMS91] so students interested in that bound can consult those references. This text will concentrate on the computing the true union bound. Craig's method is an extremely powerful analysis tool and students interested in more details than provided in the text are referred to [SA00]

The conditional union bound can be computed by using some characteristics of a finite state machine. Combining (14.43) with (14.48) gives

$$\begin{aligned} P\left(\hat{\vec{I}} \neq 0 \middle| \vec{I} = 0\right) &\leq K_b \sum_{n \in \Omega_e(1,0)} \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\Delta_E(\{n,0\},j)}{4N_0}}\right) \\ &= \frac{K_b}{\pi}\int_0^{\frac{\pi}{2}} \sum_{n \in \Omega_e(1,0)} \prod_{l=1}^{L_e(n)} \exp\left[\frac{-\Delta_E(n,0,l)}{4N_0\cos^2(n_p)}\right] dn_p. \end{aligned} \tag{14.49}$$

At this point it will be noted that for a fixed n_p if

$$\sum_{n \in \Omega_e(1,0)} \prod_{l=1}^{L_e(n)} \exp\left[\frac{-\Delta_E(n,0,l)}{4N_0\cos^2(n_p)}\right] \tag{14.50}$$

can be evaluated in a simple manner than the integral in (14.49) is numerically simple to evaluate. Consequently we will focus on evaluating (14.50). As a first step, recall that edges are entirely determined by the states they connect. Because of this characteristic first error events of length $L_e = 2$ can completely enumerated by considering all possible transient states at time 2 because for a first error event of length two, it is apparent that $\sigma_n(1) = 0$ and $\sigma_n(3) = 0$ by definition of a first error event. Consequently the terms in (14.50) that have $L_e = 2$ are given as

$$\sum_{n \in \Omega_2(1,0)} \prod_{l=1}^{L_e(n)} \exp\left[\frac{-\Delta_E(n,0,l)}{4N_0\cos^2(n_p)}\right] = \sum_{m=2}^{N_s} \exp\left[\frac{-\Delta_\sigma(0,m,0,0)}{4N_0\cos^2(n_p)}\right] \exp\left[\frac{-\Delta_\sigma(m,0,0,0)}{4N_0\cos^2(n_p)}\right] \tag{14.51}$$

where we have used the notation $\Omega_{L_e}(1,0)$ to denote the set of all error events of length L_e. Much like in Section 13.3.2 vectors can be used to simplify the notation so we define

$$\vec{S}_{gb}(n_p) = \left[\exp\left[\frac{-\Delta_\sigma(0,1,0,0)}{4N_0\cos^2(n_p)}\right] \exp\left[\frac{-\Delta_\sigma(0,2,0,0)}{4N_0\cos^2(n_p)}\right] \cdots \exp\left[\frac{-\Delta_\sigma(0,N_s,0,0)}{4N_0\cos^2(n_p)}\right]\right] \tag{14.52}$$

$$\vec{S}_{bg}(n_p) = \left[\exp\left[\frac{-\Delta_\sigma(1,0,0,0)}{4N_0\cos^2(n_p)}\right] \exp\left[\frac{-\Delta_\sigma(2,0,0,0)}{4N_0\cos^2(n_p)}\right] \cdots \exp\left[\frac{-\Delta_\sigma(N_s,0,0,0)}{4N_0\cos^2(n_p)}\right]\right]. \tag{14.53}$$

These vectors (size $1 \times N_s - 1$) represent a transition from the correct state (or *good* state[1]) to the set of incorrect states (or *bad* states) (hence subscript gb) or vice versa (hence subscript bg). Using the vector notation we have now have

$$\sum_{n \in \Omega_2(1,0)} \prod_{l=1}^{L_e(n)} \exp\left[\frac{-\Delta_E(n,0,l)}{4N_0\cos^2(n_p)}\right] = \vec{S}_{gb}(n_p)\vec{S}_{bg}^T(n_p). \tag{14.54}$$

Any first error event that is longer than $L_e = 2$ will dwell within the *bad* states in between these

[1]The *good* and *bad* notation is borrowed from [Big84]

transitions. To represent this transition among *bad* states, define a transition matrix of the form

$$\mathbf{S}_b(n_p) = \begin{bmatrix} \exp\left[\frac{-\Delta_\sigma(2,2,0,0)}{4N_0\cos^2(n_p)}\right] & \exp\left[\frac{-\Delta_\sigma(2,3,0,0)}{4N_0\cos^2(n_p)}\right] & \cdots & \exp\left[\frac{-\Delta_\sigma(2,N_s,0,0)}{4N_0\cos^2(n_p)}\right] \\ \exp\left[\frac{-\Delta_\sigma(3,2,0,0)}{4N_0\cos^2(n_p)}\right] & \exp\left[\frac{-\Delta_\sigma(3,3,0,0)}{4N_0\cos^2(n_p)}\right] & & \\ \vdots & & \ddots & \\ \exp\left[\frac{-\Delta_\sigma(N_s,2,0,0)}{4N_0\cos^2(n_p)}\right] & \cdots & & \exp\left[\frac{-\Delta_\sigma(N_s,N_s,0,0)}{4N_0\cos^2(n_p)}\right] \end{bmatrix} \tag{14.55}$$

Using this vector notation, the terms in (14.50) that have $L_e = 2 + k$ are given as

$$\sum_{n\in\Omega_{2+k}(1,0)} \prod_{l=1}^{L_e(n)} \exp\left[\frac{-\Delta_E(n,0,l)}{4N_0\cos^2(n_p)}\right] = \vec{S}_{gb}(n_p)\mathbf{S}_b(n_p)^k\vec{S}_{bg}^T(n_p) \tag{14.56}$$

and consequently (14.50) itself is given

$$\sum_{n\in\Omega_e(1,0)} \prod_{l=1}^{L_e(n)} \exp\left[\frac{-\Delta_E(n,0,l)}{4N_0\cos^2(n_p)}\right] = \vec{S}_{gb}(n_p)\left(\sum_{k=0}^{\infty}\mathbf{S}_b(n_p)^k\right)\vec{S}_{bg}^T(n_p) \tag{14.57}$$

Using the representation of a powers series gives the closed form result of

$$\sum_{n\in\Omega_e(1,0)} \prod_{l=1}^{L_e(n)} \exp\left[\frac{-\Delta_E(n,0,l)}{4N_0\cos^2(n_p)}\right] = \vec{S}_{gb}(n_p)\left(\mathbf{I}_{N_s-1} - \mathbf{S}_b(n_p)\right)^{-1}\vec{S}_{bg}^T(n_p). \tag{14.58}$$

We now have derived an algorithm to compute the union bound. The code can be summarized in pseudo code as

Union Bound Computation

- Identify $\vec{S}_{bg}^T(n_p)$, $\vec{S}_{bg}^T(n_p)$, and $\mathbf{S}_b(n_p)$ as a function of n_p and E_b/N_0.
- Choose E_b/N_0.
- For $m = 1 : M_p$ (points for numeric integration)
  1. $n_p = \frac{(m-1)\pi}{2(M_p-1)}$
  2. Compute $\left(\mathbf{I}_{N_s-1} - \mathbf{S}_b(n_p)\right)^{-1}$
  3. Compute $Z(n_p) = \vec{S}_{gb}(n_p)\left(\mathbf{I}_{N_s-1} - \mathbf{S}_b(n_p)\right)^{-1}\vec{S}_{bg}^T(n_p)$
- Compute $\frac{K_b}{\pi}\int_0^{\frac{\pi}{2}} Z(n_p)dn_p$.

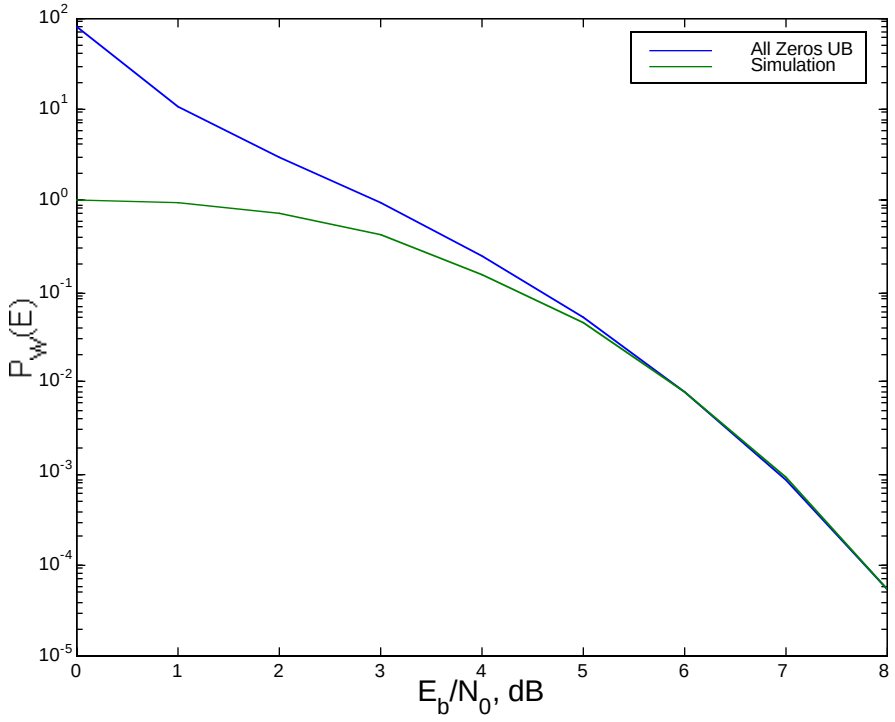


Figure 14.10: The frame error rate of the 4-PAM OMWM of Section 13.2.2 with $\vec{I} = 0$ from simulation and from the large frame size analysis. $K_b = 200$.

Example 14.12: Consider again the 4-PAM OMWM of Section 13.2.2 and assume $\vec{I} = 0$. To compute the conditional union bound the vectors $\vec{S}_{bg}^T(n_p)$ and $\vec{S}_{bg}^T(n_p)$ and the matrix $\mathbf{S}_b(n_p)$ need to be identified as a function of n_p and E_b/N_0. To this end we have

$$\vec{S}_{gb}(n_p) = \left[\exp\left[\frac{-4E_b}{5N_0\cos^2(n_p)}\right] \; 0 \; 0\right] \tag{14.59}$$

$$\vec{S}_{bg}(n_p) = \left[0 \; \exp\left[\frac{-4E_b}{5N_0\cos^2(n_p)}\right] \; 0\right]. \tag{14.60}$$

and

$$\mathbf{S}_b(n_p) = \begin{bmatrix} 0 & \exp\left[\frac{-9E_b}{5N_0\cos^2(n_p)}\right] & \exp\left[\frac{-E_b}{5N_0\cos^2(n_p)}\right] \\ 1 & 0 & 0 \\ 0 & \exp\left[\frac{-E_b}{4N_0\cos^2(n_p)}\right] & \exp\left[\frac{-9E_b}{4N_0\cos^2(n_p)}\right] \end{bmatrix}. \tag{14.61}$$

The conditional union bound and the simulation performance is plotted in Fig. 14.10 for a frame of $K_b = 200$ bits. This bound for the infinite frame length is very tight for at moderate SNR and also K_b does not have to be extremely large for it also to be useful.

14.3.1 Full Union Bound

Now that the idea of using the error state diagram to enumerate the first error events for $\vec{I} = 0$ is understood we can generalize this technique to compute the entire union bound. The computation of entire union bound requires enumeration and averaging of the first error events for all possible transmitted codewords $\vec{I} = j$. Recall when $\vec{I} = 0$ the first error event was entirely characterized by a finite state machine that could track the trajectory of the error state $\sigma_i(k)$. For considering all transmitted codewords the analysis must track the trajectory of both the transmitted state, $\sigma_i(k)$ and the error state, $\sigma_i(k)$. The error state diagram for this case is formed by associating the transmitted state and the decoder state into a product state denoted as $(\sigma_i(k) = n_i, \sigma_j(k) = n_j), n_i, n_j = 1, N_s$. For $R = 1$ each product state will have 4 branches exiting the state, each representing the joint transmitted and decoded bit pair $(I(k) = l, \hat{I}(k) = m), l, m = 0, 1$. The product states (σ_i, σ_j) are denoted as *good* if $\sigma_i = \sigma_j$ and *bad* otherwise. The number of product states for a N_s state trellis code is N_s^2 among which $N_g = N_s$ will be *good* and $N_b = N_s(N_s - 1)$ will be *bad*.

A first error event can be viewed in this error state diagram as a path that diverges from a *good* state into a *bad* state and moves among the *bad* states until it transitions back into a *good* state. Similar to $\vec{I} = 0$ case N_s additional states are added to error state diagram as absorbing states to represent the termination of an error event.

Example 14.13: Consider again the OMWM having N_s=4 of Section 13.2.2. Fig. 14.11 shows the state transition diagram for enumerating the first error events for all possible transmitted codewords for this modulation. There are four good state and twelve bad states. Each state has four branches leaving the state.

While this error state diagram is more complicated it's operation follows in much the same way as the case for $\vec{I} = 0$ case. There are transitions from the *good* states to the *bad* states and vice versa as for the $\vec{I} = 0$ case. The only difference is that there are N_s *good* states and $N_s(N_s - 1)$ *bad* states. The only major difference is that each of the paths through the trellis accounts for both of the possible transmitted sequences and the possible error sequences. The union bound wants to consider all the error sequences while averaging over all the possible transmitted sequences. This averaging requires each transition to account for the random data bits of the error event. Consequently the error event is represented transitions to and from the *good* and *bad* states with two matrices

$$\mathbf{S}_{gb}(n_p)(\text{size } N_s \times N_s(N_s - 1)) \qquad \mathbf{S}_{bg}(n_p)(\text{size } N_s(N_s - 1) \times N_s) \tag{14.62}$$

The entries in these matrices are again of the form

$$\frac{1}{2} \exp\left[\frac{-\Delta_\sigma(i, j, k, l)}{4N_0 \cos^2(n_p)}\right] \tag{14.63}$$

where recall that the notation (i, j, k, l) is an enumeration of the true transmitted edge and possible error edge. $(1, 2, 1, 1)$ implies $\sigma_i(k) = 1$, $\sigma_i(k + 1) = 2$, $\sigma_j(k) = 1$, and $\sigma_j(k + 1) = 1$. The $\frac{1}{2}$ term takes into account the random transmitted modulation. For instance, consider error state (1,1) in Fig. 14.11. The random transmitted bit can cause the transmitted state to either go to state 1 or state 2. The union bound must provide an average over these two paths and this is accomplished with this $\frac{1}{2}$ term. Likewise there is a $N_s(N_s - 1) \times N_s(N_s - 1)$ matrix to represent the transitions within the bad states and this matrix is again given as $\mathbf{S}_b(n_p)$.

Following steps in the previous section gives an integrand for the first error event that starts in *good* state (j_1, j_1) and ends up in *good* state (j_2, j_2) as

$$\sum_{n \in \Omega_e(1, j_1, j_s)} \prod_{l=1}^{L_e(n)} \exp\left[\frac{-\Delta_E(n, 0, l)}{4N_0 \cos^2(n_p)}\right] = \left[\mathbf{S}_{gb}(n_p) \left(\mathbf{I}_{N_s - 1} - \mathbf{S}_b(n_p)\right)^{-1} \mathbf{S}_{bg}^T(n_p)\right]_{j_1 j_s}. \tag{14.64}$$

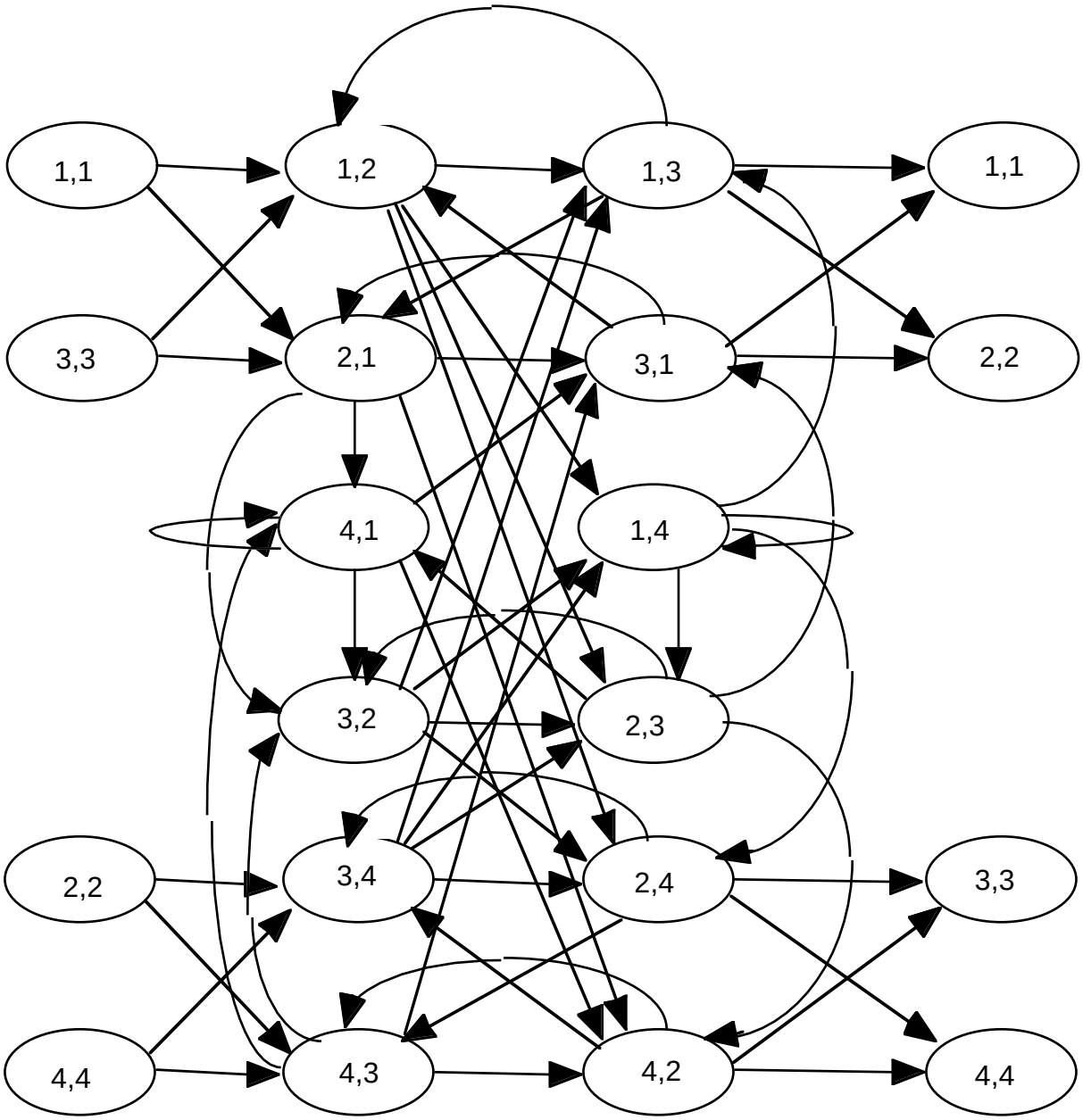


Figure 14.11: The error state diagram for the HCV OMWM.

All error events require all the elements in the matrix in (14.64) to be summed up. Consequently all first error events can be accounted for as

$$\sum_{n\in\Omega_e(1)}\prod_{l=1}^{L_e(n)}\exp\left[\frac{-\Delta_E(n,0,l)}{4N_0\cos^2(n_p)}\right]=\vec{1}_{N_s}^T\mathbf{S}_{gb}(n_p)\left(\mathbf{I}_{N_s-1}-\mathbf{S}_b(n_p)\right)^{-1}\mathbf{S}_{bg}^T(n_p)\vec{1}_{N_s}. \tag{14.65}$$

This implies the complete union bound can be expressed as

$$\begin{aligned}P\left(\hat{\vec{I}}\neq\vec{I}\right) &\leq \frac{K_b}{\pi N_s}\int_0^{\frac{\pi}{2}}\sum_{n\in\Omega_e(1)}\prod_{l=1}^{L_e(n)}\exp\left[\frac{-\Delta_E(n,0,l)}{4N_0\cos^2(n_p)}\right]dn_p.\\ &= \frac{K_b}{\pi N_s}\int_0^{\frac{\pi}{2}}\vec{1}_{N_s}^T\mathbf{S}_{gb}(n_p)\left(\mathbf{I}_{N_s-1}-\mathbf{S}_b(n_p)\right)^{-1}\mathbf{S}_{bg}^T(n_p)\vec{1}_{N_s}dn_p.\end{aligned} \tag{14.66}$$

The N_s in the denominator is to take into account the random starting good state. It can be seen that the algorithm for the complete union bound is essentially the same as the algorithm for the conditional ($\vec{I}=0$) union bound with just larger matrices and some extra constants to accomodate all the random transmitted paths that produce the error events.

Example 14.14: Consider again the OMWM having N_s=4 of Section 13.2.2. Enumerating the *bad* error states as (1,2)=1, (2,1)=1,..., (4,2)=12, the matrices needed for the union bound are given as

$$\mathbf{S}_{gb}(n_p) = \begin{bmatrix} W_1 & W_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ W_1 & W_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & W_1 & W_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & W_1 & W_1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{14.67}$$

$$\mathbf{S}_{bg}(n_p) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & W_1 & W_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & W_1 & W_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & W_1 & W_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & W_1 & W_1 \end{bmatrix} \tag{14.68}$$

and

$$\mathbf{S}_{b}(n_p) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & W_2 & 0 & W_3 & W_3 & W_3 & 0 \\ 0 & 0 & W_3 & W_3 & 0 & 0 & 0 & W_2 & 0 & 0 & 0 & W_3 \\ 0 & 0 & W_2 & W_3 & 0 & 0 & 0 & W_3 & 0 & 0 & 0 & W_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & W_3 & 0 & W_3 & W_2 & W_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & W_3 & 0 & W_3 & W_3 & W_2 & 0 \\ 0 & 0 & W_3 & W_3 & 0 & 0 & 0 & W_3 & 0 & 0 & 0 & W_2 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & W_3 & 0 & W_2 & W_3 & W_3 & 0 \\ 0 & 0 & W_3 & W_2 & 0 & 0 & 0 & W_3 & 0 & 0 & 0 & W_3 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{14.69}$$

where

$$W_1 = \frac{1}{2}\exp\left[\frac{-4E_b}{5N_0\cos^2(n_p)}\right] \qquad W_2 = \frac{1}{2}\exp\left[\frac{-9E_b}{5N_0\cos^2(n_p)}\right] \qquad W_3 = \frac{1}{2}\exp\left[\frac{-E_b}{5N_0\cos^2(n_p)}\right].$$

The union bound and the simulation performance is plotted in Fig. 14.12 for a frame of $K_b = 200$ bits.

14.3.2 State Reduction Techniques

The bound on the word error performance for OMWM can be simplified in a variety of ways. There are in general two aspects to the complexity of calculating the bound

1. The evaluation of $Z(n_p) = \vec{1}_{N_s}^T \mathbf{S}_{gb}(n_p)\left(\mathbf{I}_{N_s-1} - \mathbf{S}_b(n_p)\right)^{-1}\mathbf{S}_{bg}^T(n_p)\vec{1}_{N_s}$. This requires an $(N_s-1)N_s \times (N_s-1)N_s$ matrix inversion and two matrix multiplies.

2. The evaluation of the integral in (14.66).

In general both of these aspects can be made simpler. The numeric integration is dominated by the values of n_p near zero and the computation of $Z(n_p)$ can be simplified by using state reduction techniques. The idea of state reduction is that many states in a state diagram are equivalent and can be grouped together. Since the states are grouped together the size of the state space is significantly reduced and the computation associated with $Z(n_p)$ is significantly reduced. This idea of state reduction goes back

Figure 14.12: The frame error rate of the 4-PAM OMWM of Section 13.2.2 from simulation and from the large frame size analysis. $K_b = 200$.

$\sigma(l)$	$T(l,\sigma(l))$	$\hat{\vec{I}}(l,\sigma(l))$
1	-2	$\vec{I}_1$
2	-3	$\vec{I}_2$
3	-4	$\vec{I}_3$
4	-6	$\vec{I}_4$

Table 14.1: The Viterbi algorithm metrics at time l.

to Boolean logic synthesis and is equally applicable to the Chernoff bound computation. Detailed references for these state reduction techniques are given in [SW04, **?**].

Example 14.15: Consider again the OMWM having N_s=4. In this case the modulation is entirely represented by one *good* state and three *bad* states with

$$\vec{S}_{gb}(n_p) = \left[\exp\left[\frac{-4E_b}{5N_0\cos^2(n_p)}\right]\ 0\ 0\right] \tag{14.70}$$

$$\vec{S}_{bg}(n_p) = \left[0\ \exp\left[\frac{-4E_b}{5N_0\cos^2(n_p)}\right]\ 0\right]. \tag{14.71}$$

and

$$\mathbf{S}_b(n_p) = \begin{bmatrix} 0 & \frac{1}{2}\exp\left[\frac{-E_b}{5N_0\cos^2(n_p)}\right] + \frac{1}{2}\exp\left[\frac{-36E_b}{5N_0\cos^2(n_p)}\right] & \exp\left[\frac{-E_b}{5N_0\cos^2(n_p)}\right] \\ 1 & 0 & 0 \\ 0 & \exp\left[\frac{-E_b}{5N_0\cos^2(n_p)}\right] & \frac{1}{2}\exp\left[\frac{-E_b}{5N_0\cos^2(n_p)}\right] + \frac{1}{2}\exp\left[\frac{-36E_b}{5N_0\cos^2(n_p)}\right] \end{bmatrix} \tag{14.72}$$

14.4 Other Methods for MLWD for OMWM

Not completed this edition

14.5 Homework Problems

Problem 14.1. Consider the OMWM defined with the trellis given in Fig 14.13. The first component on the branch label is $I(m)$ and the second component is $\tilde{d}_i(m)\sqrt{E_b}$. At time l in a Viterbi decoder for this OMWM the surviving maximum likelihood metrics and sequence is given in Table 14.1. With $Q(l) = 2$, find $T(l+1,\sigma(l+1))$ and $\hat{\vec{I}}_p(l+1,\sigma(l+1))$ for $\sigma(l+1) \in \{1,2,3,4\}$.

Problem 14.2. $K_b = 4$ bits of the alternate mark inversion OMWM has a trellis description shown in Fig. 14.14.

a) If the transmitted signal is $\vec{I} = 0$ find all the simple error events for this OMWM. A clear drawing will suffice for an answer. How many error events out of the 15 possible error events have been eliminated from the union bound by the concept of the simple error event for this OMWM?

b) Using the concept of a simple error event find the tightest union bound to $P_W\left(E\left|\vec{I}=0\right.\right)$.

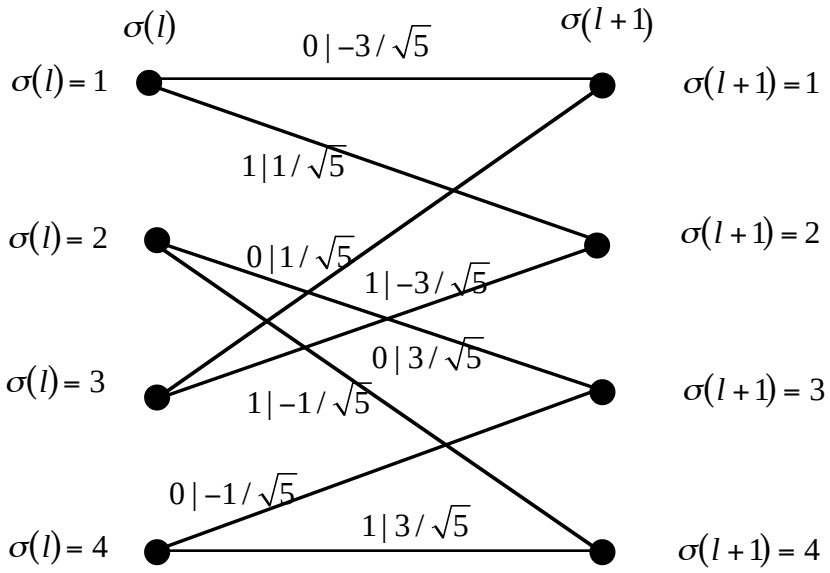


Figure 14.13: A trellis diagram for a modulation with memory.

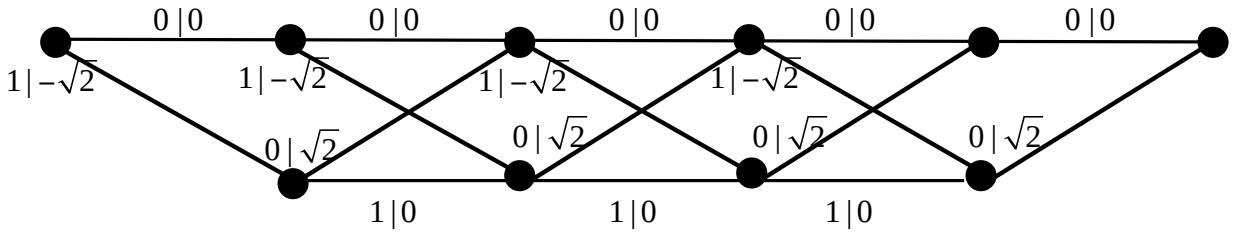


Figure 14.14: The trellis diagram for AMI with K_b=4.

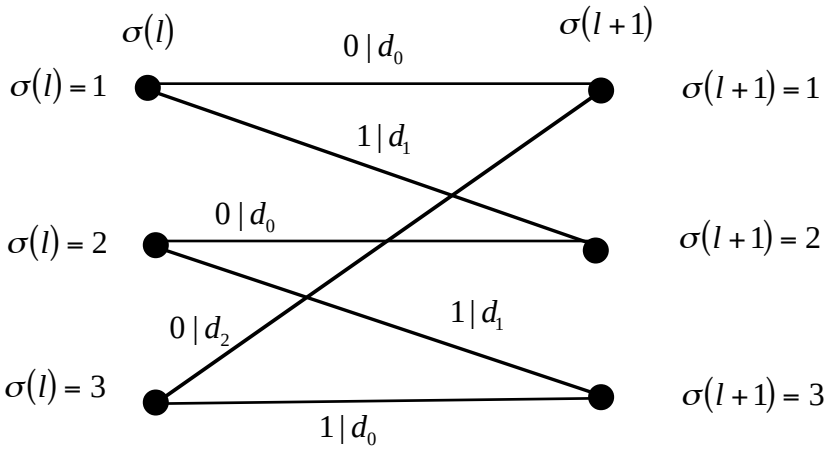


Figure 14.15: A trellis diagram for a modulation with memory.

$\sigma(l)$	$T(l,\sigma(l))$	$\hat{\vec{I}}(l,\sigma(l))$
1	2	$\vec{I}_1$
2	3	$\vec{I}_2$
3	4	$\vec{I}_3$

Table 14.2: The Viterbi algorithm metrics at time l.

Problem 14.3. Consider the orthogonal modulation with memory (OMWM) defined with the trellis given in Fig 14.15. At time l in a Viterbi decoder for this OMWM, the surviving maximum likelihood metric and sequence at each state is given in Table 14.2. Assume that $d_0 = 1,\ d_1 = \exp[j2\pi/3],\ d_2 = \exp[-j2\pi/3]$, $E_b = 1$, and $Q(l) = 0.2 + j0.1$ find $T(l+1,\sigma(l+1))$ and $\hat{\vec{I}}(l+1,\sigma(l+1))$

Problem 14.4. Prof. Fitz has four daughters who really enjoy sports and usually play on several teams at a time. A typical Saturday has multiple games ongoing at various times. On a particular Saturday the Fitz girls have multiple games ongoing at 8:00AM at locations A, B, and C, at 10:00AM at locations D and E, at 12:00PM at locations F, G, H, I, and at 2:00PM at locations J, K, and L. Professor Fitz would like to make it to one game at each time and yet since he is very cheap he would like to drive the shortest distance from home getting to the games and back home again. The important distances are given in Table 14.3. Prof. Fitz was completely flummoxed by this problem but you immediately realize that the problem can be solved with the Viterbi algorithm.

a) Identify the trellis diagram (a directed graph) for this problem of getting from home to four sporting events and back home again. Associate a distance with each edge of the directed graph.

b) Enumerate out an algorithm for finding the shortest path that does not require a brute force search of all the possible game times and locations. Using this algorithm find the shortest path to see four games and provide the locations where the four games will be watched.

Problem 14.5. CC all zeros
Problem 14.6. CC reduced state
Problem 14.7. TCM all zeros
Problem 14.8. TCM reduced state

	A	B	C	D	E	F	G	H	I	J	K	L
$Home$	5	3	7							6	9	1
A				2	4							
B				1	7							
C				6	8							
D						1	1	11	9			
E						8	3	5	9			
F										5	5	8
G										7	6	5
H										1	2	4
I										5	2	2

Table 14.3: The important distances to sporting events for the Fitz daughters' Saturday games.

14.6 Mini-Projects

14.6.1 Project 1

Project Goals

1. Engage in an implementation of a Viterbi demodulator for an orthogonal modulation with memory (OMWM).

Consider a OMWM consisting of a four state trellis code (N_s=4) using 4PAM modulation (M_s=4) as proposed by Ho, Cavers and Varaldi [HCV93]. The modulation updates are given in Table 14.4. Assume $K_b = 198$, $E_b = 1$. The OMWM produces a frame of $N_f = 200$ symbols. Assume further that the modulation starts with $\sigma(1) = 1$ and that 2 symbols are used at the end to terminate the trellis at $\sigma(201) = 1$.

a) Draw a trellis diagram and identify the additive metric for each branch used in the Viterbi algorithm implementation of MLWD.

b) Detail the termination sequence that will happen for each of the modulation states.

c) Describe the maximum likelihood sequence demodulation given the 200 matched filter outputs. In particular use the data in `project1.mat` and produce the MLSD word estimate. The file contains the 200 matched filter outputs.

$\sigma(l+1) = g_1\left(I(l), \sigma(l)\right)$

	$I(l)$	
State, $\sigma(l)$	0	1
1	1	2
2	3	4
3	1	2
4	3	4

$J(l) = g_2\left(I(l), \sigma(l)\right)$

	$I(l)$	
State, $\sigma(l)$	0	1
1	0	2
2	3	1
3	2	0
4	1	3

$D_z(l) = a(J(l))$

$J(l)$	$D_z(l)$
0	$-3/\sqrt{5}$
1	$-1/\sqrt{5}$
2	$1/\sqrt{5}$
3	$3/\sqrt{5}$

Table 14.4: The finite state machine description of an example trellis code.

14.6.2 Project 2

Project Goals

1. Engage in an implementation of a Viterbi demodulator for an OMWM.
2. Look at space–time trellis coding as a specific example of an OMWM.

In this problem we examine a packet based wireless transmission system with M-ary linear modulation that uses multiple transmit antennas. The complex envelope of the transmitted signal from the i^{th} antenna has the form

$$X_{zi}(t) = \sum_{l=1}^{N_f} D_i(l)u(t-(l-1)T) \tag{14.73}$$

where $D_i(l)$ is a complex modulation symbol with $|D_i| \leq 1$, N_f is the length of the packet, and $u(t)$ is the pulse shape. The received signal has the form

$$Y_z(t) = \sum_{i=1}^{L_t} c_i X_{zi}(t) + W(t) \tag{14.74}$$

where c_i are the known complex channel gains for the i^{th} antenna and $W(t)$ is a complex AWGN of one-sided spectral density of N_0, i.e. $R_W(\tau) = N_0\delta(\tau)$.

Assume that $L_t = 2$ and recall from Problem 9.10 that space–time transmission can be viewed as a linear modulation where the constellation at time l is a function of $D_i(l)$ and $c_i \quad i = 1, 2$. Consider the $R = 1$ bit/symbol 8 state space–time trellis code proposed by Yan and Blum [YB02]:

$\sigma(l+1) = g_1\left(I(l), \sigma(l)\right)$

	$I(l)$	
State, $\sigma(l)$	0	1
1	1	2
2	3	4
3	5	6
4	7	8
5	1	2
6	3	4
7	5	6
8	7	8

$\left(D_1(l)\ D_2(l)\right) = g_2\left(I(l), \sigma(l)\right)$

	$I(l)$	
State, $\sigma(l)$	0	1
1	(11)	$(-1-1)$
2	$(1-1)$	(-11)
3	(-11)	$(1-1)$
4	$(-1-1)$	(11)
5	$(-1-1)$	(11)
6	(-11)	$(1-1)$
7	$(1-1)$	(-11)
8	(11)	$(-1-1)$

Assuming a frame of 200 symbols in length where $\sigma(1) = 1$ and that 3 symbols are used at the end to terminate the trellis at $\sigma(201) = 1$.

a) Define the VA algorithm by identifying the effective constellation points as a function of the channel gains, c_1 and c_2 and $(D_1(l)\ D_2(l))$. Draw a trellis diagram and identify the additive metric for each branch used in the Viterbi algorithm implementation of MLWD.

b) Describe the maximum likelihood word demodulation given the 200 matched filter outputs and the known channel state. In particular use the data in `framemfstc.mat` and produce the MLWD estimate. The file contains both the channel coefficients and the 200 matched filter outputs.

Bibliography

[Ae72] M. Abramowitz and I. E. Stegun (eds). *Handbook of Mathematical Functions*. U. S. Department of Commerce, Washington, DC, 1972.

[AF70] K. Abend and B.D. Fritchman. Statistical detection for communication channels with intersymbol interface. *Proc. IEEE*, vol. 58:779–785, 1970.

[Agi00] Engineers of Agilent. Using error vector magnitude measurements to analyze and troubleshoot vector-modulated signals. Agilent product note 89400-14, Agilent Technologies, 2000.

[Ala98] S. M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Trans. on Select Areas in Comm.*, 16(8):1451–1458, October 1998.

[Arm35] Edwin A. Armstrong. A method of reducing disturbances in radio signaling by a system of frequency modulation. In *IRE Conference*, New York, November 1935.

[Aus67] M.E. Austin. Decision-feedback equalization for digital communication over dispersive channels, 1967.

[BB99] S. Benedetto and E. Biglieri. *Principles of Digital Transmission*. Klewer Academic Press, New York, 1999.

[BCJR74] L.R. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. Info. Theory*, vol. IT-20:284–287, 1974.

[BDMS91] E. Biglieri, D. Divaslar, P.J. McLane, and M.K. Simon. *Introduction to Trellis-Coded Modulations with Applications*. MacMillan, New York, NY, 1991.

[Bel57] R. E. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, 1957.

[BG96] C. Berrou and A. Glavieux. Near optimum error correcting coding and decoding: turbo codes. *IEEE Trans. Commun.*, COM-44:1261–1271, October 1996.

[BGT93] C. Berrou, A. Glavieux, and P. Thitmajshima. Near Shannon limit error correcting coding and decoding: turbo codes. In *IEEE Internation Conference on Communications*, pages 1064–1070, Geneva, Switzerland, June 1993.

[Big84] E. Biglieri. High level modulation and coding for for nonlinear satellite channels. *IEEE Trans. Commun.*, vol. 32:616–626, May 1984.

[Big86] E. Biglieri. Ungerboeck codes do not shape the signal power spectrum. *IEEE Info. Theory*, IT-32:595–596, July 1986.

[Bra95] John Bray. *The Communications Miracle*. Plenum, New York, 1995.

[Bre59] D.G. Brennan. Linear diversity combining techniques. *Proceedings of the IRE*, 47:1075–1102, 1959.

[Bur04] Russel W. Burns. *Communications: A International History of the Formative Years*. Institute of Electrical Engineers, Herts, UK, 2004.

[CAC01] K. Chugg, A. Anastasopoulos, and X. Chen. *Iterative Detection*. Kluwer Academic Press, Boston, 2001.

[Car26] J. R. Carson. *Electric Circuit Theory and Operational Calculus*. McGraw-Hill, New York, 1926.

[CH66] R.W. Chang and J.C. Hancock. On receiver structures for channels having memory. *IEEE Info. Theory*, IT-12:463–468, October 1966.

[Cha66] R. W. Chang. Synthesis of band-limited orthogonal signal for multichannel data transmission. *Bell Systems Technical Journal*, vol. 45:1775–1796, December 1966.

[Che52] H. Chernoff. A measure of the asymptotic efficiency for tests of a hypothesis based on a sum of observations. *Ann. Math Stat.*, 23:493–507, 1952.

[Com04] Federal Communication Commission. Telecommunication rules and regulations. Technical Report Title 47 of the Code of Federal Regulations, Government Printing Office, `http://wireless.fcc.gov/rules.html`, 2004.

[Cou93] L. W. Couch. *Digital and Analog Communication Systems*. MacMillan, New York, 1993.

[Cra91] J. W. Craig. A new simple and exact result for caluclating the probability of error for two dimensional signal constellations. In *IEEE Military Communication Conference*, pages 571–575, Nov 1991.

[CT92] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley Interscience, New York, 1992.

[DR87] W. B. Davenport and W. L. Root. *An Introduction to the Theory of Random Signals and Noise*. IEEE Press, New York, 1987.

[FGF99] I. Fevrier, S. B. Gelfand, and M.P. Fitz. Reduced complexity decision feedback equalization for high-speed communications over sparse channels. *IEEE Trans. Commun.*, COM-46:927–937, June 1999.

[FGL+84] G.D. Forney, R.G. Gallager, G.R. Lang, F.M. Longstaff, and S.U. Qureshi. Efficient modulation for band-limited channels. *IEEE JSAC*, vol. SAC-2:632–647, 1984.

[For72] G.D. Forney. Maximum likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *IEEE Trans. Inform. Theory*, vol. IT-18:363–378, 1972.

[For73] G.D. Forney. The Viterbi algorithm. *Proc. IEEE*, Vol. 61:268–278, 1973.

[For91] G. D. Forney. Geometrically uniform codes. *IEEE Info. Theory*, 37:1241–1960, September 1991.

[Fre98] B. Frey. *Graphical Models for Machine Learning and Digital Communication*. MIT Press, Cambridge, MA, 1998.

[Gal62] Robert Gallagher. Low density parity check codes. *IEEE Info. Theory*, IT-8:21–28, January 1962.

[Gal01] R. Gallagher. *Discrete Stochastic Processes*. Klewer, 2001.

[Gra53] F. Gray. Pulse code communications. Technical report, US Patent Office, 1953.

[Gre81] U. Grenander. *Abstract Inference*. Wiley, New York, 1981.

[Har28] R. V. L. Hartley. Transmission of information. *Bell Systems Technical Journal*, page 535, July 1928.

[Hay83] S. Haykin. *Communications Systems*. John Wiley and Sons, New York, 1983.

[HCR82] J. F. Hayes, T. M. Cover, and J. B. Riera. Optimal sequence detection and optimal symbol-by-symbol detection: Similar algorithms. *IEEE Trans. Commun.*, vol. COM-30:152–157, January 1982.

[HCV93] P. Ho, J. Cavers, and J. Varaldi. The effect of constellation density on trellis coded modulation in fading channels. *IEEE Trans. Veh. Technol.*, vol. VT-42:318–325, 1993.

[Hel91] C.W. Helstrom. *Probability and Stochastic Processes for Engineers*. Macmillan, New York, 1991.

[Huu03] Anton A. Huurdeman. *The Worldwide History of Telecommunications*. Wiley-Interscience, Hoboken, NJ, 2003.

[KH97] E.W. Kamen and B.S. Heck. *Fundamentals of Signals and Systems Using Matlab*. Prentice Hall, Englewood Cliffs, NJ, 1997.

[Kot60] V. A. Kotel'nikov. *Theory of optimum noise immunity*. McGraw Hill Book Company, New York, 1960.

[KR04] James F. Kurose and Keith W. Ross. *Computer Networking: A Top-Down Approach Featuring the Internet*. Addison Wesley, 2004.

[KSH00] T. Kailath, A. Sayed, and B. Hassibi. *Linear Estimation*. Prentic Hall, Upper Saddle River, NJ, 200.

[LC04] S. Lin and D. Costello. *Error Control Coding*. Prentice Hall, Englewood Cliffs, NJ, 2004.

[Leh86] E.L. Lehmann. *Testing Statistical Hypothesis*. Wiley Interscience, New York, 1986.

[Les69] Lawrence Lessing. *Man of High Fidelity*. Bantam, New York, 1969.

[LG89] A. Leon-Garcia. *Probability and Random Processes for Electrical Enginering*. Addison - Wesley, New York, 1989.

[LGW00] A Leon-Garcia and I. Widjaja. *Communication Networks*. McGraw Hill, Boston, 2000.

[LT85] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, New York, 1985.

[Luc65] R. W. Lucky. Automatic equalization for digital communications. *Bell System Tech J.*, vol. 44:547–588, April 1965.

[May84] Robert J Mayhan. *Discrete-Time and Continuous-Time Linear Systems*. Addison-Wesley, New York, 1984.

[MH94] U. Madhow and M.L. Honig. MMSE interference suppression for direct sequence spread spectrum CDMA. *IEEE Trans. Commun.*, vol. 42:3178–3188, Dec. 1994.

[Mid60] David Middleton. *Introduction to Statistical Communication Theory*. Peninsula, Los Altos, CA, 1960.

[Mil63] A. Miller. Transmission system. Technical report, US Patent Office, Patent 3,108,261, 1963.

[Mil64] K. S. Miller. *Multidimentional Gaussian Distributions*. John Wiley and Sons Inc., New York, 1964.

[Mil74] K. S. Miller. *Complex Stochastic Process*. Addison-Wesley, Reading MA, 1974.

[Mit98] S. K. Mitra. *Digital Signal Processing*. McGraw Hill, 1998.

[Mos96] S. Moshavi. Multi-user detection for DS-CDMA communications. *IEEE Communications Magazine*, pages 124–136, October 1996.

[MWW72] P. L. McAdam, L. Welch, and C.L. Weber. Map bit decoding of convolutional codes. In *Proc. IEEE Int. Symposium on Info. Theory*, 1972.

[MZ60] S.J. Mason and H.J. Zimmersman. *Electronic Circuits, Signal and Systems*. Wiley, New York, 1960.

[Nak60] M. Nakagami. The m-distribution - a general formula of intensity distribution of rapid fading. In W.C. Hoffman, editor, *Statistical Methods in Radio Wave Propagation*, pages 3–36. Pergamon, 1960.

[Nor43] D.O. North. Analysis of factors which determine signal-noise discrimination in pulsed carrier systems. Technical Report PTR-6C, RCA Technical Report, June 1943. Reprinted in Proc. IEEE pp. 1015-1027, July 1963.

[Nyq28] H. Nyquist. Certan topics in telegraph transmission theory. *AIEE Transactions*, vol. 47:617–644, 1928.

[OS99] A.V. Oppenheim and R.W. Schafer. *Digital Signal Processing*. Prentice Hall, Englewood Cliffs, NJ, 1999.

[OW97] A.V. Oppenheim and A.S. Willsky. *Signals and Systems*. Prentice Hall, Upper Saddle River, NJ, 1997.

[Pap84] A. Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw Hill, New York, 1984.

[Pet89] Blake Peterson. Spectrum analysis basics. Application note 150, Hewlett Packard, 1989.

[PFTV86] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes, The Art of Scientific Computing*. Cambridge University Press, Cambridge, 1986.

[PM88] J. G. Proakis and Manolakis. *Introduction to Digital Signal Processing*. McMillan, 1988.

[Poo88] H. V. Poor. *An Introduction to Signal Detection and Estimation*. Springer-Verlag, New York, 1988.

[Por97] Boaz Porat. *A Course in Digital Signal Processing*. Wiley, 1997.

[Pro89] J. G. Proakis. *Digital Communications*. McGraw-Hill, New York, 1989.

[PS94] J. G. Proakis and Masoud Salehi. *Communications System Engineering*. Prentice Hall, Englewood Cliffs, NJ, 1994.

[Ric45] S.O. Rice. Mathematical analysis of random noise. *Bell System Tech J.*, 24:96–157, 1945.

[Ric48] S. O. Rice. Statistical properties of a sine-wave plus random noise. *Bell System Technical Journal*, vol. 27:109–157, 1948.

[SA00] M. Simon and M.-S. Alouini. *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*. John Wiley and Sons, 2000.

[Sch91] L. L. Scharf. *Statistical Signal Processing*. Addison-Wesley, Reading, MA, 1991.

[SD98] M.K. Simon and D. Divsalar. Some new twists to problems involving the Gaussian probability integral. *IEEE Trans. Commun.*, COM-46:200–210, February 1998.

[Sha48] Claude Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 1948.

[Shn67] D.A. Shnidman. A generalized Nyquist criterion an an optimum linear receiver for a pulse modulation system. *Bell System Technical Journal*, vol. 46:2163–2177, November 1967.

[Shn89] D.A. Shnidman. The calculation of the probability of detection and the generalized Marcum's Q-function. *IEEE Info. Theory*, vol. IT-35:389–399, March 1989.

[Skl88] B. Sklar. *Digital Communications*. Prentice-Hall, Englewood Cliffs, NJ, 1988.

[SW49] Claude E. Shannon and Warren Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, IL, 1949.

[SW02] Henry Stark and John W. Woods. *Probability and Random Processes with Applications to Signal Processing*. Prentice Hall, Upper Saddle River, NJ, 2002.

[SW04] Jun Shi and Rick Wesel. Efficient computation of trellis code generating functions. *IEEE Trans. Commun.*, COM-52:219–227, February 2004.

[Tan02] Andrew S. Tanenbaum. *Computer Networks, Fourth Edition*. Prentice Hall, 2002.

[Ung74] G. Ungerboeck. Adaptive maximum likelihood receiver for carrier modulated data transmission systems. *IEEE Trans. Commun.*, COM-22 No. 5, May 1974.

[Ung82] Gottfried Ungerboeck. Channel coding with multilevel/phase signals. *IEEE Info. Theory*, IT-28:55–67, January 1982.

[Ver98] S. Verdu. *Multiuser Detection*. Cambridge Press, Cambridge, UK, 1998.

[Vit67] A.J. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans. Info. Theory*, vol. IT-13:260–269, 1967.

[VO79] A. J. Viterbi and J. K. Omura. *Principles of Digital Communication and Coding*. McGraw Hill, New York, 1979.

[VT68] H. L. Van Trees. *Detection, Estimation, and Modulation Theory.* John Wiley and Sons, New York, 1968.

[Wal97] H.R. Walker. VPSK and VMSK modulation transmit digital audio and video at 15 bit/s/Hz. *IEEE Trans. on Broadcasting*, 43:96–103, 1997.

[Web87] C. L. Weber. *Elements of Detection and Signal Design.* Springer Verlag, New York, 1987. Reprint of a 1967 book.

[Wes01] Douglas B. West. *Introduction to Graph Theory.* Prentice Hall, 2001.

[Wic95] S. B. Wicker. *Error Control Systems for Digital Communications and Storage.* Prentice Hall, Upper Saddle River, NJ, 1995.

[Wie49] Norbert Wiener. *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications.* Wiley, New York, 1949.

[WJ65] J. M. Wozencraft and I. M. Jacobs. *Principles of Communications Engineering.* John Wiley and Sons, New York, 1965.

[XSR90] Z. Xie, R. Short, and C. Rushforth. A family of subotimum detectors for coherent multiuser communications. *IEEE J. Select. Areas Commun.*, vol. 8:683–690, May 1990.

[YB02] Q. Yan and R. S. Blum. Improved space-time convolutional codes for quasistatic slow fading channels. *IEEE Transactions on Wireless Communications*, pages 563–571, Oct 2002.

[Zie86] A. Van Der Ziel. *Noise in Solid State Devices.* Wiley-Interscience, New York, 1986.

[ZTF89] R.E. Ziemer, W.H. Trantor, and D.R. Fannin. *Signals and Systems: Continuous and Discrete.* MacMillan, New York, 1989.

Appendix A

Detection Theory

This appendix summarizes several important results in the theory of optimal detection. This theory complements a graduate treatment of digital communications. Specifically this Appendix considers the optimum demodulation of one symbol of a binary digital modulation in additive Gaussian noise. While this case is far from general, it illustrates many important issues. The results presented here are derived from Bayesian decision/detection theory. The derivations presented are limited to binary modulations but the concepts are easily generalized for M-ary modulations. More details and examples for this material can be obtained from a variety of textbooks [VT68, Poo88].

A.1 Binary Hypothesis Testing

In binary digital communications, a demodulator's job is to make an optimum or near optimum estimate of which of two signals were transmitted by examining distorted versions of the transmitted waveform. This model only considers one symbol of a binary modulated waveform, but the results are easily extended to sequences of modulated symbols. Ω is the signal space with two mutually exclusive and collectively exhaustive sets of elements corresponding to the two possible hypotheses, $\vec{M} \in \Omega_0$ or $\vec{M} \in \Omega_1$. The vector $\vec{M}$ corresponds to a possible parameterization of the transmitted signal, i.e., $\vec{x}\left(\vec{M}\right)$ or $x\left(t, \vec{M}\right)$. In general, $\vec{M}$ is a nuisance parameter associated with the transmitted signal (an example is the random phase in noncoherent detection see Section A.4). Associated with the transmitted signal space is an *a priori* probability of the symbol being transmitted, denoted $\pi_i = P\left(\vec{M} \in \Omega_i\right),\ i = 0, 1.$ Ω_Y is the observation space (i.e., the demodulator input) we denote the observations with $\vec{Y}$ or $Y(t)$. The distortion of the channel can be represented by conditional probability distribution[1] $P\left(\vec{Y} \middle| \vec{M}\right)$ defined on Ω_Y. The decision space has two elements, H_0 and H_1, corresponding to the two possible transmitted waveforms. The decision process is represented by a probabilistic decision rule denoted $\delta\left(H_i \middle| \vec{Y}\right)$. Two important characteristics of this probabilistic decision rule are

$$0 \leq \delta\left(H_0 \middle| \vec{Y}\right) \leq 1 \tag{A.1}$$

and

$$\delta\left(H_0 \middle| \vec{Y}\right) + \delta\left(H_1 \middle| \vec{Y}\right) = 1. \tag{A.2}$$

[1] Since we want to keep the discussion generic we use $P\left(\vec{Y} \middle| \vec{M}\right)$ to represent either a conditional PDF or a conditional PMF.

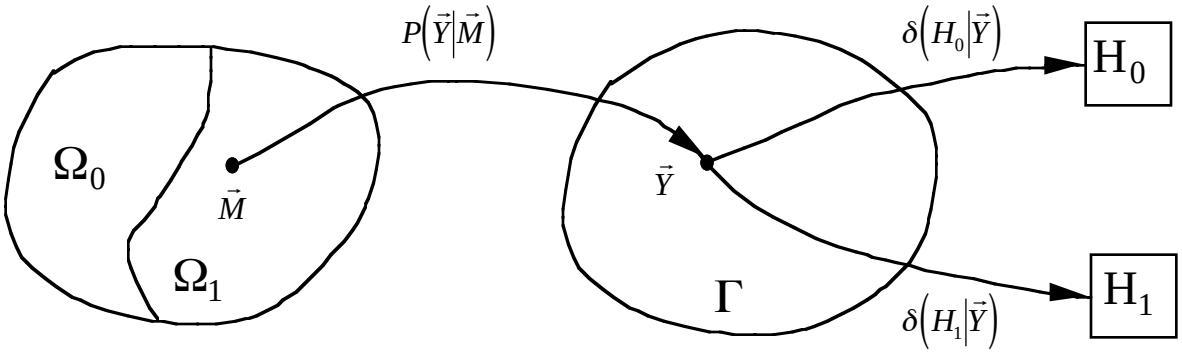


Figure A.1: Diagram of the binary detection mathematical model.

Fig. A.1 shows the mathematical model for this decision process. The model has been developed without specifications on the signal and observation spaces and the nuisance parameters. These parameters are specified as needed in the following sections.

Given this mathematical model, the minimum probability of bit error decoding rule can be derived. The goal in this section is to choose a decision rule as a function of an observed $\vec{Y} = \vec{y}$ that minimizes the bit error probability (BEP). Using total probability, the BEP is expressed as

$$\begin{aligned} P_B(E) &= P\left(H_0, \vec{M} \in \Omega_1\right) + P\left(H_1, \vec{M} \in \Omega_0\right) \\ &= \pi_1 \int_{\Omega_Y} \delta\left(H_0 \middle| \vec{y}\right) P_{\vec{Y}}\left(\vec{y} \middle| \vec{M} \in \Omega_1\right) d\vec{y} + \pi_0 \int_{\Omega_Y} \delta\left(H_1 \middle| \vec{y}\right) P_{\vec{Y}}\left(\vec{y} \middle| \vec{M} \in \Omega_0\right) d\vec{y} \end{aligned} \tag{A.3}$$

Using (A.3), the BEP is expressed as

$$\begin{aligned} P_B(E) &= \pi_1 + \int_{\Omega_Y} \delta\left(H_1 \middle| \vec{y}\right) \left[\pi_0 P_{\vec{Y}}\left(\vec{y} \middle| \vec{M} \in \Omega_0\right) - \pi_1 P_{\vec{Y}}\left(\vec{y} \middle| \vec{M} \in \Omega_1\right)\right] d\vec{y} \\ &= \pi_1 + \int_{\Omega_Y} \delta\left(H_1 \middle| \vec{y}\right) g\left(\vec{y}\right) d\vec{y} \end{aligned} \tag{A.4}$$

where

$$g\left(\vec{y}\right) = \left[\pi_0 P_{\vec{Y}}\left(\vec{y} \middle| \vec{M} \in \Omega_0\right) - \pi_1 P_{\vec{Y}}\left(\vec{y} \middle| \vec{M} \in \Omega_1\right)\right]. \tag{A.5}$$

Realize that $\pi_i \geq 0$ and is not a function of $\delta\left(H_i \middle| \vec{y}\right)$, so the BEP is minimized by making the integral in (A.5) as small as possible. Since the decision rule satisfies $0 \leq \delta\left(H_i \middle| \vec{y}\right) \leq 1$, the integral is minimized by making $\delta\left(H_i \middle| \vec{y}\right) = 0$ when $g\left(\vec{y}\right) > 0$, $\delta\left(H_i \middle| \vec{y}\right) = 1$ when $g\left(\vec{y}\right) < 0$, and $\delta\left(H_i \middle| \vec{y}\right)$ arbitrary when $g\left(\vec{y}\right) = 0$. This decision rule is expressed as

$$\pi_0 P_{\vec{Y}}\left(\vec{y} \middle| \vec{M} \in \Omega_0\right) \underset{H_1}{\overset{H_0}{\gtrless}} \pi_1 P_{\vec{Y}}\left(\vec{y} \middle| \vec{M} \in \Omega_0\right) \tag{A.6}$$

It is interesting to note that the decoding model allows for a probabilistic decision rule but the resulting decision rule is a deterministic function of $\vec{y}$. If both sides of are divided[2] by $P_{\vec{Y}}\left(\vec{y}\right)$ and noting

$$P\left(H_j \middle| \vec{Y} = \vec{y}\right) = \frac{\pi_j P_{\vec{Y}}\left(\vec{y} \middle| \vec{M} \in \Omega_j\right)}{P_{\vec{Y}}\left(\vec{y}\right)}, \tag{A.7}$$

[2] $P_{\vec{Y}}\left(\vec{y}\right)$ is nonzero for any $\vec{Y} = \vec{y}$ observed at the demodulator input.

the decoding rule in (A.6) is equivalent to a rule that chooses the most likely hypothesis given the observed $\vec{y}$. This rule, i.e.,

$$P\left(H_0 \left| \vec{Y} = \vec{y} \right.\right) \underset{H_1}{\overset{H_0}{\gtrless}} P\left(H_1 \left| \vec{Y} = \vec{y} \right.\right) \tag{A.8}$$

is referred to as the maximum *a posteriori* (MAP) decision rule.

A.2 Karhunen Loève Expansion

The Karhunen-Loève (KL) expansion is a mathematical tool that permits a random process to be represented using random vector spaces. The expansion first uses Mercer's Theorem to show that a covariance function can be expanded into a sum involving the associated eigenfunctions. These eigenfunctions are then the basis for the vector space (KL) expansion of the random process. This expansion is theoretically advantageous since it permits a derivation of the optimal demodulators for digital communication signals. Unfortunately, the practicality of the expansion is limited due to the difficulty in finding the appropriate set of orthonormal basis function. A good treatment of the KL expansion is given in [VT68].

Theorem A.1 *A covariance function, $R_N(\tau)$, is a non-negative definite kernel for any $g \in L^2[a,b]$*[3]

Proof: Without loss of generality assume N(t) is a zero mean random process. A non-negative definite kernel implies

$$K_N(g,g) = \int_a^b \int_a^b g^*(t) R_N(t,s) g(s) ds dt \geq 0. \tag{A.9}$$

Using the definition of the covariance function and that $g \in L^2[a,b]$ gives

$$\begin{aligned} K_N(g,g) &= \int_a^b g^*(t) \int_a^b E[x(t)x^*(s)] g(s) ds dt \\ &= E\left[\left| \int_a^b g^*(t) x(t) dt \right|^2 \right] \geq 0. \qquad \square \end{aligned} \tag{A.10}$$

Theorem A.2 (Mercer) *If a linear transformation generated by the continuous Hermitian kernel $R_N(t_1,t_2)$ is non-negative definite, then $R_N(t_1,t_2)$ can be expanded into the uniformly convergent series*

$$R_N(t_1,t_2) = \sum_{k=1}^{\infty} \lambda_k \Psi_k(t_1) \Psi_k^*(t_2) \qquad t_1, t_2 \in [a,b] \tag{A.11}$$

where $\Psi_k(t_i)$'s are the orthonormal eigenfunctions of $R_N(t_1,t_2)$ and λ_k's are the eigenvalues which are solutions of the following integral equation:

$$\lambda_k \Psi_k(t_1) = \int_a^b R_N(t_1,t_2) \Psi_k^*(t_2) dt_2 \qquad a \leq t_1 \leq b. \tag{A.12}$$

[3] $g \in L^2[a,b]$ implies that $\int_a^b |g(t)|^2 dt < \infty$.

Proof: See [Gre81] □

Theorem A.3 (Karhunen-Loève Expansion) *Suppose that $N(t), t \in [a,b]$ is a complex, zero mean, second order random process with covariance function $R_N(t_1,t_2)$ continuous on $[a,b]\times[a,b]$ (i.e., $N(t)$ is mean square continuous on $[a,b]$), then $N(t)$ can be expanded into the mean square convergent series*

$$N(t) = \sum_{k=1}^{\infty} N_k \Psi_k(t) \qquad a \le t \le b \tag{A.13}$$

where

$$N_k = \int_a^b N(t)\Psi_k^*(t)dt. \tag{A.14}$$

Proof: Recall mean square convergence implies

$$\lim_{N\to\infty} E\left[|N(t) - \sum_{k=1}^{N} N_k\Psi_k(t)|^2\right] = 0. \tag{A.15}$$

This limit can be broken into three terms

$$\begin{aligned} E\left[|N(t) - \sum_{k=1}^{N} N_k\Psi_k(t)|^2\right] &= E\left[N(t)N^*(t)\right] - 2\Re\left\{E\left[N(t)\sum_{k=1}^{N} N_k^*\Psi_k^*(t)\right]\right\} \\ &\quad + E\left[\sum_{j=1}^{N} N_j\Psi_j(t)\sum_{k=1}^{N} N_k\Psi_k(t)\right] \\ &= Z_1 - Z_2 + Z_3. \end{aligned} \tag{A.16}$$

These three components are reduced as

1.
$$Z_1 = \lim_{N\to\infty} E\left[N(t)N^*(t)\right] = R_N(t,t) \tag{A.17}$$

2.
$$Z_2 = \lim_{N\to\infty} 2\Re\left\{E\left[N(t)\sum_{k=1}^{N} N_k^*\Psi_k^*(t)\right]\right\}. \tag{A.18}$$

Using (A.14) gives

$$\begin{aligned} Z_2 &= \lim_{N\to\infty} 2\Re\left\{E\left[N(t)\sum_{k=1}^{N}\int_a^b N^*(s)\Psi_k(s)ds\Psi_k^*(t)\right]\right\} \\ &= \lim_{N\to\infty} 2\Re\left\{\sum_{k=1}^{N}\int_a^b E\left[N(t)N^*(s)\right]\Psi_k(s)ds\Psi_k^*(t)\right\}. \end{aligned} \tag{A.19}$$

By Mercer's Theorem this is given as

$$Z_2 = \lim_{N\to\infty} 2\Re\left\{\sum_{k=1}^{N}\int_a^b \sum_{i=1}^{\infty}\lambda_i\Psi_i(t)\Psi_i^*(s)\Psi_k(s)ds\Psi_k^*(t)\right\}. \tag{A.20}$$

The orthogonality of the basis functions and Mercer's theorem reduce this to

$$Z_2 = \lim_{N\to\infty} 2\Re\left\{\sum_{k=1}^{N}\lambda_k\Psi_k(t)\Psi_k^*(t)\right\} = 2R_N(t,t). \tag{A.21}$$

3.

$$\begin{aligned} Z_3 &= \lim_{N\to\infty} E\left[\sum_{j=1}^{N} N_j\Psi_j(t)\sum_{k=1}^{N} N_k^*\Psi_k^*(t)\right] = \lim_{N\to\infty}\sum_{j=1}^{N}\Psi_j(t)\sum_{k=1}^{N} E\left[N_jN_k^*\right]\Psi_k^*(t) \\ &= \lim_{N\to\infty}\sum_{k=1}^{N}\lambda_k\Psi_k(t)\Psi_k^*(t). \end{aligned} \tag{A.22}$$

Using these three results give the limit as

$$\lim_{N\to\infty} E\left[|N(t)-\sum_{k=1}^{N} N_k\Psi_k(t)|^2\right] = -R_N(t,t) + \lim_{N\to\infty}\sum_{k=1}^{N}\lambda_k\Psi_k(t)\Psi_k^*(t) \tag{A.23}$$

which is zero as a consequence of Mercer's Theorem. □

A.2.1 Characteristics of the KL Expansion

1. The KL expansion coefficients are orthogonal random variables, i.e.,

$$E\left[N_iN_j^*\right] = E\left[\int_a^b N(t_1)\Psi_i^*(t_1)dt_1\int_a^b N(t_2)\Psi_j^*(t_2)dt_2\right] \tag{A.24}$$

$$= \int_a^b \Psi_i^*(t_1)\int_a^b R_N(t_1,t_2)\Psi_j(t_2)dt_2dt_1 \tag{A.25}$$

$$= \int_a^b \Psi_i^*(t_1)\lambda_j\Psi_j(t_1)dt_1 = 0 \qquad \forall\, i\neq j. \tag{A.26}$$

   This characteristic is the power of the KL expansion compared to other orthonormal expansions.

2. The function $\Psi_k(t)$ (the eigenfunctions) define a coordinate system much the same way as eigenvectors do in linear algebra.

3. The eigenvalues are the average energy in that coordinate, i.e., $E[|N_i|^2] = \lambda_i$. This implies that λ_i is real and since $R_N(t_1,t_2)$ is non-negative definite $\lambda_i \geq 0$.

4. The sum of the eigenvalues is the average energy of $N(t)$ on the interval $[a,b]$

$$E\left[\int_a^b |N(t)|^2dt\right] = \int_a^b R_N(t,t)dt = \int_a^b\sum_{k=1}^{\infty}\lambda_k|\Psi_k(t)|^2dt = \sum_{k=1}^{\infty}\lambda_k. \tag{A.27}$$

5. If $N(t)$ is a Gaussian process then the coefficients of the KL expansion are statistically independent Gaussian random variables. This characteristic of the KL expansion is due to linear transformations of a Gaussian random process producing another Gaussian random process and the fact that orthogonality and independence are equivalent for Gaussian random variables. Consequently the KL expansion reduces a Gaussian random process to a Gaussian random vector with independent components. This proves to be the characteristic most valuable in the derivation of optimum signal processors.

6. If $R_N(t_1, t_2)$ is positive definite then the $\psi_i(t)$ are a set of complete orthonormal basis functions. Completeness implies that any deterministic signal, $s(t)$, defined on $[a, b]$ can be represented by the $\psi_i(t)$ as

$$s(t) = \sum_{k=1}^{\infty} s_k \psi_k(t) \qquad a \leq t \leq b \tag{A.28}$$

with

$$s_k = \int_a^b s(t)\psi_k^*(t)dt. \tag{A.29}$$

It should be noted that equality is in the L^2 sense.

Example A.1: The correlation function of a sinusoid with a random phase has the form

$$R_N(t_1, t_2) = \cos(2\pi f(t_1 - t_2)). \tag{A.30}$$

For simplicity take [a,b] = $[0, \frac{1}{f}]$ then it is easy to show that only two nonzero eigenvalues exist and the associated eigenfunctions are

$$\psi_1(t) = \sqrt{2f}\cos(2\pi ft) \tag{A.31}$$
$$\psi_2(t) = \sqrt{2f}\sin(2\pi ft) \tag{A.32}$$

with

$$\lambda_1 = \lambda_2 = \frac{1}{2}. \tag{A.33}$$

$N(t)$ is represented as

$$N(t) = N_I\sqrt{2f}\cos(2\pi ft) + N_Q\sqrt{2f}\sin(2\pi ft). \tag{A.34}$$

Example A.2: Recall AWGN is an approximation to a random process with a continuous covariance function. Since AWGN has

$$R_N(t_1, t_2) = N_0\delta(t_1 - t_2) \tag{A.35}$$

it is useful to examine the KL expansion knowing the discontinuous correlation function is only due to the mathematical modeling of wideband noise. The integral equation defining the KL expansion of AWGN is given by

$$\begin{aligned} \lambda_k\psi_k(t_1) &= \int_a^b N_0\delta(t_1 - t_2)\psi_k(t_2)dt_2 & \text{(A.36)} \\ &= N_0\psi_k(t_1). & \text{(A.37)} \end{aligned}$$

Consequently, any orthonormal expansion is the KL expansion for AWGN and the eigenvalues are $\lambda_i = N_0$. It is interesting to note that the KL expansion of AWGN and Mercer's theorem imply that

$$\delta(t_1 - t_2) = \sum_{k=1}^{\infty} \psi_k(t_1)\psi_k(t_2) \qquad t_1, t_2 \in [a, b]. \tag{A.38}$$

Other examples of KL expansions for random processes commonly used in the analysis of communication systems are given in [VT68].

A.3 Coherent Detection

This section will review the optimum demodulator of binary signals in additive Gaussian noise where all parameters of the two possible signals are known. A great deal of insight into the real problems in communication system design is gained from this exercise. A simple model of this problem is given as:

1. $x_0(t)$ and $x_1(t)$ $t \in [0, T_p]$ are the two possible finite energy, complex baseband transmitted signals. There are no nuisance parameters, so $\Omega_j = \{x_j(t)\}$.

2. The channel is modeled as frequency flat with propagation delay τ_p, a carrier phase rotation of ϕ_p, and propagation loss L_p.

3. The received signal is corrupted by an additive complex Gaussian baseband noise, $N_z(t)$, with a positive definite correlation function given by $C_{N_z}(t, u)$.

Consequently, the input to the demodulator has the form

$$Y_z(t) = x_j(t - \tau_p) L_p \exp[j\phi_p] + N_z(t) \tag{A.39}$$

The derivation in the remainder of this section does not discuss several mathematical technicalities (e.g., singular detection), since these technicalities detract from a fluid presentation of the main ideas. A more detailed derivation of optimum decoding procedures is available in [VT68, Poo88].

The minimum BEP decoder satisfies (A.6), but an equivalent form of this decoding rule is obtained by forming the likelihood function

$$L\left(\vec{Y}\right) = \frac{f_{\vec{Y}}(\vec{y}|H_1)}{f_{\vec{Y}}(\vec{y}|H_0)} \tag{A.40}$$

and making a threshold test

$$L\left(\vec{Y}\right) \underset{H_1}{\overset{H_0}{\gtrless}} \frac{\pi_0}{\pi_1}. \tag{A.41}$$

The computing of the likelihood function is trivial if the observations of vectors. For the case of continuous time observations the computing of the likelihood function is more involved and requires both the KL expansion and Grenander's Therorem.

Theorem A.4 (Grenander) *Suppose that $\vec{Y}_N$ is a continuous random vector under H_0 and H_1 with a PDF $f_{\vec{Y}_N}(\vec{y}_N|H_i)$, for each value of N and define a vector likelihood function $L_N\left(\vec{Y}\right)$ on Γ_Y as*

$$L_N\left(\vec{Y}\right) = \frac{f_{\vec{Y}_N}(\vec{y}|H_1)}{f_{\vec{Y}_N}(\vec{y}|H_0)}. \tag{A.42}$$

If $L_N(\vec{Y})$ is nonsingular for each N then

$$\lim_{N\to\infty} L_N\left(\vec{Y}\right) = L\left(\vec{Y}\right) \qquad \textit{in probability.} \tag{A.43}$$

Proof: See [Gre81] □

This theorem enables the computing of the likelihood ratio by first truncating the vector representing $\vec{Y}$ (generated by the KL expansion) and forming a likelihood ratio, then taking the limit.

This technique is easily applied to this coherent detection problem since the received signal has the form in (A.39) and the KL expansion of $N_z(t)$ can be used to represent $Y(t)$ in terms of a sequence of random variables. This representation is given as

$$Y(t) = \lim_{N\to\infty}\sum_{i=1}^{N} Y_i\Psi_i(t) = \lim_{N\to\infty}\sum_{i=1}^{N}\left(L_p e^{j\phi_p}x_{ij} + N_i\right)\Psi_i(t) \tag{A.44}$$

where

$$x_{ij} = \int_a^b x_j(t-\tau_p)\Psi_i^*(t)\,dt. \tag{A.45}$$

Since the noise is Gaussian, the coefficients in this expansion are independent Gaussian random variables with $E(Y_i) = L_p e^{j\phi_p}x_{ij}$ and $\mathrm{var}(Y_i) = \lambda_i$ where λ_i is the i^{th} eigenvalue associated with the correlation function $R_N(t,u)$. Since Y_i, for $i = 1,\ldots,N$ is a Gaussian vector, the corresponding conditional density function, and hence the N term likelihood ratio, can be expressed in closed form for finite N. The conditional PDF of the truncated vector is given by

$$f_{\vec{Y}_n}(\vec{y}_N|H_j) = f_{\vec{Y}_n}(y_1,\ldots,y_N|H_j) = \frac{1}{\pi^N\prod_{i=1}^N\lambda_i}\exp\left(-\sum_{i=1}^{N}\frac{\left|y_i - L_p e^{j\phi_p}x_{ij}\right|^2}{\lambda_i}\right) \tag{A.46}$$

and the likelihood ratio for the truncated vector is given by

$$L_N\left(\vec{Y}\right) = \exp\left(-\sum_{i=1}^{N}\frac{\left|y_i - L_p e^{j\phi_p}x_{i1}\right|^2}{\lambda_i} + \sum_{i=1}^{N}\frac{\left|y_i - L_p e^{j\phi_p}x_{i0}\right|^2}{\lambda_i}\right) \tag{A.47}$$

This can be algebraically reduced to

$$L_N\left(\vec{Y}\right) = \exp\left(2\Re\left\{L_p e^{j\phi_p}\sum_{i=1}^{N}\frac{y_i\left(x_{i1}-x_{i0}\right)^*}{\lambda_i}\right\} + L_p^2\sum_{i=1}^{N}\frac{|x_{i0}|^2-|x_{i1}|^2}{\lambda_i}\right) \tag{A.48}$$

The second term in the exponent is not a function of $\vec{Y}$ so the optimum demodulator consists of the following test

$$\exp\left(2\Re\left\{L_p e^{j\phi_p}\sum_{i=1}^{N}\frac{y_i\left(x_{i1}-x_{i0}\right)^*}{\lambda_i}\right\}\right) \underset{H_0}{\overset{H_1}{\gtrless}} R\left(\vec{x}_1,\vec{x}_0,\vec{\Lambda},L_p,\pi_0\right) \tag{A.49}$$

At this point the limit of this ratio can be taken and by Grenander's Theorem, this limit converges to the optimum demodulator. A more useful form for the limit is obtained by Pitcher's Theorem.

Theorem A.5 (Pitcher) *Defining*

$$q_j(t) = \lim_{N\to\infty}\sum_{i=1}^{N}\frac{x_{ij}}{\lambda_i}\Psi_i(t) = \lim_{N\to\infty}\sum_{i=1}^{N}q_{ij}\Psi_i(t) \tag{A.50}$$

and assuming that the support of $x_j(t-\tau_p)$ is contained in $[a,b]$, then

i) The optimum filter is defined by the integral equation

$$\int_a^b R_{N_z}(t,u)\,q_j(u)\,du = x_j(t-\tau_p), \tag{A.51}$$

ii) The optimum filtering operation is

$$\lim_{N\to\infty}\sum_{i=1}^{N}\frac{y_i x_{ij}^*}{\lambda_i} = \int_a^b y_z(t)\,q_j^*(t)\,dt, \tag{A.52}$$

iii) The optimum threshold in the threshold test is

$$\begin{aligned}\lim_{N\to\infty}R\left(\vec{x}_1,\vec{x}_0,\vec{\Lambda},L_p,\pi_0\right) &= \exp\left(L_p^2\left(\int_a^b x_1(t-\tau_p)\,q_1^*(t)\,dt - \int_a^b x_0(t-\tau_p)\,q_0^*(t)\,dt\right)\right)\frac{\pi_0}{\pi_1}\\ &= R\left(x_1(t),x_0(t),R_{N_z}(\tau),L_p,\pi_0\right).\end{aligned} \tag{A.53}$$

Proof:

i) Using Mercer's theorem in the integral gives

$$\begin{aligned}\int_a^b R_N(t,u)\,q_j(u)\,du &= \int_a^b\sum_{k=1}^{\infty}\lambda_k\Psi_k(t)\,\Psi_k^*(u)\sum_{i=1}^{\infty}\frac{x_{ij}}{\lambda_i}\Psi_i(u)\,du\\ &= \sum_{k=1}^{\infty}\lambda_k\Psi_k(t)\sum_{i=1}^{\infty}\frac{x_{ij}}{\lambda_i}\int_a^b\Psi_k^*(u)\,\Psi_i(u)\,du\\ &= \sum_{k=1}^{\infty}x_{kj}\Psi_k(t)\\ &= x_j(t-\tau_p)\end{aligned} \tag{A.54}$$

ii)

$$\begin{aligned}\lim_{N\to\infty}\sum_{i=1}^{N}\frac{y_i x_{ij}^*}{\lambda_i} &= \lim_{N\to\infty}\sum_{i=1}^{N}\int_a^b y_z(t)\,\Psi_i^*(t)\,dtq_{ij}^* \\ &= \int_a^b y_z(t)\,q_j^*(t)\,dt \end{aligned} \tag{A.55}$$

iii)

$$\begin{aligned}\lim_{N\to\infty}\sum_{i=1}^{N}\frac{x_{ij}x_{ij}^*}{\lambda_i} &= \lim_{N\to\infty}\sum_{i=1}^{N}\int_a^b x_j(t-\tau_p)\,\Psi_i^*(t)\,dtq_{ij}^* \\ &= \int_a^b x_j(t-\varepsilon)\,q_j^*(t)\,dt. \qquad \square \end{aligned} \tag{A.56}$$

This property implies that the optimum receiver has the form

$$\exp\left(2L_p\Re\left\{e^{-j\phi_p}\int_a^b y_z(t)\left[q_1(t)-q_0(t)\right]^*dt\right\}\right)\underset{H_0}{\overset{H_1}{\gtrless}} R\left(x_1(t),x_0(t),R_{N_z}(\tau),L_p,\pi_0\right). \tag{A.57}$$

Since the natural log is a monotonic function, taking the log of both sides of the threshold test simplifies the expression. Consequently the optimum demodulator can equivalently calculate the log-likelihood function and compare it to a threshold. The optimum demodulator satisfies the following modified threshold test

$$\Re\left\{e^{-j\phi_p}\int_a^b y_z(t)\left[q_1(t)-q_0(t)\right]^*dt\right\}\underset{H_0}{\overset{H_1}{\gtrless}}\frac{1}{2L_p}\log\left[R\left(x_1(t),x_0(t),R_N(\tau),L_p,\pi_0\right)\right]=R. \tag{A.58}$$

It is worth noting that

$$R=\frac{L_p}{2}\left(\int_a^b u_1(t-\tau_p)\,q_1^*(t)\,dt-\int_a^b u_0(t-\tau_p)\,q_0^*(t)\,dt\right)+\frac{1}{2L_p}\log\left(\frac{\pi_0}{\pi_1}\right). \tag{A.59}$$

The resulting form of the optimum demodulator computes the correlation between the received waveform and a waveform that is a function of the pulse shapes and the noise statistics. This is a generalized matched filtering operation. The output of this correlation is derotated (in phase) by ϕ_p. The real part of the derotated value is compared to a threshold to produce the binary decisions. It should be noted that

$$Q_1 = \int_a^b Y_z(t)\,q_1^*(t)\,dt \tag{A.60}$$

$$Q_0 = \int_a^b Y_z(t)\,q_0^*(t)\,dt \tag{A.61}$$

are sufficient statistics for the optimum coherent binary detection problem considered in this section. This characteristic is the major contribution of Pitcher's theorem; the continuous time observation $y_z(t)$ (an uncountable number of random variables) can be reduced to two random variables without any loss of information about the transmitted information bit. This idea is what makes implementation of optimum demodulators for digital communications in the presence of additive Gaussian noise feasible.

A.3.1 Coherent Detection in AWGN

Results for the AWGN channel can be greatly simplified and an intuitive demodulation structure results. Recall AWGN is characterized by

$$R_{N_z}(\tau) = N_0\delta(\tau). \tag{A.62}$$

Again the AWGN model does not have a continuous correlation function but the model is a mathematical approximation of a noise that does have a continuous correlation function. A more mathematical rigorous derivation using the Weiner process is given in [Poo88]. The generalized matched filter for AWGN will satisfy the following integral equation

$$\int_a^b N_0\delta(t-u)q_j(u)du = N_0 q_j(t) = x_j(t-\tau_p). \tag{A.63}$$

The matched filter for AWGN is given as

$$q_j(t) = \frac{1}{N_0}x_j(t-\tau_p). \tag{A.64}$$

This property implies that the optimum receiver has the form

$$\exp\left(\frac{2L_p}{N_0}\Re\left\{e^{-j\phi_p}\int_a^b y_z(t)\left[x_1(t-\tau_p)-x_0(t-\tau_p)\right]^* dt\right\}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \exp\left(\frac{L_p^2}{N_0}\left(\int_a^b |x_1(t-\tau_p)|^2 dt\right)\right) \times \exp\left(-\frac{L_p^2}{N_0}\left(\int_a^b |x_0(t-\tau_p)|^2 dt\right)\right)\frac{\pi_0}{\pi_1}. \tag{A.65}$$

Noting that

$$\int_a^b |x_j(t-\tau_p)|^2 dt = E_j \tag{A.66}$$

Rearranging gives the matched filter output/energy correction formulation that is used throughout this text, i.e.,

$$\exp\left(\frac{2L_p}{N_0}\Re\left\{e^{-j\phi_p}Q_1\right\} - \frac{L_p^2E_1}{N_0}\right)\pi_1 \underset{H_0}{\overset{H_1}{\gtrless}} \exp\left(\frac{2L_p}{N_0}\Re\left\{e^{-j\phi_p}Q_0\right\} - \frac{L_p^2E_0}{N_0}\right)\pi_0. \tag{A.67}$$

A.4 Noncoherent Detection

Noncoherent detection typically refers to detection where no phase estimation is attempted. While the performance of noncoherent detection in AWGN is lower than coherent detection noncoherent detection is often used in practical communications systems. The same signal model is retained, i.e.,

$$Y_z(t) = L_p x_j(t) e^{j\phi_p} + N_z(t) \tag{A.68}$$

except that ϕ_p is assumed unknown and uniformly distributed in $[-\pi,\pi]$. Consequently, using the models in Section A.1 implies that $\Omega_j = \left\{u_j(t)e^{j\phi_p}, \phi_p \in [-\pi,\pi]\right\}$, i.e., ϕ_p is the nuisance parameter in the detection problem. The likelihood ratio is now given using total probability as

$$L\left(\vec{Y}\right) = \frac{f_{\vec{Y}}\left(\vec{y}\,|x(t,m)\in\Omega_1\right)}{f_{\vec{Y}}\left(\vec{y}\,|x(t,m)\in\Omega_0\right)} = \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi} f_Y\left(\vec{y}\,\middle|x_1(t-\tau_p)e^{j\phi_p}\right)d\phi_p}{\frac{1}{2\pi}\int_{-\pi}^{\pi} f_Y\left(\vec{y}\,\middle|x_0(t-\tau_p)e^{j\phi_p}\right)d\phi_p} \tag{A.69}$$

Grenander's theorem can again be used to evaluate (A.69) by truncating the KL expansion and examining the likelihood function for this truncated sequence, i.e.,

$$L_N\left(\vec{Y}\right) = \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\left(-\sum_{i=1}^{N} \frac{\left|y_i - L_p e^{j\phi_p} x_{i1}\right|^2}{\lambda_i}\right) d\phi_p}{\frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\left(-\sum_{i=1}^{N} \frac{\left|y_i - L_p e^{j\phi_p} x_{i0}\right|^2}{\lambda_i}\right) d\phi_p}. \tag{A.70}$$

Mulitplying out the terms in the argument of the exponential and grouping the terms that depend on both $y_z(t)$ and ϕ_p, the optimum demodulator for the truncated vector is then given as

$$L_N\left(\vec{Y}\right) = \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\left(2\Re\left\{L_p e^{j\phi_p} \sum_{i=1}^{N} \frac{y_i x_{i1}^*}{\lambda_i}\right\}\right) d\phi_p}{\frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\left(2\Re\left\{L_p e^{j\phi_p} \sum_{i=1}^{N} \frac{y_i x_{i0}^*}{\lambda_i}\right\}\right) d\phi_p} \underset{H_0}{\overset{H_1}{\gtrless}} R\left(\vec{x}_1, \vec{x}_0, \vec{\Lambda}, L_p, \pi_0\right). \tag{A.71}$$

Using Pitcher's property and Grenander's theorem again, the optimum demodulator converges to

$$\frac{\frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\left(2\Re\left\{L_p e^{j\phi_p} \int_a^b y_z(t)\, q_1^*(t)\, dt\right\}\right) d\phi_p}{\frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\left(2\Re\left\{L_p e^{j\phi_p} \int_a^b y_z(t)\, q_0^*(t)\, dt\right\}\right) d\phi_p} \underset{H_0}{\overset{H_1}{\gtrless}} R\left(x_1(t), x_0(t), R_{N_z}(\tau), L_p, \pi_0\right). \tag{A.72}$$

Using the definition of $I_0(x)$ [Ae72] one gets

$$\frac{I_0\left(2L_p\left|\int_a^b y_z(t)\, q_1^*(t)\, dt\right|\right)}{I_0\left(2L_p\left|\int_a^b y_z(t)\, q_0^*(t)\, dt\right|\right)} = \frac{I_0\left(2L_p|q_1|\right)}{I_0\left(2L_p|q_0|\right)} \underset{H_0}{\overset{H_1}{\gtrless}} R\left(x_1(t), x_0(t), R_{N_z}(\tau), L_p, \pi_0\right). \tag{A.73}$$

A.4.1 Noncoherent Detection In AWGN

Results for the AWGN channel can be greatly simplified and an intuitive demodulation structure results. Recall AWGN is characterized by

$$R_{N_z}(\tau) = N_0 \delta(\tau). \tag{A.74}$$

The generalized matched filter for AWGN will satisfy the following integral equation

$$\int_a^b N_0 \delta(t-u) q_j(u) du = N_0 q_j(t) = x_j(t - \tau_p). \tag{A.75}$$

The matched filter for AWGN is given as

$$q_j(t) = \frac{1}{N_0} x_j(t - \tau_p). \tag{A.76}$$

This property implies that the optimum receiver noncoherent detection in AWGN has the form

$$\frac{I_0\left(\frac{2L_p}{N_0}\left|\int_a^b y_z(t)\, x_1^*(t-\tau_p)\, dt\right|\right)}{I_0\left(\frac{2L_p}{N_0}\left|\int_a^b y_z(t)\, x_0^*(t-\tau_p)\, dt\right|\right)} = \frac{I_0\left(\frac{2L_p}{N_0}|q_1|\right)}{I_0\left(\frac{2L_p}{N_0}|q_0|\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \exp\left(\frac{L_p^2(E_1 - E_0)}{N_0}\right) \frac{\pi_0}{\pi_1}. \tag{A.77}$$

Appendix B

Minimum Mean Square Error Estimation

This appendix summarizes several important results for minimum mean square error (MMSE) estimation. Detailed results and detailed proofs for this material can be obtained from a wide variety of textbooks [KSH00, Poo88, Sch91, VT68]. The results needed for this text correspond to situations where the observations are vectors of length N or an infinite time series. The problem notation is given as

- The observations will be denoted $\vec{Y}$ or $Y(k)$ where k is the discrete time index.
- The desired output is B or $B(k)$.
- The estimator is denoted as $\hat{B}$ or $\hat{B}(k)$.
- The error is denoted as $E = B - \hat{B}$ or $E(k) = B(k) - \hat{B}(k)$.

Minimum mean square estimation is a Bayesian estimator that minimizes the cost function $E\left[|E|^2\right]$ or $E\left[|E(k)|^2\right]$

Property B.1 *The unconstrained MMSE estimator is the conditional mean estimator, i.e.,*

$$\hat{B} = E\left[B\left|\vec{Y}\right.\right] \qquad \hat{B}(k) = E\left[B(k)\left|\vec{Y}\right.\right] \tag{B.1}$$

Linear estimation is often used in practice because it is optimal (with Gaussian signals and noise) or the complexity is managable. Linear estimation usually occurs when the random observations and desired output are zero mean signals. If non-zero means exist then estimators have an affine characteristic. The linear estimator is given as

$$\hat{B} = \vec{w}^H \vec{Y} \qquad \hat{B}(k) = \sum_l w^*(l) Y(k-l) \tag{B.2}$$

Property B.2 *The linear MMSE estimator exists if and only if* $E\left[EZ^*\right] = 0$ *where* $Z = \vec{H}^H \vec{Y}$.

This property states that the error resulting from MMSE linear estimation is orthogonal to any linear function of the observations. A useful special case of this property is

Property B.3 *The MMSE estimator exists if and only if* $E\left[E\vec{Y}_i^*\right] = 0$ *for any* i.

Stacking this orthogonality condition for all of the observed data results in the Wiener-Hopf equations

$$E\left[E\vec{Y}^H\right] = \vec{0}_N^T = E\left[B\vec{Y}^H\right] - \vec{w}^H E\left[\vec{Y}\vec{Y}^H\right]. \tag{B.3}$$

Rearranging and solving gives

$$\vec{w}^H = E\left[B\vec{Y}^H\right]\left(E\left[\vec{Y}\vec{Y}^H\right]\right)^{-1} \tag{B.4}$$

The minimum error is also given as

$$E\left[|E|^2\right] = E\left[|B|^2\right] - E\left[B\vec{Y}^H\right]\left(E\left[\vec{Y}\vec{Y}^H\right]\right)^{-1} E\left[\vec{Y}B^*\right] \tag{B.5}$$

If we consider the case of the observations and the desired output being jointly stationary time series the orthogonality condition becomes

$$0 = E\left[\left(B(k) - \sum_l w(l)Y(k-l)\right)Y(i)^*\right] \qquad \forall i \tag{B.6}$$

and Wiener-Hopf equations then become

$$E\left[B(k)Y(i)^*\right] = \sum_l w(l)E\left[Y(k-l)Y(i)^*\right] \qquad \forall i \tag{B.7}$$

$$R_{BY}(m) = \sum_l w(l)R_Y(m-l) \qquad \forall m. \tag{B.8}$$

Note that this corresponds to an infinte number of equations and an infinite number of unknowns. A more intuitive solution can be obtain by taking the Fourier transform of both sides. This results in

$$S_{BY}\left(e^{j2\pi f}\right) = W\left(e^{j2\pi f}\right)S_Y\left(e^{j2\pi f}\right) \tag{B.9}$$

where

$$W\left(e^{j2\pi f}\right) = \sum_k w(k)e^{j2\pi fk} \qquad S_Y\left(e^{j2\pi f}\right) = \sum_k R_Y(k)(k)e^{j2\pi fk}. \tag{B.10}$$

Consequently we have

$$W\left(e^{j2\pi f}\right) = \frac{S_{BY}\left(e^{j2\pi f}\right)}{S_Y\left(e^{j2\pi f}\right)} \qquad w(k) = \int_{-0.5}^{0.5} W\left(e^{j2\pi f}\right)e^{j2\pi fk}df \tag{B.11}$$

This type of a filter is often referred to as noncausal Wiener-Kolmogorov filter.

Here we consider the concept of causal (anti-causal) Wiener-Kolmogorov filters. This paragraph will summarize the results for causal filters as causal filters are more commonly implemented in practice, but the results are easily redone for anti-causal filters (which is actually the main result we will need in this text). Two characteristics enable causal filtering

1. Any correlation function can be decomposed into a causal and an anti-causal factorization (spectral factorization).

2. If inputs into a MMSE filter are white then the optimal unconstrained filter can be truncated (either causally or anti-causally) and the orthogonality conditions are satisfied.

The procedure for finding causal filters is summarized as 1) form a spectral factorization of the input 2) use the spectral factorization to causally whiten the input, and 3) form the unconstrained MMSE filter for the whitened input and truncate. The important first step is the spectral factorization theorem.

Theorem B.1 (Spectral Factorization) *If the power spectrum of a discrete random process, $S_X\left(e^{j2\pi f}\right)$, does not go to zero then*

$$S_X\left(e^{j2\pi f}\right) = \gamma_x F_x^-\left(e^{j2\pi f}\right) F_x^+\left(e^{j2\pi f}\right) \tag{B.12}$$

where $F_x^-\left(e^{j2\pi f}\right) = 1 + \sum_{l=1}^{\infty} f^*(l)\left(e^{j2\pi f}\right)^l$ and $F_x^+\left(e^{j2\pi f}\right) = 1 + \sum_{l=1}^{\infty} f(l)\left(e^{j2\pi f}\right)^{-l}$ and γ_x a positive constant.

Proof: See [Poo88, KSH00] for discussions and proof. □

Property B.4 *The correlation function is obtained as*

$$R_x(m) = \gamma_x \sum_{l=0}^{\infty} f(l) f^*(l-m) \tag{B.13}$$

Definition B.1 *A monic filter is a filter $F\left(e^{j2\pi f}\right)$ whose impulse response value at zero is unity, i.e., $f(0) = 1$.*

The spectral factorization is unique in the sense that $F_x^+\left(e^{j2\pi f}\right)$ and $F_x^-\left(e^{j2\pi f}\right)$ are both monic.

Property B.5 *An important property of the spectral factorization is*

$$\left(F_x^+\left(e^{j2\pi f}\right)\right)^* = F_x^-\left(e^{j2\pi f}\right) \tag{B.14}$$

Property B.6 *A second important property of the spectral factorization is that*

$$F_i\left(e^{j2\pi f}\right) = \left(F_x^+\left(e^{j2\pi f}\right)\right)^{-1} \tag{B.15}$$

is also a causal monic filter.

Property B.7 *If one considers an unconstrained MMSE filter, $\hat{B}(k)$, and if the input $Y(k)$ is a white process then truncating the unconstrained filter to form a causal filter $\hat{B}^+(k)$ will still satisfy the Wiener-Hopf equations for all causally considered observations.*

Proof: The orthogonality condition for the truncated filter for the causally considered observations can be restated as

$$\begin{aligned} E\left[\left(\hat{B}(k) - \hat{B}^+(k)\right) Y(k-m)^*\right] &= \sum_{l=1}^{\infty} w(-l) E\left[Y(k+l)Y(k-m)^*\right] \qquad \forall m \geq 0 \\ &= \sum_{l=1}^{\infty} w(-l) R_Y(m+l) = 0 \qquad \forall m \geq 0 \end{aligned} \tag{B.16}$$

□

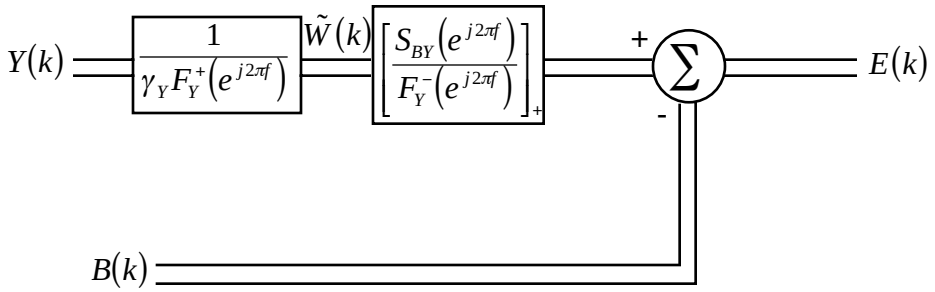


Figure B.1: The block diagram for the causal MMSE linear filter.

The spectral factorization theorem can be used to causally whiten the observation process by putting $Y(k)$ through the filter

$$W_1\left(e^{j2\pi f}\right) = \frac{1}{\sqrt{\gamma_y} F_y^+\left(e^{j2\pi f}\right)} \tag{B.17}$$

to obtain a process $\tilde{W}(k)$ where

$$S_{\tilde{W}}\left(e^{j2\pi f}\right) = \left|W_1\left(e^{j2\pi f}\right)\right|^2 S_Y\left(e^{j2\pi f}\right) = \frac{S_Y\left(e^{j2\pi f}\right)}{\gamma_y \left|F_x^+\left(e^{j2\pi f}\right)\right|^2} = 1. \tag{B.18}$$

The remainder of the unconstrained MMSE filter,

$$W_2\left(e^{j2\pi f}\right) = \frac{S_{BY}\left(e^{j2\pi f}\right)}{\sqrt{\gamma_y} F_y^-\left(e^{j2\pi f}\right)}, \tag{B.19}$$

can be truncated to be causal while still maintaining the desired orthogonality. The overall structure of the causal MMSE filter is shown in Fig. B.1.

Appendix C

Notation

Signals and Systems

$x(t)$=A signal
$X(f)$=The Fourier transform of $x(t)$
P_x=The power of $x(t)$
E_x=The energy of $x(t)$
$V_x(\tau)$=The correlation function of $x(t)$
$G_x(f)$=The energy spectrum of $x(t)$
$h(t)$=A filter impulse response
$H(f)$=A filter transfer function

Bandpass Signals

$x_c(t)$=Bandpass signal
$x_I(t)$=In-phase signal
$x_Q(t)$=Quadrature signal
$x_z(t)$=Complex envelope signal
$x_A(t)$=Amplitude signal
$x_P(t)$=Phase signal
B_T=Transmission bandwidth
f_c=Carrier frequency

Analog Modulations

$m(t)$=Analog message signal
W=Message signal bandwidth
τ_p=Propagation time delay
L_p=Propagation loss
ϕ_p=Propagation phase shift
E_T=Transmission efficiency
E_B=Bandwidth efficiency
A_c=Transmission amplitude
a=Modulation coefficient
$f_i(t)$=Instantaneous frequency
$f_d(t)$=Instantaneous frequency deviation
k_p=Phase deviation constant
k_f=Frequency deviation constant (radians/s/volt)
f_k=Frequency deviation constant (Hertz/volt)
f_p=Peak frequency deviation
D=Bandwidth expansion factor
E_M=Multiplexing Efficiency

Digital Communications Basics

E_b=Energy per bit
T=Symbol time
W_b=Bit rate
T_p=Transmission duration
ν_B=Spectral efficiency (bits/s/Hz)
$\vec{I},\ I$=Information bits
$\hat{\vec{I}},\ \hat{I}$=Information bit estimates
$\Delta_E(\alpha,\beta)$=Square Euclidean distance between $x_\alpha(t)$ and $x_\beta(t)$
$\{A_d(k),\Delta_E(k)\}$=Squared Euclidean distance spectrum of a modulation
K_b=Number of bits to be transmitted per frame
π_i=A prior probability of $\vec{I}=i$

Linear Modulations

$u(t)$=Modulation pulse shape
$D_i(k)$=Modulation symbol
$a(\bullet)$=Constellation mapping, i.e., $d_i = a(i),\ i = 0, M-1$
Q=Output of filter matched to $u(t)$

Orthogonal Modulations

$I(l)$=Information bit to be transmitted at step l
$D_z(l)$=Modulation symbol for orthogonal modulation
T_u=Pulse shape duration for stream modulation

Orthogonal Modulations with Memory

$\sigma(l)$=Modulation state at time l
$N_s(l)$=Number of states possible at time l
$J(l)$=Constellation label sequence, i.e., $D_z(l) = a(J(l))$
M_s=Number of points in the constellation, i.e., $J(l) = i,\ i = 0, M_s - 1$
L_c=Number of taps in the discrete time equivalent channel model for channels with ISI
$g_1(\bullet,\bullet)$=The modulation finite state machine, i.e., $\sigma(l+1) = g_1\left(I(l), \sigma(l)\right)$
$g_2(\bullet,\bullet)$=The output symbol selector mapping, i.e., $J(l) = g_2\left(I(l), \sigma(l)\right)$
L_e=Length of a simple error event

General Linear Modulations with Memory

$\vec{I}(l),\ I(l)$=Information bits to be transmitted at symbol time l
N_c=Number of symbols transmitted
K_s=Number of bits per trellis transition
M_i=Number of branches per trellis transition, i.e., $M_i = 2^{K_s}$
N_s=Number of output symbols per trellis transition
R=Number of bits transmitted per symbol, i.e., $R = K_s/N_s$

Nonlinear Modulations with Memory

L_m=Phase response length of a CPM modulation
h=CPM modulation index
$q_m(t)$=Phase smoothing response
$g_f(t)$=Frequency pulse
$\theta(k)$=Phase state at time k

Appendix D

Acronyms

Signals and Systems

ADC=Analog to digital converter
BPF=Bandpass filter
DC=Direct current, often refers to zero frequency
DTFT=Discrete time Fourier transform
I=In-phase
LPF=Lowpass filter
LTI=Linear Time Invariant
Q=Quadrature
SNR=Signal to noise ratio

Random Variables and Processes

AWGN=Additive white Gaussian noise
CDF=Cumulative distribution function
ChF=Characteristic function
CLT=Central limit theorem
PDF=Probability density function
PMF=Probability mass function
PSD=Power spectral density
RV=Random variable
WSS=Wide Sense Stationary

Analog Communication

AM=Amplitude modulation
DSB-AM=Double sideband AM
FDM=Frequency division multiplexing
FM=Frequency modulation
LC-AM=Large carrier AM
PLL=Phase-locked loop
PM=Phase modulation
QCM=Quadrature carrier multiplexing
SSB-AM=Single sideband AM
TV=Television
VCO=Voltage controlled oscillator
VSB-AM=Vestigial sideband AM

Digital Communication Basics

APP=A posteriori probability
BEP=Bit error probability
BFSK=Binary frequency shift keying
BPSK=Binary phase shift keying
DTMF=Discrete tone multiple frequency
EVM=Error vector magnitude
IEEE=Institute of electrical and electronic engineers
ISI=Intersymbol interference
MAP=Maximum a posteriori
MAPBD=Maximum a posteriori bit demodulation
MLBD=Maximum likelihood bit demodulation
MLWD=Maximum likelihood word demodulation
MFSK=M-ary frequency shift keying
MPSK=M-ary phase shift keying
MSK=Minimum shift keying
OCDM=Orthogonal code division multiplexing
OFDM=Orthogonal frequency division multiplexing
PPM=Pulse position modulation
PWM=Pulse width modulation
PWEP=Pair-wise error probability
QPSK=Quadri-phase shift keying
VPSK=Variable phase shift keying

Advanced Digital Communications

ICI=Intercarrier interference
=Not completed this edition