
Principles Of Digital Communications

Bixio Rimoldi
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Switzerland

14.3.2007

Chapter 1

Introduction and Objectives

The evolution of communication technology during the past few decades has been impressive. In spite of an enormous progress, many of the challenges still lay ahead of us. While any prediction of the next big technological revolution is likely to be wrong, it is safe to say that communication devices will become smaller, lighter, more powerful, more integrated, more ubiquitous, and more reliable than they are today. Perhaps one day the input/output interface and the communication/computation hardware will be separated. The former will be the only part that we will carry on us and it will communicate wirelessly with the latter. Perhaps the communication/computation hardware will be part of the infrastructure. It will be built into cars, trains, airplanes, public places, homes, offices, etc. With the the input/output device that we carry around we will have virtually unlimited access to communication and computation facilities. Search engines may be much more powerful than they are today, giving instant access to any information digitally stored. The input/output device may contain all of our preferences so that, for instance, when we sit down in front of a computer, we see the environment that we like regardless of location (home, office, someone else's desk) and regardless of the hardware and operating system. The input device may also allow us to unlock doors and make payments, making keys, credit cards, and wallets obsolete. Getting there will require joint efforts from almost all branches of electrical engineering, computer science, and system engineering.

In this course we focus on the system aspects of digital communications. Digital communications is a rather unique field in engineering in which theoretical ideas have had an extraordinary impact on actual system design. Our goal is to get acquainted with some of these ideas. Hopefully, you will appreciate the way that many of the mathematical tools you have learned so far will turn out to be exactly what we need. These tools include probability theory, stochastic processes, linear algebra, and Fourier analysis.

We will focus on systems that consist of a single transmitter, a channel, and a receiver as shown in Figure 1.1. The channel filters the incoming signal and corrupts it with

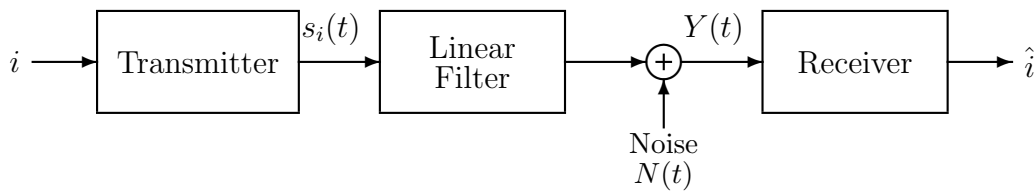


Figure 1.1: Basic point-to-point communication system over a bandlimited Gaussian channel.

noise. The noise is Gaussian since it represents the contribution of various noise sources.¹ The filter in the channel model has both a physical and a conceptual justification. The conceptual justification stems from the fact that most wireless communication systems are subject to a license that dictates, among other things, the frequency band that the signal is allowed to occupy. A convenient way to enforce this constraint is to tell the system designers that the channel contains an ideal filter that blocks everything outside the intended band. The physical reason has to do with the observation that the signal emitted from the transmit antenna typically encounters obstacles that create reflections and scattering. Hence the receive antenna may capture the superposition of a number of delayed and attenuated replicas of the transmitted signal (plus noise). It is a straightforward exercise to check that this physical channel is linear and time-invariant. Thus it may be modeled by a linear filter as shown in the figure.² In some cases the transmit and/or the receive antennas also filter the signal. This happens, for instance, when the signal's bandwidth is so large that the antenna characteristic varies over the frequency interval spanned by the signal. The filter in Figure 1.1 accounts for these and possibly other linear time-invariant transformations that act upon the communication signals as it travels from the sender to the receiver. The channel model of Figure 1.1 is meaningful for both wireline and wireless communication channels. It is referred to as the bandlimited Gaussian channels.

Since communication means different things for different people, we need to clarify the role of the transmitter/receiver pair depicted in Figure 1.1. For the purpose of this class a transmitter implements a mapping between a message set and a signal set, both of the same cardinality, say m . The number m of elements of the message set is important but the nature of its elements is not. (More on this later.) Without loss of generality we can let the message set consist of the integers $\{0, 1, \dots, m - 1\}$. The elements of the message set are called messages. There is a one-to-one correspondence between messages

¹Individual noise sources do not necessarily have Gaussian statistics. However, due to the central limit theorem, their aggregate contribution is often quite well approximated by a Gaussian random process.

²If the scattering and reflecting objects move with respect to the transmit/receive antennae then the filter is time-varying but this case is deferred to the advanced digital communication class.

and elements of the signal set. The “nature” (e.g. discrete vs continuous time) of the signals is important since signals have to be compatible with the channel. The channel is always assumed to be given to the designer who has no control over it. By assumption, the designer can only control the design of the transmitter/receiver pair. A user communicates by selecting a message $i \in \{0, 1, \dots, m - 1\}$ which is converted by the transmitter into the corresponding signal s_i . The channel reacts to the signal by producing the observable y . Based on y , the receiver generates an estimate $\hat{i}(y)$ of i . Hence the receiver is a map from the space of channel output signals to the message set. Hopefully $i = \hat{i}$ most of the time. When this is not the case we say that an error event occurred. In all situations of interest to us it is not possible to reduce the probability of error to zero. This is so since, with positive probability, the channels is capable of producing an output y that could have stemmed from more than one message. One of the performance measures of a transmitter/receiver pair for a given channel is thus the probability of error. Another performance measure is the rate at which we communicate. Conceptually, we may label every message with a unique sequence of $\log m$ bits so that communicating the message is equivalent to communicating the corresponding bit sequence. (This is why earlier we said that the nature of the messages is not relevant). Hence we are sending the equivalent of $\log m$ bits every time we use the channel. By increasing the value of m we increase the rate in bits per channel use but, as we will see, under normal circumstances this increase can not be done indefinitely without increasing the probability of error.

At the end of this course you should have a good understanding of a basic communication system and be able to make sensible design choices. In particular, you should know what a receiver does to minimize the probability of error, be able to do a quantitative analysis of some of the most important performance figures, and know which tradeoffs you have as a system designer.

A few words about the big picture and the approach that we will take are in order. We will discover that a natural way to design, analyze, and implement a transmitter/receiver pair is in terms of the modules shown in Figure 1.2. These modules allow us to focus on selected issues while hiding others. For instance, at the very bottom level we exchange messages. At this level we may think of all modules as being inside a “black box” that hides all the implementation details and lets us see only what the user has to see from the outside. The “black box” is an abstract channel model that takes messages and delivers messages not always without making errors. At this level of granularity the visible performance figures are the cardinality of the message set, how long we have to wait until we are allowed to choose the next message, and the probability of error. The first two determine how many bits we send per unit of time, i.e., the rate at which we communicate. At the top level of Figure 1.2 we focus on the characteristics of the actual signals being sent over the physical medium, such as the average power of the transmitted signal and the frequency band it occupies. We will see that at the second level from the bottom we communicate n -tuples. It is at this level that we will understand the heart of the receiver. We will understand how the receiver should base its decision so as to minimize the probability of error and see how to compute the resulting error probability. Finally, one layer up we communicate using low-frequency (as opposed to radio frequency) signals. Separating the

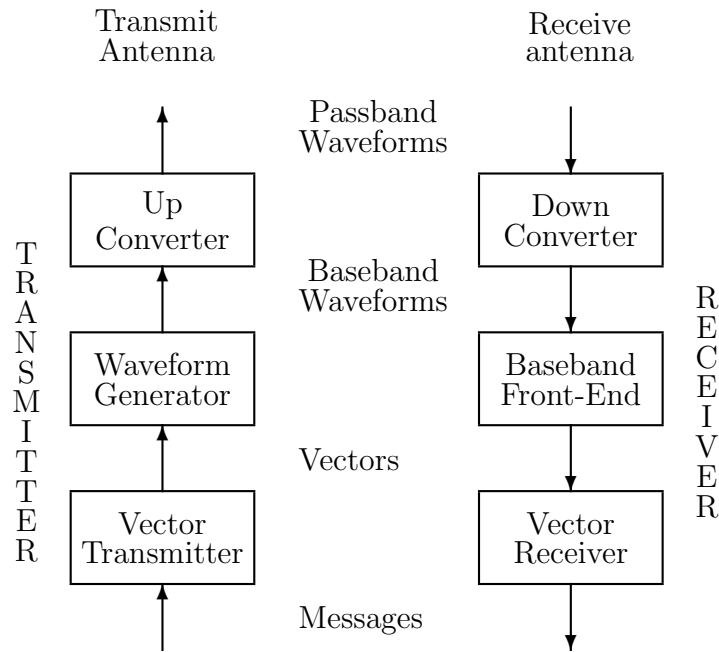


Figure 1.2: Decomposed transmitter and receiver.

top two layers is important for implementation purposes.

There is more than one way to organize the discussion around the modules of Figure 1.2. Following the signal path, i.e., starting from the first module of the transmitter and working our way through the system until we reach the final stage of the receiver would not be a good idea since it makes little sense to study the transmitter design without having an appreciation of the task and limitations of a receiver. We will instead make many passes over the block diagram of Figure 1.2, each time at a different level and focussing on different issues as discussed in the previous paragraph, but each time considering the sender and the receiver together. We will start with the channel seen by the bottom modules in Figure 1.2. This approach has the advantage that you will quickly be able to appreciate what the transmitter and the receiver should do. One may argue that this approach has the disadvantage of asking the student to accept an abstract channel model that seems to be oversimplified. (It is not, but this will not be immediately clear). On the other hand one can also argue in favor of the pedagogical value of starting with highly simplified models. Shannon, the founding father of modern digital communication theory and one of the most profound engineers and mathematicians of the 20th century, was known to solve difficult problems by first reducing the problem to a much simpler version that he could almost solve “by inspection.” Only after having familiarized himself with the simpler problem would he work his way back to the next level of difficulty.

The choice of material covered in this course is by now more or less standard for an introductory course on digital communications. The approach depicted in Figure 1.2 has been made popular by J.M. Wozencraft and I. M. Jacobs in *Principles of Communication*

Engineering – a textbook appeared in 1965. However, the field has evolved since then and these notes reflect such evolution. Some of the exposition has benefited from the notes *Introduction to Digital Communication*, written by Profs. A. Lapidoth and R. Gallager for the MIT course Nr. 6.401/6.450, 1999. I am indebted to them for letting me use their notes during the first few editions of this course.

There is only so much that one can do in one semester. EPFL offers various possibilities for those who want to know more about digital communications and related topics. Classes for which this course is a recommended prerequisite are *Advanced Digital Communications*, *Information Theory and Coding*, *Principles of Diversity in Wireless Networks*, and *Coding Theory*. For the student interested in hands-on experience, EPFL offers *Software-Defined Radio: A Hands On Course*.

Networking is another branch of communications that has developed almost independently of the material treated in this class. It relies on quite a different set of mathematical models and tools. Networking assumes that there is a network of bit pipes which is reliable most of the time but that can fail once in a while, e.g., due to network congestion, hardware failure, queue overflow, etc. Queues are used to temporarily store packets when the next link is congested. Networking deals with problems such as finding a route for a packet, computing the delay incurred by a packet as it goes from source to destination considering the queueing delay and the fact that packets are retransmitted if their reception is not acknowledged. We will not be dealing with networking problems in this class.

We conclude this introduction with a very brief overview of the various chapters. Not everything in this overview will make sense to you now. Nevertheless we advise you to read it now and read it again when you feel that it is time to step back and take a look at the “big picture.” It will also give you an idea of which fundamental concepts will play a role in this course.

Chapter 2 deals with the vector channel case of Figure 1.2. The emphasis will be on the design of an optimal Vector Receiver, assuming that the Vector Transmitter and the Vector Channel are given. This is an application of what is known in the statistical literature as hypothesis testing (to be developed in Chapter 2). After a rather general start we will spend some time on the Gaussian Vector Channel. (In Chapter 9 you will realize that the Gaussian Vector Channel is a cornerstone of digital communications.)

In Chapter 3 we will focus on the Waveform Generator and on the Baseband Front-End of Figure 1.2. The mathematical tool behind the description of the Waveform Generator is the notion of orthonormal expansion from linear algebra. We will fix an orthonormal basis and we will let the output of the Vector Transmitter be the vector of coefficients that determine the signal produced by the Waveform Transmitter (with respect to the given orthonormal basis). The Baseband Front-End of the receiver reduces the received waveform to a vector (n -tuple) that contains just as much information as needed to decide about the message selected by the sender. To do so the Baseband Front-End projects the received waveform onto each element of the mentioned orthonormal basis. The resulting n -tuple is passed to the Vector Receiver. Together, the Vector Transmitter

and the Waveform Generator form the Waveform Transmitter. Together the Baseband Front-End and the Vector Receiver form the Waveform Receiver. What we do in Chapter 3 holds irrespectively of the specific set of signals that we use to communicate.

Chapter 4 deals with general high level implications of a specific signal set.

Chapter ?? is about choosing a convenient orthonormal basis for the Waveform Generators. We will see that it is possible to choose in such a way that the transmitted signals have a desirable power spectral density and, at the same time, satisfy the desire for a relatively low-cost Baseband Front-End. The main concept here is what is called Nyquist criterion.

Chapter ?? deals with the Up/Down Converters. The idea is to learn how to shift the spectrum of the transmitted signal so that we can place its center frequency at any desired location in the frequency axis, without changing what we have called the Waveform Transmitter and the Waveform Receiver. This will be done using one of the fundamental properties of the Fourier transform. Given our ability to shift the center frequency of the transmitted signal to any desired location, it makes sense to let the Waveform Transmitter and the Waveform Receiver operate in some fixed frequency range if this simplifies their implementation. Implementing signal processing (amplification, filtering, multiplication of signals, etc.) becomes more and more challenging as the center frequency of the signals being processed increases. This is so since simple wires meant to carry the signal inside the circuit may act as transmit antenna and irradiate the signal. This may cause all kinds of problems, including the fact that signals get mixed “in the air” and, even worse, are reabsorbed into the circuit by some short wire that acts as receive antenna causing interference, oscillations due to unwanted feedback, etc.. To minimize such problems, it is common practice to let the Waveform Transmitter and Waveform Receiver operate at “baseband”, i.e. process signals that have $f = 0$ as their center frequency. As it turns out, the baseband representation of a general signal is complex-valued, even if the signal being represented is real-valued. This means that the Waveform Transmitter/Receiver pairs have to deal with complex-valued signals. This is not a problem per se. In fact working with complex-valued signals simplifies the notation. However, it requires a small overhead in terms of having to learn how to deal with complex-valued stochastic processes and complex-valued random vectors.

Dealing with complex-valued Gaussian processes and vectors is the topic of Chapter 8.

Chapter 9 “closes the loop” by showing that the channel “seen” by the Vector Transmitter and the Vector Receiver is indeed the abstract Gaussian Vector Channel that we have assumed in Chapter 2. To emphasize the importance of the Vector Channel we mention that in a typical information theory course (mandatory at the master-level at EPFL) as well as in a typical coding theory course (offered at EPFL in the Ph.D. program), the channel is a Vector Channel (perhaps not called this way) and one takes it for granted that the student knows where it comes from. (The material treated in this class is also assumed as being assimilated in *Advanced Digital Communications* as well as in *Software-Defined Radio: A Hands on Course*, both of which are offered at EPFL at the master

level.)

Chapter 6 is a case study on coding. The communication model is that of Chapter 2 with the Vector Channel being Gaussian. The Vector Transmitter will incorporate a convolutional encoder and the Vector Receiver will be based on the Viterbi algorithm. The performance of the resulting scheme will be analyzed and compared to the uncoded case.

Chapter 2

Optimal Receivers for Vector Channels

2.1 Introduction

As pointed out in the introduction, we will study point-to-point communications from various abstraction levels. In this chapter we will be dealing with the vector channel. In the next chapter it will become clear why the vector channel is an important abstraction model. For now it suffices to say that it is the channel that we see from the input to the output of the dotted box in Figure 2.1. The goal of this chapter is to understand how to design and analyze the vector receiver when the channel and the transmitter are given.

We start with the communication system depicted in Figure 2.2 (in which the channel is a bit more general than the vector channel depicted in Figure 2.1). Its components are:

- The source: It is responsible for producing the message $H \in \mathcal{H} = \{0, 1, \dots, (m-1)\}$. The task of the receiver would be extremely simple if the source selected the message according to some deterministic rule. In this case the receiver could reproduce the source message by following the same algorithm and there would be no need for a communication system. For this reason, in communication we always assume that the source is modeled by a random variable, here denoted by the capital letter H . As usual, a random variable taking values on a finite alphabet is described by its probability mass function $P_H(i)$, $i \in \mathcal{H}$. In most cases of interest to us, H is uniformly distributed and/or $m = 2$.
- The transmitter: It is a mapping from \mathcal{H} to $\mathcal{S} = \{\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_{m-1}\}$ where $\mathbf{s}_i \in \mathbb{C}^n$ for some n . (We will start with $\mathbf{s}_i \in \mathbb{R}^n$ but we will see in the last chapter that it is crucial that we allow for $\mathbf{s}_i \in \mathbb{C}^n$).
- The channel: It is described by the probability density of the output for each of the possible inputs. When the channel input is \mathbf{s}_i , the probability density of \mathbf{Y} will be denoted by $f_{\mathbf{Y}|\mathcal{S}}(\mathbf{y}|\mathbf{s}_i)$.

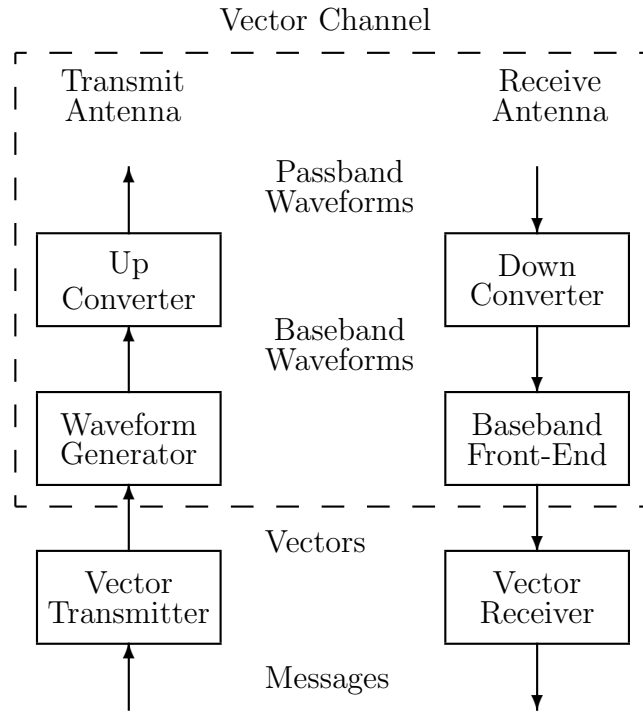


Figure 2.1: Vector channel abstraction.



Figure 2.2: Main setup considered Part I.

- The receiver: The receiver's task is to “guess” H from \mathbf{Y} . The decision made by the receiver is denoted by \hat{H} . Unless specified otherwise, the receiver will always be designed to minimize the probability of error defined as the probability that \hat{H} differs from H . This is the so-called *hypothesis testing* problem that comes up in various contexts (not only in communications).

First we give a few examples.

EXAMPLE 1. A common source model consist of $\mathcal{H} = \{0, 1\}$ and $P_H(0) = P_H(1) = 1/2$. This models individual bits of, say, a file. Alternatively, one could model an entire file of, say, 1 Mbit by saying that $\mathcal{H} = \{0, 1, \dots, (2^{10^6} - 1)\}$ and $P_H(i) = \frac{1}{2^{10^6}}, i \in \mathcal{H}$.

EXAMPLE 2. A transmitter for a binary source could be a map from $\mathcal{H} = \{0, 1\}$ to

$\mathcal{S} = \{-a, a\}$ for some real-valued constant a . Alternatively, a transmitter for a 4-ary source could be a map from $\mathcal{H} = \{0, 1, 2, 3\}$ to $\mathcal{S} = \{a, ia, -a, -ia\}$, where $i = \sqrt{-1}$.

EXAMPLE 3. The channel model that we will use mostly in this chapter is the additive white Gaussian (AWGN) channel that maps a channel input $\mathbf{s} \in \mathbb{R}^n$ into $\mathbf{Y} = \mathbf{s} + \mathbf{Z}$, where \mathbf{Z} is a Gaussian random vector with independent components.

Specifying the decision rule implemented by the receiver is straightforward once we understand the hypothesis testing problem studied in the next section.

2.2 Hypothesis Testing

Detection, decision, and hypothesis testing are all synonyms. They refer to the problem of deciding the outcome of a random variable H that takes values on a finite alphabet $\mathcal{H} = \{0, 1, \dots, m-1\}$, from the outcome of some related random variable Y . The random variable H is called the *Hypothesis* and Y the *observation*.

The problem that a receiver has to solve is a detection problem in the above sense. Here the hypothesis H is the message selected by the source. The transmitter sends a signal (typically a distinct signal for each letter of \mathcal{H}) and the receiver observes the channel response Y . The receiver decides the value of H based on Y . The receiver's decision will be denoted by \hat{H} . We wish to make $\hat{H} = H$, but this is not always possible. The goal is to devise a decision that makes $P_c = Pr\{\hat{H} = H\}$ as large as possible.¹

The standard assumption is that we know the *a priori* probability P_H and for each $i \in \mathcal{H}$ we know the conditional probability density function² (pdf) $f_{Y|H}(y|i)$ of Y .

EXAMPLE 4. Here is a good example of a typical hypothesis testing problem. The problem is that of communicating one bit of information (or more generally a sequence of bits) across an optical fiber. The bit being transmitted is modeled by the random variable $H \in \{0, 1\}$, $P_H(0) = 1/2$. If $H = 1$, we switch on a LED whose light is carried across an optical fiber to a photodetector at the receiver front end. The photodetector outputs the number of photons $Y \in \mathbb{N}$ it detects. The problem is to decide whether $H = 0$ or $H = 1$. Our decision may only be based on whatever prior information we have about the model and on the actual observation y . What makes the problem interesting is that it is impossible to determine H from Y with certainty. Even if the LED is off, the detector is likely to detect some photons (e.g. due to "ambient light"). A good assumption is that Y is Poisson distributed with intensity λ that depends on whether the LED is on or off.

¹Pr is a short-hand for *probability of the enclosed event*.

²In most cases of interest in communication, the random variable Y is continuous. That's why in the above discussion we have implicitly assumed that, given $H = i$, Y has a pdf $f_{Y|H}(y|i)$. If Y is a discrete random variable, then we assume that we know the conditional probability $p_{Y|H}(y|i)$.

Mathematically, the situation is as follows:

$$\begin{aligned} H = 0, \quad Y &\sim p_{Y|H}(y|0) = \frac{\lambda_0^y}{y!} e^{-\lambda_0} \\ H = 1, \quad Y &\sim p_{Y|H}(y|1) = \frac{\lambda_1^y}{y!} e^{-\lambda_1} \end{aligned}$$

We read the first row as follows: “When the hypothesis is $H = 0$ then the observable Y is Poisson distributed with intensity λ_0 ”.

The problem of deciding the value of H from the observable Y when we know the distribution of H and that of Y for each value of H is a standard hypothesis testing problem. \square

From P_H and $f_{Y|H}$, via Bayes rule, we obtain

$$P_{H|Y}(i|y) = \frac{P_H(i)f_{Y|H}(y|i)}{f_Y(y)}$$

where $f_Y(y) = \sum_i P_H(i)f_{Y|H}(y|i)$. $P_{H|Y}(i|y)$ is the *posterior* (also called a *posteriori probability* of H given Y). Once we have observed that $Y = y$, the probability that $H = i$ becomes $P_{H|Y}(i|y)$.

If we choose $\hat{H} = i$, then $P_{H|Y}(i|y)$ is the probability that we made the correct decision. Since our goal is to maximize the probability of being correct, the optimum decision rule is

$$\hat{H}(y) = \arg \max_i P_{H|Y}(i|y) \quad (\text{MAP decision rule}). \quad (2.1)$$

This is called *maximum a posteriori (MAP) decision rule*. In case of ties, i.e. if $P_{H|Y}(j|y)$ equals $P_{H|Y}(k|y)$ equals $\max_i P_{H|Y}(i|y)$, then it does not matter if we decide for $\hat{H} = k$ or for $\hat{H} = j$. In either case the probability that we have decided correctly is the same.

Since the MAP rule maximizes the probability of being correct for each observation y , it also maximizes the unconditional probability of being correct P_c . The former is $P_{H|Y}(\hat{H}(y)|y)$. If we plug in the random variable Y instead of y , then we obtain a random variable. (A real-valued function of a random variable is a random variable.) The expected value of this random variable is the (unconditional) probability of being correct, i.e.,

$$P_c = E[P_{H|Y}(\hat{H}(Y)|Y)] = \int_y P_{H|Y}(\hat{H}(y)|y) f_Y(y) dy.$$

There is an important special case, namely when H is uniformly distributed. In this case, $P_{H|Y}(i|y)$, as a function of i , is proportional to $f_{Y|H}(y|i)/m$. Therefore, the argument that maximizes $P_{H|Y}(i|y)$ also maximizes $f_{Y|H}(y|i)$. Then the MAP decision rule is equivalent to *the maximum likelihood (ML) decision rule*:

$$\hat{H}(y) = \arg \max_i f_{Y|H}(y|i) \quad (\text{ML decision rule}). \quad (2.2)$$

2.2.1 Binary Hypothesis Testing

The special case in which we have to make a binary decision, i.e., $H \in \mathcal{H} = \{0, 1\}$, is both instructive and of practical relevance. Since there are only two alternatives to be tested, the MAP test may now be written as

$$\begin{array}{c} \hat{H} = 1 \\ \frac{f_{Y|H}(y|1)P_H(1)}{f_Y(y)} \geq \frac{f_{Y|H}(y|0)P_H(0)}{f_Y(y)} \\ \hat{H} = 0 \end{array}.$$

An equivalent rule is

$$\Lambda(y) = \frac{f_{Y|H}(y|1)}{f_{Y|H}(y|0)} \underset{\hat{H} = 0}{\overset{\hat{H} = 1}{\geq}} \frac{P_H(0)}{P_H(1)} = \eta \quad (\text{binary MAP rule}). \quad (2.3)$$

The left side of the above test is called the *likelihood ratio* denoted by $\Lambda(y)$ whereas the right side is the *threshold* η . Notice that if $P_H(0)$ increases, so does the threshold. In turn the region $\{y : \hat{H}(y) = 0\}$ becomes bigger. This is intuitive.

When $P_H(0) = P_H(1) = 1/2$ the threshold becomes unity and the MAP test becomes a ML test that may be written as

$$\begin{array}{c} \hat{H} = 1 \\ f_{Y|H}(y|1) \underset{\hat{H} = 0}{\geq} f_{Y|H}(y|0) \\ \hat{H} = 0 \end{array} \quad (\text{binary ML rule}).$$

The decoding region \mathcal{R}_i is the set of y for which the decision is $\hat{H} = i$, $i \in \{0, 1\}$.

To compute the probability of error it is often convenient to compute the error probability for each hypothesis and then take the average. When $H = 0$, we make an incorrect decision if $Y \in \mathcal{R}_1$ or, equivalently, if $\Lambda(y) \geq \eta$. Hence, denoting by $P_e(i)$ the probability of making an error when $H = i$,

$$P_e(0) = Pr\{Y \in \mathcal{R}_1 | H = 0\} = \int_{\mathcal{R}_1} f_{Y|H}(y|0) dy \quad (2.4)$$

$$= Pr\{\Lambda(Y) \geq \eta | H = 0\}. \quad (2.5)$$

Whether it is easier to work with the right side of (2.4) or of (2.5) depends on whether it is easier to work with the conditional density of Y or of $\Lambda(Y)$. We will see examples of both cases.

Similar expressions hold for the probability of error conditioned on $H = 1$, denoted by $P_e(1)$. The unconditional error probability is then

$$P_e = P_e(1)p_H(1) + P_e(0)p_H(0).$$

From (2.3) we see that, for the purpose of performing a MAP test, having $\Lambda(Y)$ is as good as having the observable Y . Any random variable obtained from Y that has this property is called a *sufficient statistic*.

2.3 The Q Function

The Q function is defined as:

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{\xi^2}{2}} d\xi.$$

Hence, if $Z \sim \mathcal{N}(0, 1)$ (meaning that Z is a Normally distributed zero-mean random variable of unit variance) then $Pr\{Z \geq x\} = Q(x)$.

If $Z \sim \mathcal{N}(m, \sigma^2)$, then the probability $Pr\{Z \geq x\}$ can be written using the Q function by noticing that $\{Z \geq x\}$ is equivalent to $\{\frac{Z-m}{\sigma} \geq \frac{x-m}{\sigma}\}$. But $\frac{Z-m}{\sigma} \sim \mathcal{N}(0, 1)$. Hence $Pr\{Z \geq x\} = Q(\frac{x-m}{\sigma})$. Make sure you are familiar with these steps. We will use them frequently.

We now describe some of the key properties of $Q(x)$.

- (a) If $Z \sim \mathcal{N}(0, 1)$, $F_Z(z) = Pr\{Z \leq z\} = 1 - Q(z)$. (Draw a picture that expresses this relationship in terms of areas under the probability density function of Z .)
- (b) $Q(0) = 1/2$, $Q(-\infty) = 1$, $Q(\infty) = 0$.
- (c) $Q(-x) + Q(x) = 1$. (Again, draw a picture.)
- (d) $\frac{1}{\sqrt{2\pi\alpha}} e^{-\frac{\alpha^2}{2}} (1 - \frac{1}{\alpha^2}) < Q(\alpha) < \frac{1}{\sqrt{2\pi\alpha}} e^{-\frac{\alpha^2}{2}}$, $\alpha > 0$.
- (e) An alternative expression with fixed integration limits is $Q(x) = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} e^{-\frac{x^2}{2\sin^2\theta}} d\theta$. It holds for $x \geq 0$.
- (f) $Q(\alpha) \leq \frac{1}{2} e^{-\frac{\alpha^2}{2}}$, $\alpha \geq 0$.

Proofs: The proofs of (a), (b), and (c) are immediate (a picture suffices). The proof of part (d) is omitted. To prove (e), let $X \sim \mathcal{N}(0, 1)$ and $Y \sim \mathcal{N}(0, 1)$ be independent. Hence $Pr\{X \geq 0, Y \geq \xi\} = Q(0)Q(\xi) = \frac{Q(\xi)}{2}$.

Using Polar coordinates

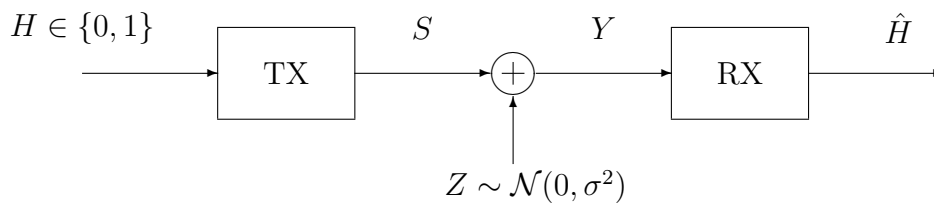
$$\frac{Q(\xi)}{2} = \int_0^{\frac{\pi}{2}} \int_{\frac{\xi}{\sin\theta}}^{\infty} \frac{e^{-\frac{r^2}{2}}}{2\pi} r dr d\theta = \frac{1}{2\pi} \int_0^{\frac{\pi}{2}} \int_{\frac{\xi^2}{2\sin^2\theta}}^{\infty} e^{-t} dt d\theta = \frac{1}{2\pi} \int_0^{\frac{\pi}{2}} e^{-\frac{\xi^2}{2\sin^2\theta}} d\theta.$$

To prove (f) we use (e) and the fact that $e^{-\frac{\xi^2}{2\sin^2\theta}} \leq e^{-\frac{\xi^2}{2}}$ for $\theta \in [0, \frac{\pi}{2}]$. Hence

$$Q(\xi) \leq \frac{1}{\pi} \int_0^{\frac{\pi}{2}} e^{-\frac{\xi^2}{2}} d\theta = \frac{1}{2} e^{-\frac{\xi^2}{2}}.$$

2.4 Binary Communication Across the Scalar Gaussian Channel

We consider the following setup



We assume that the transmitter maps $H = 0$ into $a \in \mathbb{R}$ and $H = 1$ into $b \in \mathbb{R}$. The output statistic for the various hypotheses is as follows:

$$H = 0 : Y \sim \mathcal{N}(a, \sigma^2)$$

$$H = 1 : Y \sim \mathcal{N}(b, \sigma^2).$$

An equivalent way to say this is

$$f_{Y|H}(y | 0) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y-a)^2}{2\sigma^2} \right\}$$

$$f_{Y|H}(y | 1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y-b)^2}{2\sigma^2} \right\}.$$

We compute the likelihood ratio

$$\Lambda(y) = \frac{f_{Y|H}(y | 1)}{f_{Y|H}(y | 0)} = \exp \left\{ -\frac{(y-b)^2 - (y-a)^2}{2\sigma^2} \right\} = \exp \left\{ \frac{b-a}{\sigma^2} \left(y - \frac{a+b}{2} \right) \right\}.$$

The threshold is $\eta = \frac{P_0}{P_1}$. Now we have all the ingredients for the MAP rule. Comparing $\Lambda(y)$ to η is the same as comparing $\log \Lambda(y)$ to $\log \eta$. The function $\log \Lambda(y)$ is called *log likelihood ratio*. Hence the MAP decision rule is

$$\begin{array}{l} \hat{H} = 1 \\ \frac{b-a}{\sigma^2} \left(y - \frac{a+b}{2} \right) \geq \ln \eta \\ \hat{H} = 0 \end{array}$$

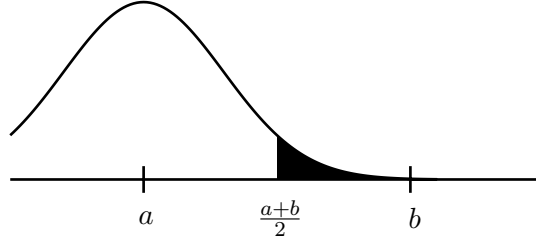


Figure 2.3: The probability of error when $H = 0$ is the black area. It is the probability that the noise makes y exceed the threshold when $H = 0$. The value of the threshold, half way between a and b , is determined assuming $P_H(0) = P_H(1)$.

If $b > a$, then we can divide both sides by $\frac{b-a}{\sigma^2}$ without changing the outcome of the above comparison. In this case we obtain

$$\hat{H}_{\text{MAP}}(y) = \begin{cases} 1, & y > \theta \\ 0, & \text{otherwise,} \end{cases}$$

where $\theta = \frac{\sigma^2}{b-a} \ln \eta + \frac{a+b}{2}$. Notice that if $P_H(0) = P_H(1)$, then $\ln \eta = 0$ and the threshold θ becomes the midpoint $\frac{a+b}{2}$.

We now determine the probability of error. Recall that

$$P_e(0) = Pr\{Y > \theta \mid H = 0\} = \int_{\mathcal{R}_1} f_{Y|H}(y \mid 0) dy.$$

This is the probability that a Gaussian random variable with mean a and variance σ^2 exceeds the threshold θ . From our review on the Q function we know immediately that $P_e(0) = Q\left(\frac{\theta-a}{\sigma}\right)$. Similarly, $P_e(1) = Q\left(\frac{b-\theta}{\sigma}\right)$. Finally, $P_e = P_H(0)Q\left(\frac{\theta-a}{\sigma}\right) + P_H(1)Q\left(\frac{b-\theta}{\sigma}\right)$.

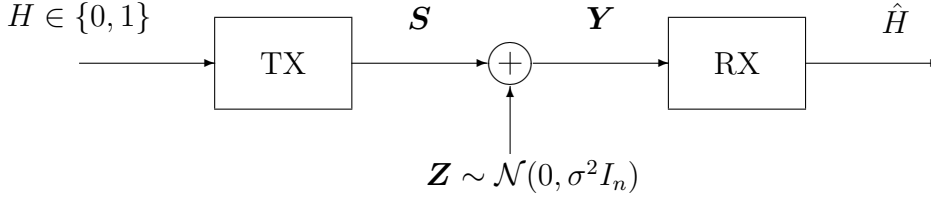
The most common case is when $P_H(0) = P_H(1) = 1/2$. Then $\frac{\theta-a}{\sigma} = \frac{b-\theta}{\sigma} = \frac{b-a}{2\sigma} = \frac{d}{2\sigma}$, where d is the distance between a and b . In this case

$$P_e = Q\left(\frac{d}{2\sigma}\right).$$

Figure 2.3, which holds for the ML decision rule, leads immediately to P_e . Make sure that you understand it. It will be used frequently.

2.5 Binary Communication Across the Vector Gaussian Channel

The setup is the same as for the scalar case except that the transmitter output \mathbf{s} , the noise \mathbf{z} , and the observation \mathbf{y} are now n -tuples over \mathbb{R} . The new setting is represented in the figure below. Before going on we recommend reviewing the background material in Appendices 2.C and 2.E



We now assume that the hypothesis i is mapped into the transmitter output $X(i)$ defined by

$$X(i) = \begin{cases} \mathbf{a} \in \mathbb{R}^n, & i = 0 \\ \mathbf{b} \in \mathbb{R}^n, & i = 1. \end{cases}$$

We also assume that $\mathbf{Z} \sim \mathcal{N}(0, \sigma^2 I_n)$.

As we did earlier, we start writing down the output statistic for each hypothesis

$$H = 0 : \quad \mathbf{Y} = \mathbf{a} + \mathbf{Z} \sim \mathcal{N}(\mathbf{a}, \sigma^2 I_n)$$

$$H = 1 : \quad \mathbf{Y} = \mathbf{b} + \mathbf{Z} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I_n).$$

Recall that

$$f_{\mathbf{Z}}(\mathbf{z}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{z_i^2}{2\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{\|\mathbf{z}\|^2}{2\sigma^2}}.$$

Similarly,

$$f_{\mathbf{Y}|H}(\mathbf{y} | 0) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{\|\mathbf{y}-\mathbf{a}\|^2}{2\sigma^2}}$$

$$f_{\mathbf{Y}|H}(\mathbf{y} | 1) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{\|\mathbf{y}-\mathbf{b}\|^2}{2\sigma^2}}.$$

Hence

$$\Lambda(\mathbf{y}) = \frac{f_{\mathbf{Y}|H}(\mathbf{y} | 1)}{f_{\mathbf{Y}|H}(\mathbf{y} | 0)} = \exp \left\{ \frac{\|\mathbf{y} - \mathbf{a}\|^2 - \|\mathbf{y} - \mathbf{b}\|^2}{2\sigma^2} \right\},$$

and

$$LLR(\mathbf{y}) = \frac{\|\mathbf{y} - \mathbf{a}\|^2 - \|\mathbf{y} - \mathbf{b}\|^2}{2\sigma^2} \quad (2.6)$$

$$= \frac{\|\mathbf{a}\|^2 - \|\mathbf{b}\|^2}{2\sigma^2} + \frac{1}{\sigma^2} \langle \mathbf{y}, \mathbf{b} - \mathbf{a} \rangle \quad (2.7)$$

$$= \left\langle \mathbf{y} - \frac{\mathbf{a} + \mathbf{b}}{2}, \frac{\mathbf{b} - \mathbf{a}}{\|\mathbf{b} - \mathbf{a}\|^2} \right\rangle \frac{\|\mathbf{b} - \mathbf{a}\|^2}{\sigma^2}, \quad (2.8)$$

where in the last equation we used the fact that for real-valued vectors \mathbf{a} and \mathbf{b} , $\langle \mathbf{a} + \mathbf{b}, \mathbf{a} - \mathbf{b} \rangle = \|\mathbf{a}\|^2 - \|\mathbf{b}\|^2$.

From (2.7) the MAP rule is

$$\begin{aligned} \hat{H} &= 1 \\ \langle \mathbf{y}, \mathbf{b} - \mathbf{a} \rangle &\geq T, \\ \hat{H} &= 0 \end{aligned}$$

where $T = \sigma^2 \ln \eta + \frac{\|\mathbf{b}\|^2 - \|\mathbf{a}\|^2}{2}$ is a threshold and $\eta = \frac{P_H(0)}{P_H(1)}$. This says that \mathcal{R}_0 and \mathcal{R}_1 are separated by the hyperplane

$$\{\mathbf{y} \in \mathbb{R}^n : \langle \mathbf{y}, \mathbf{b} - \mathbf{a} \rangle = T\},$$

When $P_H(0) = P_H(1) = 1/2$, the separating hyperplane separates the points that are closer to \mathbf{a} from those that are closer to \mathbf{b} . We see this by solving for \mathbf{y} in

$$LLR(\mathbf{y}) = \ln \eta$$

when $\ln \eta = 0$. The \mathbf{y} that satisfy this relationship are the ones for which

$$\|\mathbf{y} - \mathbf{a}\|^2 - \|\mathbf{y} - \mathbf{b}\|^2 = 0.$$

These are the \mathbf{y} that are at the same distance from \mathbf{a} and from \mathbf{b} . Hence the ML decision rule for the AWGN channel decides for the transmitted vector that is closer to the observed vector.

We also see that the separating hyperplane moves towards \mathbf{b} when ϕ increases, which is the case when $\frac{P_H(0)}{P_H(1)}$ increases. This makes sense: if the prior probability becomes more in favor of $H = 0$ then the decoding region \mathcal{R}_0 becomes larger. Moreover, if $\frac{P_H(0)}{P_H(1)}$ exceeds 1, then $\ln \eta$ is positive and ϕ increases with σ^2 . This also makes sense: as the observation becomes noisier, we pay more attention to the prior (which favors $H = 0$).

2.6 Multi-Hypothesis Testing

In Section 2.2 we have defined the hypothesis testing problem and derived the maximum a posteriori (MAP) and maximum likelihood (ML) decision rules. This was done for the

general case of m hypotheses, that is when $\mathcal{H} = \{0, 1, \dots, (m-1)\}$. We then turned our attention to binary hypotheses, i.e. $\mathcal{H} = \{0, 1\}$, and deepened our understanding paying particular attention to the special case in which the observation Y is a Gaussian random variable (or random vector \mathbf{Y}) whose mean depends on H . Now we go back to the m hypothesis testing problem.

Recall that the MAP decision rule, which minimizes the probability of making an error, is

$$\begin{aligned}\hat{H}_{MAP}(\mathbf{y}) &= \arg \max_i P_{H|\mathbf{Y}}(i|\mathbf{y}) \\ &= \arg \max_i \frac{f_{\mathbf{Y}|H}(\mathbf{y}|i)P_H(i)}{f_{\mathbf{Y}}(\mathbf{y})} \\ &= \arg \max_i f_{\mathbf{Y}|H}(\mathbf{y}|i)P_H(i),\end{aligned}$$

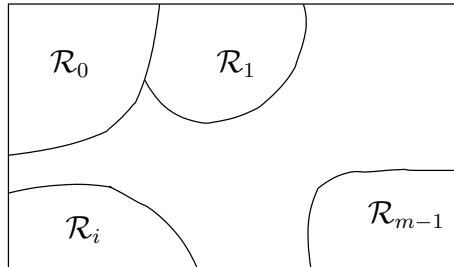
where $f_{\mathbf{Y}|H}(\cdot|i)$ is the probability density function of the observable \mathbf{Y} when the hypothesis is i and $P_H(i)$ is the probability of the i th hypothesis. This rule is well defined up to ties. If there is more than one i that achieves the maximum on the right side of one (and thus all) of the above expressions, then we may decide for any such i without affecting the probability of error. If we want the decision rule to be unambiguous, we can agree that in case of ties we pick the largest i that achieves the maximum.

When all hypotheses have the same probability, then the MAP rule specializes to the ML rule, i.e.,

$$\hat{H}_{ML}(\mathbf{y}) = \arg \max_i f_{\mathbf{Y}|H}(\mathbf{y}|i).$$

We will always assume that $f_{\mathbf{Y}|H}$ is known. If the transmitter maps the hypothesis i into the channel input \mathbf{s}_i , then $f_{\mathbf{Y}|H}(\mathbf{y}|i) = f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{s}_i)$, where $f_{\mathbf{Y}|\mathbf{X}}(\cdot|\mathbf{x})$, also denoted by $f_{\mathbf{Y}|\mathbf{x}}$, is the probability density function of the channel output when the channel input is \mathbf{x} .

Note that the decision (or decoding) function \hat{H} assigns an $i \in \mathcal{H}$ to each $\mathbf{y} \in \mathbb{R}^n$. It can be equivalently described by the decision (or decoding) regions \mathcal{R}_i , $i \in \mathcal{H}$, where \mathcal{R}_i consists of those \mathbf{y} for which $\hat{H}(\mathbf{y}) = i$. It is convenient to think of \mathbb{R}^n as being partitioned by decoding regions as depicted in the following figure.



We use the decoding regions to express the error probability P_e or, equivalently, the probability of deciding correctly P_c .

$$\begin{aligned} P_e(i) &= 1 - P_c(i) \\ &= 1 - \int_{\mathcal{R}_i} f_{\mathbf{Y}|H}(\mathbf{y}|i) d\mathbf{y}. \end{aligned}$$

Now assume the AWGN channel. When $H = i$, $i \in \mathcal{H}$, let $\mathbf{S} = \mathbf{s}_i$. Assume $P_H(i) = \frac{1}{m}$ (this is a common assumption in communications). The ML decision rule is

$$\begin{aligned} \hat{H}_{ML}(\mathbf{y}) &= \arg \max_i f_{\mathbf{Y}|H}(\mathbf{y}|i) \\ &= \arg \max_i \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{\|\mathbf{y} - \mathbf{s}_i\|^2}{2\sigma^2}\right\} \\ &= \arg \min_i \|\mathbf{y} - \mathbf{s}_i\|^2. \end{aligned}$$

Hence a ML decision rule for the AWGN channel is a minimum-distance decision rule as shown in Figure 2.4. Up to ties, \mathcal{R}_i corresponds to the *Voronoi region* of \mathbf{s}_i , defined as the set of points in \mathbb{R}^n that are at least as close to \mathbf{s}_i as to any other \mathbf{s}_j .

EXAMPLE 5. (PAM) Figure 2.5 shows the signal points and the decoding regions of a ML decoder for 6-ary Pulse Amplitude Modulation (why the name makes sense will become clear in the next chapter), assuming that the channel is the AWGN channel. The signal points are elements of \mathbb{R} and the ML decoder chooses according to the minimum-distance rule. When the hypothesis is $H = 0$, the receiver makes the wrong decision if the observation $y \in \mathbb{R}$ falls outside the decoding region \mathcal{R}_0 . This is the case if the noise $Z \in \mathbb{R}$ is larger than $d/2$, where $d = s_i - s_{i-1}$, $i = 1, \dots, 5$. Thus

$$P_e(0) = \Pr\left\{Z > \frac{d}{2}\right\} = Q\left(\frac{d}{2\sigma}\right).$$

By symmetry, $P_e(5) = P_e(0)$. For $i \in \{1, 2, 3, 4\}$, the probability of error when $H = i$ is the probability that the event $\{Z \geq \frac{d}{2}\} \cup \{Z < -\frac{d}{2}\}$ occurs. This even is the union

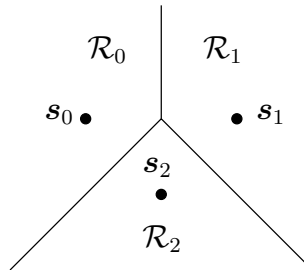


Figure 2.4: Example of Voronoi regions.

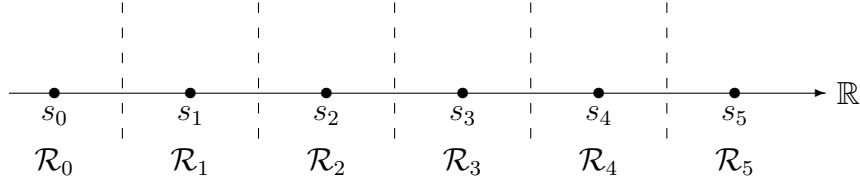


Figure 2.5: PAM signal constellation.

of disjoint events. Its probability is the sum of the probability of the individual events. Hence

$$P_e(i) = Pr\{\{Z \geq \frac{d}{2}\} \cup \{Z < -\frac{d}{2}\}\} = 2Pr\{Z \geq \frac{d}{2}\} = 2Q\left(\frac{d}{2\sigma}\right), \quad i \in \{1, 2, 3, 4\}.$$

Finally,

$$P_e = \frac{2}{6}Q\left(\frac{d}{2\sigma}\right) + \frac{4}{6}2Q\left(\frac{d}{2\sigma}\right) = \frac{5}{3}Q\left(\frac{d}{2\sigma}\right).$$

□

EXAMPLE 6. (4-ary QAM) Figure 2.6 shows the signal set $\{\mathbf{s}_0, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3\}$ for 4-ary Quadrature Amplitude Modulation (QAM). We may consider signals as points in \mathbb{R}^2 or in \mathbb{C} . We choose the former since we don't know how to deal with complex valued noise yet. The noise is $\mathbf{Z} \sim \mathcal{N}(0, \sigma^2 I_2)$ and the observable, when $H = i$, is $\mathbf{Y} = \mathbf{s}_i + \mathbf{Z}$. We assume that the receiver implements a ML decision rule, which for the AWGN channel means minimum-distance decoding. The decoding region for \mathbf{s}_0 is the first quadrant, for \mathbf{s}_1 the second quadrant, etc.. When $H = 0$, the decoder makes the correct decision if $\{Z_1 > -\frac{d}{2}\} \cap \{Z_2 \geq -\frac{d}{2}\}$, where d is the minimum distance among signal points. This is the intersection of independent events. Hence the probability of the intersection is the product of the probability of each event, i.e.

$$P_c(0) = \left[Pr\left\{Z_i \geq -\frac{d}{2}\right\}\right]^2 = Q^2\left(-\frac{d}{2\sigma}\right) = \left[1 - Q\left(\frac{d}{2\sigma}\right)\right]^2.$$

By symmetry, for all i , $P_c(i) = P_c(0)$. Hence,

$$P_e = P_e(0) = 1 - P_c(0) = 2Q\left(\frac{d}{2\sigma}\right) - Q^2\left(\frac{d}{2\sigma}\right).$$

When the channel is Gaussian and the decoding regions are bounded by affine planes, like in this and the previous example, one can express the error probability by means of the Q function. Another observation is worth mentioning: In this example we decided to focus on computing $P_c(0)$. It would have been possible to compute $P_e(0)$ instead of $P_c(0)$ but it would have costed a bit more work. To compute $P_e(0)$ we evaluate the

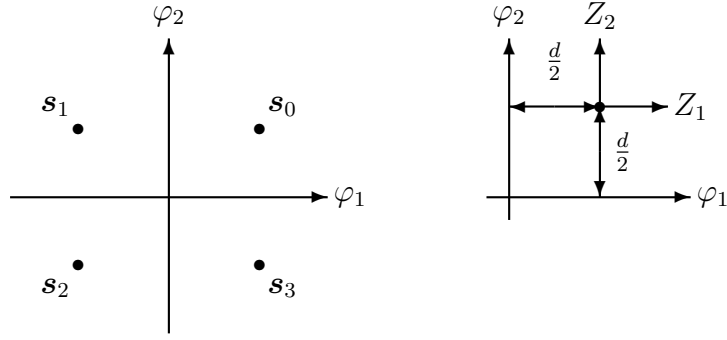


Figure 2.6: QAM signal constellation.

probability of the union $\{Z_1 \leq -\frac{d}{2}\} \cup \{Z_2 \leq -\frac{d}{2}\}$. These are not disjoint events. In fact they are independent events that can very well occur together. Thus the probability of the union is not the sum of the individual probabilities. Computing the probability of the union is not difficult but requires slightly more work than obtaining the probability of the intersection needed to determine $P_e(0)$. (In fact you are encouraged to verify the details.) \square

2.7 Union of Events Bound

Here is a simple and extremely useful bound. Recall that for general events \mathcal{A}, \mathcal{B}

$$\begin{aligned} P(\mathcal{A} \cup \mathcal{B}) &= P(\mathcal{A}) + P(\mathcal{B}) - P(\mathcal{A} \cap \mathcal{B}) \\ &\leq P(\mathcal{A}) + P(\mathcal{B}). \end{aligned}$$

More generally, using induction, we obtain the the *Union of Events Bound*

$$P\left(\bigcup_{i=1}^M \mathcal{A}_i\right) \leq \sum_{i=1}^M P(\mathcal{A}_i), \quad (UEB)$$

that applies to any collection of sets \mathcal{A}_i , $i = 1, \dots, M$. We now apply the union of events bound to approximate the probability of error in multi-hypothesis testing. Recall that

$$P_e(i) = \Pr\{\mathbf{Y} \notin \mathcal{R}_i | H = i\} = \int_{\mathcal{R}_i^c} f_{\mathbf{Y}|H}(\mathbf{y}|i) d\mathbf{y},$$

where \mathcal{R}_i^c denotes the complement of \mathcal{R}_i . If we are able to evaluate the above integral for every i , then we are able to determine the probability of error exactly. The bound that we derive is useful if we are unable to evaluate the above integral.

For $i \neq j$ define

$$\mathcal{B}_{i,j} = \{\mathbf{y} : P_H(j) f_{\mathbf{Y}|H}(\mathbf{y}|j) \geq P_H(i) f_{\mathbf{Y}|H}(\mathbf{y}|i)\}.$$

$\mathcal{B}_{i,j}$ is the set of \mathbf{y} for which the a posteriori probability when $H = j$ is at least as high as when $H = i$. Moreover,

$$\mathcal{R}_i^c \subseteq \bigcup_{j:j \neq i} \mathcal{B}_{i,j},$$

with equality if ties are always resolved against i . In fact, by definition, the right side contains all the ties whereas the left side may or may not contain them. Here ties refers to those \mathbf{y} for which equality holds in the definition of $\mathcal{B}_{i,j}$.

Now we use the union of events bound:

$$\begin{aligned} P_e(i) &= Pr \{ \mathbf{Y} \in \mathcal{R}_i^c | H = i \} \\ &\leq Pr \left\{ \mathbf{Y} \in \bigcup_{j:j \neq i} \mathcal{B}_{i,j} | H = i \right\} \\ &\leq \sum_{j:j \neq i} Pr \{ \mathbf{Y} \in \mathcal{B}_{i,j} | H = i \} \\ &= \sum_{j:j \neq i} \int_{\mathcal{B}_{i,j}} f_{\mathbf{Y}|H}(\mathbf{y}|i) d\mathbf{y}. \end{aligned} \tag{2.9}$$

To see exactly how we have applied the union of events bound, the second row above should be read as $P(\bigcup_{j:j \neq i} \{ \mathbf{Y} \in \mathcal{B}_{i,j} \})$ where where $P(\mathcal{A})$ is the probability of the event $\{ \mathbf{Y} \in \mathcal{A} \}$ conditioned on $H = i$.

What we have gained is that it is typically easier to integrate over $\mathcal{B}_{i,j}$ than over \mathcal{R}_j^c . For instance, for the AWGN channel and ML decision rule

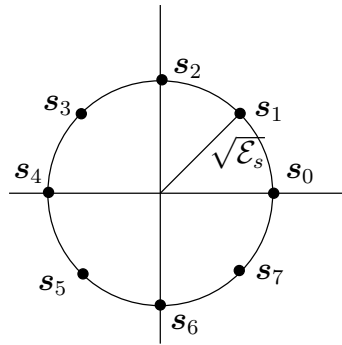
$$\int_{\mathcal{B}_{i,j}} f_{\mathbf{Y}|H}(\mathbf{y}|i) d\mathbf{y} = Q \left(\frac{\| \mathbf{s}_j - \mathbf{s}_i \|}{2\sigma} \right).$$

Moreover, in the next section we derive an easy-to-compute tight upperbound on

$$\int_{\mathcal{B}_{i,j}} f_{\mathbf{Y}|H}(\mathbf{y}|i) d\mathbf{y}$$

for a general $f_{\mathbf{Y}|H}$. Notice that the above integral is the probability of error under $H = i$ when there are only two hypotheses and the other hypothesis is $H = j$.

EXAMPLE 7. (*m*-PSK) The figure below shows a signal set for *m*-ary PSK (phase-shift keying) when *m* = 8.



Formally, the signal transmitted when $H = i$, $i \in \mathcal{H} = \{0, 1, \dots, m-1\}$, is

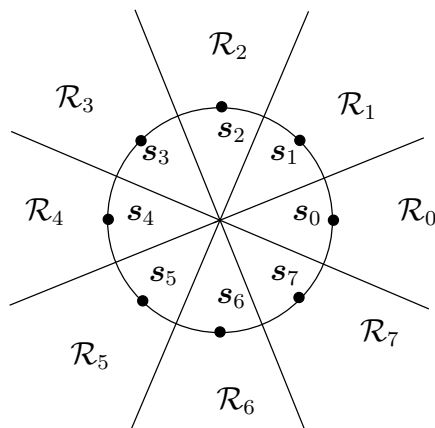
$$\mathbf{s}_i = \sqrt{\mathcal{E}_s} \begin{pmatrix} \cos 2\pi i/m \\ \sin 2\pi i/m \end{pmatrix}.$$

The hypothesis testing problem is specified by

$$H = i : \mathbf{Y} \sim \mathcal{N}(\mathbf{s}_i, \sigma^2 I_2)$$

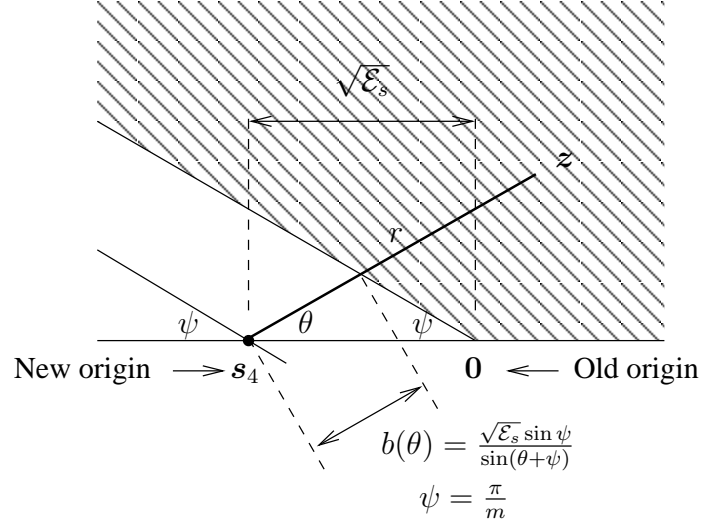
and the prior $P_H(i)$ is assumed to be uniformly distributed.

Since we have a uniform prior, the MAP and the ML decision rule are identical. Due to the circular symmetry of the additive noise, the ML decoder is a minimum-distance decoder. The decoding regions (up to ties) are shown in the picture below.



Now we proceed to compute the error probability. By symmetry, $P_e(i)$ is independent of i . Hence $P_e = P_e(i)$. To determine $P_e(i)$, it is convenient to put the coordinate system at \mathbf{s}_i as shown in the figure below.

$$\begin{aligned} P_e(i) &= 2Pr\{\mathbf{Z} \in \text{shaded area}\} \\ &= 2 \int_{\mathbf{z} \in \text{shaded area}} \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{\|\mathbf{z}\|^2}{2\sigma^2}\right\} d\mathbf{z}. \end{aligned}$$



Passing to polar coordinates with

$$z_1 = r \cos \theta$$

$$z_2 = r \sin \theta$$

we obtain

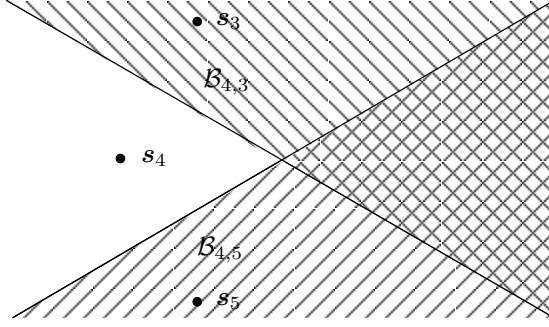
$$\begin{aligned}
 P_e(i) &= 2 \int_0^{\pi-\psi} \int_{b(\theta)}^{\infty} \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{r^2}{2\sigma^2}\right\} r dr d\theta \\
 &= \int_0^{\pi-\psi} \frac{1}{\pi} \exp\left\{-\frac{b^2(\theta)}{2\sigma^2}\right\} d\theta \\
 &= \frac{1}{\pi} \int_0^{\pi-\psi} \exp\left\{-\frac{\sin^2 \psi}{\sin^2(\theta + \psi)} \frac{\mathcal{E}_s}{2\sigma^2}\right\} d\theta. \quad (\text{Exact analysis of PSK}).
 \end{aligned}$$

This is as far as we can go with the exact analysis.

Now we use the union of events bound to determine an upperbound to the error probability. With reference to the figure below we have:

$$\begin{aligned}
 P_e(i) &= Pr\{\mathbf{Y} \in \mathcal{B}_{i,i-1} \cup \mathcal{B}_{i,i+1} | H = i\} \\
 &\leq Pr\{\mathbf{Y} \in \mathcal{B}_{i,i-1} | H = i\} + Pr\{\mathbf{Y} \in \mathcal{B}_{i,i+1} | H = i\} \\
 &= 2Pr\{\mathbf{Y} \in \mathcal{B}_{i,i-1} | H = i\} \\
 &= 2Q\left(\frac{\|\mathbf{s}_i - \mathbf{s}_{i-1}\|}{2\sigma}\right) \\
 &= 2Q\left(\frac{\sqrt{\mathcal{E}_s}}{\sigma} \sin \psi\right).
 \end{aligned}$$

Notice that we have been using a version of the union of events bound adapted to the problem: we are getting a tighter bound by using the fact that $\mathcal{R}_i^c = \mathcal{B}_{i,i-1} \cup \mathcal{B}_{i,i+1}$ rather than $\mathcal{R}_i^c \subset \cup_{j \neq i} \mathcal{B}_{i,j}$.



How good is the upperbound? Notice that

$$P_e = Pr\{\mathbf{Y} \in \mathcal{B}_{i,i-1}|H = i\} + Pr\{\mathbf{Y} \in \mathcal{B}_{i,i+1}|H = i\} - Pr\{\mathbf{Y} \in \mathcal{B}_{i,i-1} \cap \mathcal{B}_{i,i+1}|H = i\}$$

and we obtained an upper bound by lower-bounding the last term with 0. We now obtain a lower bound to P_e by upperbounding the same term. To do so, observe from the above picture that $\mathcal{B}_{i,i-1} \cap \mathcal{B}_{i,i+1}$ is the decoding region of the point which is furthest away from \mathbf{s}_i . Hence the probability that \mathbf{y} ends up in that region given that \mathbf{s}_i was sent is smaller than the probability that \mathbf{y} end up in any of the remaining regions not counting the decoding region of \mathbf{s}_i . There are $M - 1$ regions leading to an error and the probability that \mathbf{y} is in one of them is P_e . Hence the probability that it is in the one of the $M - 1$ regions that has the smallest probability can not exceed $\frac{P_e}{M-1}$:

$$Pr\{\mathbf{Y} \in (\mathcal{B}_{i,i-1} \cap \mathcal{B}_{i,i+1})|H = i\} \leq \frac{P_e(i)}{m-1} = \frac{P_e}{m-1}.$$

Hence,

$$\begin{aligned} P_e &= Pr\{\mathbf{Y} \in \mathcal{B}_{i,i-1}|H = i\} + Pr\{\mathbf{Y} \in \mathcal{B}_{i,i+1}|H = i\} - Pr\{\mathbf{Y} \in \mathcal{B}_{4,3} \cap \mathcal{B}_{i,i+1}|H = i\} \\ &\geq 2Q\left(\sqrt{\frac{\mathcal{E}_s}{\sigma^2}} \sin \psi\right) - \frac{P_e}{m-1}. \end{aligned}$$

Solving for P_e we obtain the desired lower bound

$$P_e \geq 2Q\left(\sqrt{\frac{\mathcal{E}_s}{\sigma^2}} \sin \psi\right) \frac{m-1}{m}.$$

The ratio between the upper and the lower bound is the constant $\frac{m}{m-1}$. For m large, the bounds become very tight. One can come up with lower bounds for which this ratio goes to 1 as $\mathcal{E}_s/\sigma^2 \rightarrow \infty$. One such bound is obtained by upperbounding $Pr\{\mathbf{Y} \in \mathcal{B}_{i,i-1} \cap \mathcal{B}_{i,i+1}|H = i\}$ with the probability $Q\left(\frac{\sqrt{\mathcal{E}_s}}{\sigma}\right)$ that Y_1 is positive given $H = i$. \square

2.8 Union Bhattacharyya Bound

Let us summarize. From the union of events bound applied to

$$\mathcal{R}_i^c \subseteq \bigcup_{j:j \neq i} \mathcal{B}_{i,j}$$

we have obtained the upper bound

$$\begin{aligned} P_e(i) &= Pr\{\mathbf{Y} \in \mathcal{R}_i^c | H = i\} \\ &\leq \sum_{j:j \neq i} Pr\{\mathbf{Y} \in \mathcal{B}_{i,j} | H = i\} \end{aligned}$$

and we have used this bound for the AWGN channel. What we have gained with the bound is that instead of having to compute

$$Pr\{\mathbf{Y} \in \mathcal{R}_i^c | H = i\} = \int_{\mathcal{R}_i^c} f_{\mathbf{Y}|H}(\mathbf{y}|i) d\mathbf{y},$$

which requires integrating over a possibly complicated region \mathcal{R}_i^c , we only have to compute

$$Pr\{\mathbf{Y} \in \mathcal{B}_{i,j} | H = i\} = \int_{\mathcal{B}_{i,j}} f_{\mathbf{Y}|H}(\mathbf{y}|i) d\mathbf{y}.$$

The latter integral is simply $Q(\frac{a}{\sigma})$, where a is the distance between \mathbf{s}_i and the hyperplane bounding $\mathcal{B}_{i,j}$. For a ML decision rule, $a = \frac{\|\mathbf{s}_i - \mathbf{s}_j\|}{2}$.

What if the channel is *not* AWGN? Is there a relatively simple expression for $Pr\{\mathbf{Y} \in \mathcal{B}_{i,j} | H = i\}$ that applies for general channels? Such an expression does exist. It is the *Bhattacharyya bound* that we now derive.³

Given a set \mathcal{A} , the indicator function $1_{\mathcal{A}}$ is defined as

$$1_{\mathcal{A}}(x) = \begin{cases} 1, & x \in \mathcal{A} \\ 0, & \text{otherwise.} \end{cases}$$

From

$$\mathcal{B}_{i,j} = \{\mathbf{y} \in \mathbb{R}^n : P_H(i) f_{\mathbf{Y}|H}(\mathbf{y}|i) \leq P_H(j) f_{\mathbf{Y}|H}(\mathbf{y}|j)\},$$

we immediately verify that

$$1_{\mathcal{B}_{i,j}}(\mathbf{y}) \leq \sqrt{\frac{P_H(j) f_{\mathbf{Y}|H}(\mathbf{y}|j)}{P_H(i) f_{\mathbf{Y}|H}(\mathbf{y}|i)}}.$$

With this we obtain the Bhattacharyya bound as follows:

$$\begin{aligned} Pr\{\mathbf{Y} \in \mathcal{B}_{i,j} | H = i\} &= \int_{\mathbf{y} \in \mathcal{B}_{i,j}} f_{\mathbf{Y}|H}(\mathbf{y}|i) d\mathbf{y} \\ &= \int_{\mathbf{y} \in \mathbb{R}^n} f_{\mathbf{Y}|H}(\mathbf{y}|i) 1_{\mathcal{B}_{i,j}}(\mathbf{y}) d\mathbf{y} \\ &\leq \int_{\mathbf{y} \in \mathbb{R}^n} f_{\mathbf{Y}|H}(\mathbf{y}|i) \sqrt{\frac{P_H(j) f_{\mathbf{Y}|H}(\mathbf{y}|j)}{P_H(i) f_{\mathbf{Y}|H}(\mathbf{y}|i)}} d\mathbf{y} \\ &= \sqrt{\frac{P_H(j)}{P_H(i)}} \int_{\mathbf{y} \in \mathbb{R}^n} \sqrt{f_{\mathbf{Y}|H}(\mathbf{y}|i) f_{\mathbf{Y}|H}(\mathbf{y}|j)} d\mathbf{y}. \end{aligned} \quad (2.10)$$

³There are two versions of the Bhattacharyya bound. Here we derive the one that has the simpler derivation. The other version, which is tighter by a factor 2, is left as an exercise.

What makes the last integral appealing is that we integrate over the entire \mathbb{R}^n . As shown in Problem 22 (Bhattacharyya Bound for DMCs), for *discrete memoryless channels* the bound further simplifies.

As the name indicates, the *Union Bhattacharyya bound* is the union of events bound (2.10). Inserting yields

$$P_e(i) \leq \sum_{j:j \neq i} Pr\{\mathbf{Y} \in \mathcal{B}_{i,j} | H = i\} \leq \sum_{j:j \neq i} \sqrt{\frac{P_H(j)}{P_H(i)}} \int_{\mathbf{y} \in \mathbb{R}^n} \sqrt{f_{\mathbf{Y}|H}(\mathbf{y}|i)f_{\mathbf{Y}|H}(\mathbf{y}|j)} d\mathbf{y}.$$

We can now remove the conditioning on $H = i$ and obtain

$$Pr\{e\} \leq \sum_i \sum_{j:j \neq i} \sqrt{P_H(i)P_H(j)} \int_{\mathbf{y} \in \mathbb{R}^n} \sqrt{f_{\mathbf{Y}|H}(\mathbf{y}|i)f_{\mathbf{Y}|H}(\mathbf{y}|j)} d\mathbf{y}.$$

EXAMPLE 8. (Tightness of the Bhattacharyya Bound) Consider the following scenario

$$\begin{aligned} H = 0 : \quad \mathbf{S} = \mathbf{s}_0 &= (0, 0, \dots, 0)^T \\ H = 1 : \quad \mathbf{S} = \mathbf{s}_1 &= (1, 1, \dots, 1)^T \end{aligned}$$

with $P_H(0) = 0.5$, and where the channel is the binary erasure channel described in the figure:

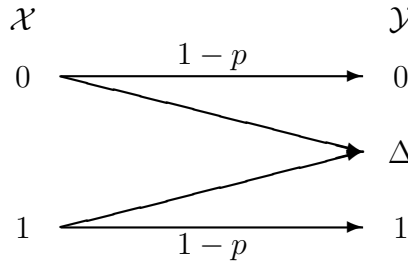


Figure 2.7: Binary erasure channel.

The Bhattacharyya bound is :

$$\begin{aligned} Pr\{\mathbf{Y} \in \mathcal{B}_{0,1} | H = 0\} &\leq \sum_{\mathbf{y} \in \{0,1,\Delta\}^n} \sqrt{P_{\mathbf{Y}|H}(\mathbf{y} | 1)P_{\mathbf{Y}|H}(\mathbf{y} | 0)} \\ &= \sum_{\mathbf{y} \in \{0,1,\Delta\}^n} \sqrt{P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} | \mathbf{s}_1)P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} | \mathbf{s}_0)} \\ &\stackrel{(a)}{=} \sqrt{P_{\mathbf{Y}|\mathbf{X}}((\Delta, \dots, \Delta)^T | \mathbf{s}_0)P_{\mathbf{Y}|\mathbf{X}}((\Delta, \dots, \Delta)^T | \mathbf{s}_1)} \\ &= p^n, \end{aligned}$$

where in (a) we used the fact that the first factor under the square root vanishes if \mathbf{y} contains ones and the second vanishes if \mathbf{y} contains zeros. The same bound applies for $H = 1$. Hence $P_e \leq \frac{1}{2}p^n + \frac{1}{2}p^n = p^n$.

If we use the tighter version of the union Bhattacharyya bound, which as mentioned earlier is tighter by a factor of 2, then we obtain

$$P_e \stackrel{(UBB)}{\leq} \frac{1}{2}p^n.$$

For the Binary Erasure Channel and the two codewords \mathbf{s}_0 and \mathbf{s}_1 we can actually compute the probability of error exactly:

$$P_e = \frac{1}{2}Pr\{\mathbf{Y} = (\Delta, \Delta, \dots, \Delta)^T\} = \frac{1}{2}p^n.$$

For this channel the Bhattacharyya bound is tight! □

2.9 Problems

PROBLEM 1. (Warmup Problem: Background Material)

- (a) Assume that X_1 and X_2 are independent random variables and that they are uniformly distributed in the interval $[0, 1]$. What is the probability that $X_1 + X_2 < 1$ and $X_2 \geq \frac{1}{2}$?
- (b) Let $\phi(t) = A \frac{\sin \frac{\pi}{T} t}{\frac{\pi}{T} t}$, $t \in \mathbb{R}$. Sketch $\phi(t)$ and its Fourier transform $\phi_{\mathcal{F}}(f)$. Label your figures appropriately.
- (c) Determine A so that $\phi(t)$ has unit energy.

PROBLEM 2. (Weather Frog)

Let us assume that a “weather frog” bases his forecast for tomorrow’s weather entirely on today’s air pressure. Determining a weather forecast is a hypothesis testing problem. For simplicity, let us assume that the weather frog only needs to tell us if the forecast for tomorrow’s weather is “sunshine” or “rain”. Hence we are dealing with a binary hypothesis testing problem. Let $H = 0$ mean “sunshine” and $H = 1$ mean “rain”. We will assume that both values of H are equally likely, i.e. $p_H(0) = p_H(1) = 1/2$.

Measurements over several years have led the weather frog to conclude that on a day that precedes sunshine the pressure may be modeled as a random variable y with the following probability density function:

$$f_{Y|H}(y|0) = \begin{cases} A - \frac{A}{2}y, & 0 \leq y \leq 1 \\ 0, & \text{otherwise.} \end{cases} \quad (2.11)$$

Similarly, the pressure on a day that precedes a rainy day is distributed according to

$$f_{Y|H}(y|1) = \begin{cases} B + \frac{B}{3}y, & 0 \leq y \leq 1 \\ 0, & \text{otherwise.} \end{cases} \quad (2.12)$$

The weather frog’s goal in life is to guess the value of H after measuring Y .

- (i) Determine A and B .
- (ii) Find the probability $p_{H|Y}(0|y)$ for all values of y . This probability is often called the a posteriori probability of hypothesis $H = 0$ given that $Y = y$. Also find the probability $p_{H|Y}(1|y)$ for all values of y . Hint: Use Bayes’ rule.
- (iii) Plot $p_{H|Y}(0|y)$ and $p_{H|Y}(1|y)$ as a function of y . Is it true that the decision rule may be written as

$$\hat{H}(y) = \begin{cases} 0, & \text{if } y \leq \theta \\ 1, & \text{otherwise,} \end{cases} \quad (2.13)$$

for some threshold θ ? If yes specify θ .

(iv) Determine, as a function of θ , the probability that the decision rule in (iii) decides $\hat{H} = 1$ when, in reality, $H = 0$. This probability is denoted $Pr(\hat{H}(y) = 1|H = 0)$.

(v) Determine, as a function of θ , the probability of error for the decision rule that you have derived in (iii). Evaluate your expression at the value of θ that you have found in (iii).

(vi) Among decision rules that compare the pressure y to a threshold like in Eqn. (2.13), is there a decision rule that results in a smaller probability of error than the rule derived in (iii)? You should be able to answer this question without further calculations. However, to double check, find the θ that maximizes the expression you have found in part (iv).

PROBLEM 3. (Enhanced Weather Frog)

A TV weather frog bases his weather forecast for tomorrow entirely on today's air pressure, which is thus his observable Y . Here, we consider an ambitious weather frog who wants to distinguish three kinds of weather. This means, that tomorrow's weather is represented by a random variable H which take on value 0 if the sun shines tomorrow, 1 if it rains or 2 if the weather is unstable. We assume that the three hypotheses are a priori equally likely, i.e. $p_H(0) = p_H(1) = p_H(2) = 1/3$.

Measurements over several years have led to the following estimate of the probability density function of today's air pressure provided that the sun shines tomorrow,

$$f_{Y|H}(y|0) = \begin{cases} A - 2Ay & , 0 \leq y \leq 0.5 \\ 0 & , \text{otherwise.} \end{cases}$$

The estimate of the probability density function of today's air pressure provided that it rains tomorrow, is

$$f_{Y|H}(y|1) = \begin{cases} B + \frac{B}{2}y & , 0 \leq y \leq 1 \\ 0 & , \text{otherwise.} \end{cases}$$

Finally, the estimate of the probability density function of today's air pressure provided that the weather is unstable tomorrow, is

$$f_{Y|H}(y|2) = \begin{cases} C & , 0 \leq y \leq 1 \\ 0 & , \text{otherwise.} \end{cases}$$

The weather frog's goal is to guess the value of H after measuring Y .

(i) Determine A , B and C .

(ii) Write down the optimal decision rule (i.e. the rule that minimize the probability of a wrong forecast) in general terms.

(iii) For all values y , draw into one graph $f_{y|H}(y|0)$, $f_{y|H}(y|1)$ and $f_{y|H}(y|2)$. Show on the graph the decision regions corresponding to the optimal decision rule. If we let $\hat{H}(y)$

denote the frog's forecast for a value y of the measurement, can the decision rule be written in the following form:

$$\hat{H}(y) \begin{cases} 0 & , \text{ if } y \leq \theta_1 \\ 2 & , \text{ if } \theta_1 < y < \theta_2 \\ 1 & , \text{ if } y \geq \theta_2, \end{cases}$$

where θ_1 and θ_2 are some thresholds? If so, determine the values θ_1 and θ_2 ?

(iv) Find the probability of a wrong forecast knowing that tomorrow's weather is unstable, i.e., determine the probability that the decision \hat{H} is different from 2 knowing that, in reality, $H = 2$. This probability is denoted $\Pr\{e|H = 2\}$.

(v) If we assume that, instead of using the optimal rule, our weather frog always decides that tomorrow's weather is sunny, what will be his probability of error (probability of a wrong forecast)? Explain.

PROBLEM 4. (Hypothesis Testing in Laplacian Noise) Consider the following hypothesis testing problem between two equally likely hypotheses. Under hypothesis $H = 0$, the observable Y is equal to $a + Z$ where Z is a random variable with Laplacian distribution

$$f_Z(z) = \frac{1}{2}e^{-|z|}. \quad (2.14)$$

Under hypothesis $H = 1$, the observable is given by $-a + Z$.

(i) Find and draw the density $f_{Y|H}(y|0)$ of the observable under hypothesis $H = 0$, and the density $f_{Y|H}(y|1)$ of the observable under hypothesis $H = 1$.

(ii) Find the optimal decision rule to minimize the probability of error. Write out the expression for the likelihood ratio.

(iii) Compute the probability of error of the optimal decision rule.

PROBLEM 5. (Poisson Parameter Estimation) In this example there are two hypotheses, $H = 0$ and $H = 1$ which occur with probabilities $p_H(0) = p_0$ and $p_H(1) = 1 - p_0$, respectively. The observable is $y \in \mathbb{N}_0$, i.e. y is a nonnegative integer. Under hypothesis $H = 0$, y is distributed according to a Poisson law with parameter λ_0 , i.e.

$$p_{Y|H}(y|0) = \frac{\lambda_0^y}{y!}e^{-\lambda_0}. \quad (2.15)$$

Under hypothesis $H = 1$,

$$p_{Y|H}(y|1) = \frac{\lambda_1^y}{y!}e^{-\lambda_1}. \quad (2.16)$$

This example is in fact modeling the reception of photons in an optical fiber (for more details, see the Example in Section 2.2 of these notes).

(i) Derive the MAP decision rule by indicating likelihood and log-likelihood ratios.

Hint: The direction of an inequality changes if both sides are multiplied by a negative number.

(ii) Derive the formula for the probability of error of the MAP decision rule.

(iii) For $p_0 = 1/3$, $\lambda_0 = 2$ and $\lambda_1 = 10$, compute the probability of error of the MAP decision rule. You may want to use a computer program to do this.

(iv) Repeat (iv) with $\lambda_1 = 20$ and comment.

PROBLEM 6. (IID versus First-Order Markov Model) Consider testing two equally likely hypotheses $H = 0$ and $H = 1$. The observable

$$Y = (Y_1, \dots, Y_k) \quad (2.17)$$

is a k -dimensional binary vector. Under $H = 0$ the components of the vector Y are independent uniform random variables (also called Bernoulli(1/2) random variables). Under $H = 1$, the component Y_1 is also uniform, but the components Y_i , $2 \leq i \leq k$, are distributed as follows:

$$Pr(Y_i = y_i | Y_{i-1} = y_{i-1}, \dots, Y_1 = y_1) = \begin{cases} 3/4, & \text{if } y_i = y_{i-1} \\ 1/4, & \text{otherwise.} \end{cases} \quad (2.18)$$

(i) Find the decision rule that minimizes the probability of error. Hint: Write down a short sample sequence (y_1, \dots, y_k) and determine its probability under each hypothesis. Then generalize.

(ii) Give a simple sufficient statistic for this decision.

(iii) Suppose that the observed sequence alternates between 0 and 1 except for one string of ones of length s , i.e. the observed sequence y looks something like

$$y = 0101010111111 \dots 111111010101 \dots \quad (2.19)$$

What is the least s such that we decide for hypothesis $H = 1$? Evaluate your formula for $k = 20$.

PROBLEM 7. (Real-Valued Gaussian Random Variables) For the purpose of this problem, two zero-mean real-valued Gaussian random variables X and Y are called jointly Gaussian if and only if their joint density is

$$f_{XY}(x, y) = \frac{1}{2\pi\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2} \begin{pmatrix} x & y \end{pmatrix} \Sigma^{-1} \begin{pmatrix} x \\ y \end{pmatrix}\right), \quad (2.20)$$

where (for zero-mean random vectors) the so-called covariance matrix Σ is

$$\Sigma = E \left[\begin{pmatrix} X \\ Y \end{pmatrix} (X, Y) \right] = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}. \quad (2.21)$$

- (i) Show that if X and Y are jointly Gaussian random variables, then X is a Gaussian random variable, and so is Y .
- (ii) How does your answer change if you use the definition of jointly Gaussian random variables given in these notes?
- (iii) Show that if X and Y are independent Gaussian random variables, then X and Y are jointly Gaussian random variables.
- (iv) However, if X and Y are Gaussian random variables but not independent, then X and Y are not necessarily jointly Gaussian. Give an example where X and Y are Gaussian random variables, yet they are not jointly Gaussian.
- (v) Let X and Y be independent Gaussian random variables with zero mean and variance σ_X^2 and σ_Y^2 , respectively. Find the probability density function of $Z = X + Y$.

PROBLEM 8. (Correlation and Independence) Let Z be a random variable with p.d.f.:

$$f_Z(z) = \begin{cases} 1/2, & -1 \leq z \leq 1 \\ 0, & \text{otherwise.} \end{cases} \quad (2.22)$$

Also, let $X = Z$ and $Y = Z^2$.

- (i) Show that X and Y are uncorrelated.
- (ii) Are X and Y independent?
- (iii) Now let X and Y be jointly Gaussian, zero mean, uncorrelated with variances σ_X^2 and σ_Y^2 respectively. Are X and Y independent? Justify your answer.

PROBLEM 9. (Transformation of Random Vectors) Let R and Φ be independent random variables. R is distributed uniformly over the unit interval, Φ is distributed uniformly over the interval $[0, 2\pi)$.⁴

- (i) Interpret R and Φ as the polar coordinates of a point in the plane. It is clear that the point lies inside (or on) the unit circle. Is the distribution of the point uniform over the unit disk? Take a guess!

⁴This notation means: 0 is included, but 2π is excluded. It is the current standard notation in the anglo-saxon world. In the French world, the current standard for the same thing is $[0, 2\pi[$.

(ii) Define the random variables

$$X = R \cos \Phi \quad (2.23)$$

$$Y = R \sin \Phi. \quad (2.24)$$

Find the joint distribution of the random variables X and Y using the Jacobian determinant.

Do you recognize a relationship between this method and the method derived in class to determine the probability density after a linear non-singular transformation?

(iii) Does the result of part (ii) support or contradict your guess from part (i)? Explain.

PROBLEM 10. (Theorem Of Irrelevance and Sufficient Statistics) Have you ever tried to drink from a fire hydrant? There are situations in which the observable Y contains too much data. You would like to have a many-to-one function T so that $T(Y)$ contains enough information to make a MAP decision but not too much to be impractical to work with. The Theorem of irrelevance gives a test to check if you have such a function.

Consider two hypotheses with probabilities $p_H(0) = p_0$ and $p_H(1) = 1 - p_0$. The observable is $Y = (Y_1, \dots, Y_k)$. Let $f_{Y|H}(y|0)$ and $f_{Y|H}(y|1)$ be given.

(i) (Theorem of irrelevance): Suppose it is possible to write

$$f_{Y|H}(y|0) = g_0(T(y))h(y) \quad (2.25)$$

$$f_{Y|H}(y|1) = g_1(T(y))h(y), \quad (2.26)$$

where $T(\cdot) : \mathbb{R}^k \rightarrow \mathbb{R}^d$ is a function from the observation space \mathbb{R}^k to some space of choice \mathbb{R}^d , $g_0(\cdot), g_1(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^+$ and $h(\cdot) : \mathbb{R}^k \rightarrow \mathbb{R}^+$.

Prove that if you have $T(y)$ you don't need y to make a MAP decision. For this reason $T(y)$ is called a sufficient statistic for the hypothesis testing problem.

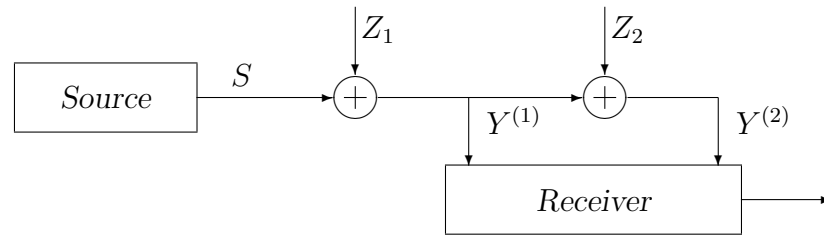
(ii) Sometimes we can partition the observable $Y \in \mathbb{R}^k$ into two vectors $Y' = (Y_1, \dots, Y_r)$ and $Y'' = (Y_{r+1}, \dots, Y_k)$. Show that the irrelevance theorem implies the following statement: If $f_{Y''|H, Y'}(y''|i, y')$ does not depend on i , then Y' is a sufficient statistic, i.e. Y'' is irrelevant to the decision problem.

(iii) Use (ii) to answer the following communications problem (see the picture below): Under $H = 0$, the source emits $S = 1$; under $H = 1$, the source emits $S = -1$. The receiver has access to two noisy versions of the source output, namely

$$Y^{(1)} = S + Z_1 \quad (2.27)$$

$$Y^{(2)} = S + Z_1 + Z_2, \quad (2.28)$$

where Z_1 and Z_2 are zero-mean Gaussian random variables of variance σ^2 . Is $Y^{(2)}$ relevant to the hypothesis testing problem? Prove your answer.



PROBLEM 11. (Sufficient Statistic) Consider a binary hypothesis testing problem specified by:

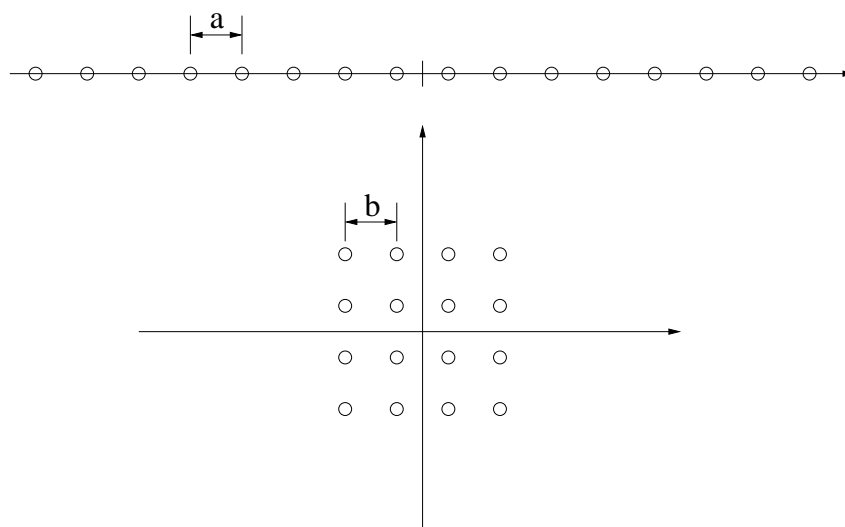
$$\begin{aligned}
 H = 0 & : \begin{cases} Y_1 = Z_1 \\ Y_2 = Z_1 Z_2 \end{cases} \\
 H = 1 & : \begin{cases} Y_1 = -Z_1 \\ Y_2 = -Z_1 Z_2 \end{cases}
 \end{aligned}$$

where Z_1 , Z_2 and H are independent random variables.

(i) Is Y_1 a sufficient statistic? Recall that Y_1 is a sufficient statistic if a MAP decoder that observes (Y_1, Y_2) makes the same decision (up to ties) as a MAP decoder that observes Y_1 alone.

(Hint: If $Y = aZ$, where a is a scalar then $f_Y(y) = \frac{1}{|a|} f_Z(\frac{y}{a})$).

PROBLEM 12. (Comparison of 16-PAM and 16-QAM) The following two signal constellations are used to communicate across an additive white Gaussian noise channel. Let the noise variance be σ^2 .



Each point represents a signal \mathbf{s}_i for some i . Assume each signal is used with the same probability.

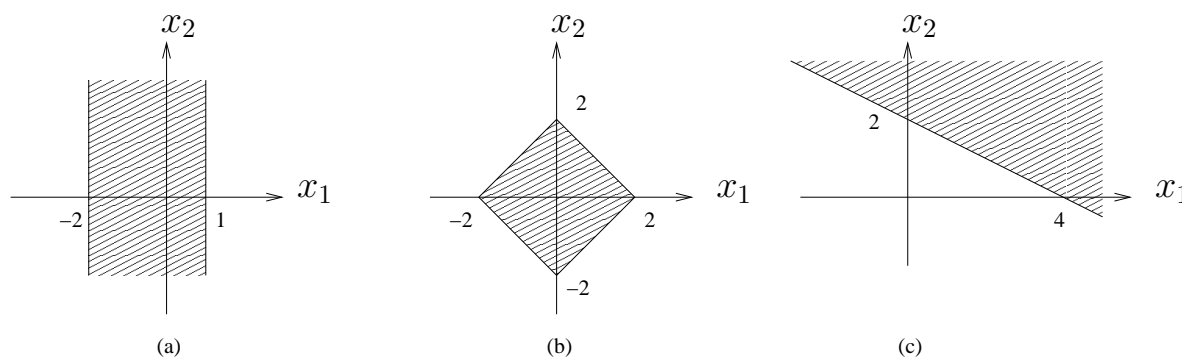
(i) For each signal constellation, compute the average probability of error, P_e , as a function of the parameters a and b , respectively.

(ii) For each signal constellation, compute the average energy per symbol, E_s , as a function of the parameters a and b , respectively:

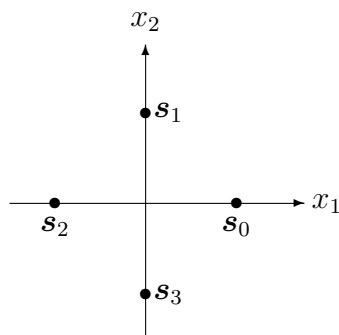
$$E_s = \sum_{i=1}^{16} p_H(i) \|\mathbf{s}_i\|^2 \tag{2.29}$$

(iii) Plot P_e versus E_s for both signal constellations and comment.

PROBLEM 13. (A Gaussian Vector and Three Regions) [Wozencraft and Jacobs] Let $\mathbf{X} \sim \mathcal{N}(0, \sigma^2 I_2)$. For each of the three figures below, express the probability that \mathbf{X} lies in the shaded region. You may use the Q -function when appropriate.



PROBLEM 14. (QPSK Decision Regions) Let $H \in \{0, 1, 2, 3\}$ and assume that when $H = i$ you transmit the signal \mathbf{s}_i shown in the figure. Under $H = i$, the receiver observes $\mathbf{Y} = \mathbf{s}_i + \mathbf{Z}$.



- (a) Draw the decoding regions assuming that $\mathbf{Z} \sim \mathcal{N}(0, \sigma^2 I_2)$ and that $P_H(i) = 1/4$, $i \in \{0, 1, 2, 3\}$.
- (b) Draw the decoding regions (qualitatively) assuming $\mathbf{Z} \sim \mathcal{N}(0, \sigma^2 I)$ and $P_H(0) = P_H(2) > P_H(1) = P_H(3)$. Justify your answer.
- (c) Assume again that $P_H(i) = 1/4$, $i \in \{0, 1, 2, 3\}$ and that $\mathbf{Z} \sim \mathcal{N}(0, K)$, where $K = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 4\sigma^2 \end{pmatrix}$. How do you decode now? Justify your answer.

PROBLEM 15. (Antenna Array) *The following problem relates to the design of multi-antenna systems. The situation that we have in mind is one where one of two signals is transmitted over a Gaussian channel and is received through two different antennas. We shall assume that the noises at the two terminals are independent but not necessarily of equal variance. You are asked to design a receiver for this situation, and to assess its performance. This situation is made more precise as follows:*

Consider the binary equiprobable hypothesis testing problem:

$$\begin{aligned} H = 0 & : Y_1 = A + Z_1, & Y_2 = A + Z_2 \\ H = 1 & : Y_1 = -A + Z_1, & Y_2 = -A + Z_2, \end{aligned}$$

where Z_1, Z_2 are independent Gaussian random variables with different variances $\sigma_1^2 \neq \sigma_2^2$, that is, $Z_1 \sim \mathcal{N}(0, \sigma_1^2)$ and $Z_2 \sim \mathcal{N}(0, \sigma_2^2)$. $A > 0$ is a constant.

- (a) Show that the decision rule that minimizes the probability of error (based on the observable Y_1 and Y_2) can be stated as

$$\sigma_2^2 y_1 + \sigma_1^2 y_2 \underset{1}{\overset{0}{\geq}} 0. \quad (2.30)$$

- (b) Draw the decision regions in the (Y_1, Y_2) plane for the special case where $\sigma_1 = 2\sigma_2$.
- (c) Evaluate the probability of error for the optimal detector as a function of σ_1^2 , σ_2^2 and A .

PROBLEM 16. (Multiple Choice Exam) *You are taking a multiple choice exam. Question number 5 allows for two possible answers. According to your first impression, answer 1 is correct with probability 1/4 and answer 2 is correct with probability 3/4.*

You would like to maximize your chance of giving the correct answer and you decide to have a look at what your left and right neighbors have to say.

The left neighbor has answered $\hat{H}_L = 1$. He is an excellent student who has a record of being correct 90% of the time.

The right neighbor has answered $\hat{H}_R = 2$. He is a weaker student who is correct 70% of the time.

- (a) You decide to use your first impression as a prior and to consider \hat{H}_L and \hat{H}_R as observations. Describe the corresponding hypothesis testing problem.
- (b) What is your answer \hat{H} ? Justify it.

PROBLEM 17. (QAM with an Erasure) Consider a QAM receiver that outputs a special symbol called “erasure” and denoted by δ whenever the observation falls in the shaded area shown in Figure (2.8). Assume that \mathbf{s}_0 is transmitted and that $\mathbf{Y} = \mathbf{s}_0 + \mathbf{N}$ is received where $\mathbf{N} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_2)$. Let P_{0i} , $i = 0, 1, 2, 3$ be the probability that the receiver outputs $\hat{H} = i$ and let $P_{0\delta}$ be the probability that it outputs δ . Determine P_{00} , P_{01} , P_{02} , P_{03} and $P_{0\delta}$.

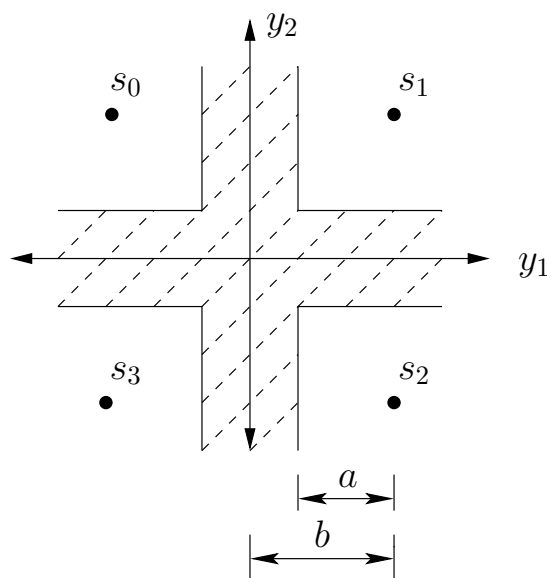


Figure 2.8: Modified QAM demodulator

PROBLEM 18. (Repeat Codes and Bhattacharyya Bound) A repeat code is a code that transmits each source output N times across the channel. It is clear that the probability of error at the decoder decreases with increasing N .

Consider two equally likely hypotheses (or, as we could also say, source output values). Under hypothesis $H = 0$, the signal $(X_1, \dots, X_N) = (1, \dots, 1)$ is put onto the channel; under hypothesis $H = 1$, the signal is $(X_1, \dots, X_N) = (-1, \dots, -1)$. The transmission channel adds zero-mean independent Gaussian noise of variance σ^2 . At the receiver, we observe

$$(Y_1, \dots, Y_N) = (X_1 + Z_1, \dots, X_N + Z_N). \quad (2.31)$$

Based on this observation, we can find the MAP estimator. In fact, it turns out that a sufficient statistic is the sum of the received values, $Y_1 + Y_2 + \dots + Y_N$. The corresponding probability of error was found to be

$$Pr^{(1)}\{e\} = Q\left(\frac{\sqrt{N}}{\sigma}\right). \quad (2.32)$$

However, in this case, the receiver has to be able to perform addition of real numbers, and we also have to store them. This is not always possible. Therefore, suppose now that the decoder has access only to the sign of Y_i , $1 \leq i \leq N$. That is, the observation is

$$W = (W_1, \dots, W_N) = (\text{sgn}(Y_1), \dots, \text{sgn}(Y_N)) = (\text{sgn}(X_1 + Z_1), \dots, \text{sgn}(X_N + Z_N))$$

where $Z_i \sim \mathcal{N}(0, \sigma^2)$.

(i) Determine the MAP decision rule based on the observation (W_1, \dots, W_N) . Give a simple sufficient statistic, and draw a diagram of the optimal receiver.

(ii) Find the expression for the probability of error $Pr^{(2)}\{e\}$. You may assume that N is odd.

(iii) Your answer to (ii) contains a sum that cannot be solved in closed form. Therefore, find the Bhattacharyya bound on $Pr^{(2)}\{e\}$.

(iv) For $N = 1, 3, 5, 7$, find the numerical values of $Pr^{(1)}\{e\}$, $Pr^{(2)}\{e\}$, and the Bhattacharyya bound on $Pr^{(2)}\{e\}$.

PROBLEM 19. (Tighter Union Bhattacharyya Bound: Binary Case) In this problem we derive a tighter version of the Union Bhattacharyya Bound for binary hypotheses.

Let

$$\begin{aligned} H = 0 & : Y \sim f_{Y|H}(y | 0) \\ H = 1 & : Y \sim f_{Y|H}(y | 1). \end{aligned}$$

The MAP decision rule is

$$\hat{H}(y) = \arg \max_i P_H(i) f_{Y|H}(y | i),$$

and the resulting probability of error is

$$Pr\{e\} = P_H(0) \int_{\mathcal{R}_1} f_{Y|H}(y | 0) dy + P_H(1) \int_{\mathcal{R}_0} f_{Y|H}(y | 1) dy. \quad (2.34)$$

(i) Argue that

$$Pr\{e\} = \int_y \min \{P_H(0) f_{Y|H}(y | 0), P_H(1) f_{Y|H}(y | 1)\} dy.$$

(ii) Prove that for $a, b \geq 0$, $\min(a, b) \leq \sqrt{ab} \leq \frac{a+b}{2}$. Use this to prove the tighter version of Bhattacharyya Bound, i.e.,

$$Pr\{e\} \leq \frac{1}{2} \int_y \sqrt{f_{Y|H}(y|0)f_{Y|H}(y|1)} dy.$$

(iii) Compare the above bound to the one derived in class when $P_H(0) = \frac{1}{2}$. How do you explain the improvement by a factor $\frac{1}{2}$?

PROBLEM 20. (Tighter Union Bhattacharyya Bound: M -ary Case)

In class we have derived the Union Bhattacharyya Bound. Is this a tight bound or can we do better? To be specific, let us analyze the following M -ary MAP detector:

$$\hat{H}(y) = \text{smallest } i \text{ such that} \quad (2.35)$$

$$P_H(i)f_{Y|H}(y|i) = \max_j \{P_H(j)f_{Y|H}(y/j)\} \quad (2.36)$$

Let

$$\mathcal{B}_{ij} = \begin{cases} y : P_H(j)f_{Y|H}(y|j) \geq P_H(i)f_{Y|H}(y|i), & j < i \\ y : P_H(j)f_{Y|H}(y|j) > P_H(i)f_{Y|H}(y|i), & j > i \end{cases} \quad (2.37)$$

(i) Verify that $\mathcal{B}_{ij} = \mathcal{B}_{ji}^c$.

Given $H = i$, the detector will make an error iff:

$$y \in \bigcup_{j:j \neq i} \mathcal{B}_{ij} \quad (2.38)$$

We calculate the probability of error as:

$$Pr\{e\} = \sum_{i=0}^{M-1} Pr\{e | H = i\}P_H(i) \quad (2.39)$$

(ii) Show that:

$$Pr\{e\} \leq \sum_{i=0}^{M-1} \sum_{j>i} [Pr\{\mathcal{B}_{ij} | H = i\}P_H(i) + Pr\{\mathcal{B}_{ji} | H = j\}P_H(j)] \quad (2.40)$$

$$= \sum_{i=0}^{M-1} \sum_{j>i} \left[\int_{\mathcal{B}_{ij}} f_{Y|H}(y|i)P_H(i)dy + \int_{\mathcal{B}_{ji}^c} f_{Y|H}(y|j)P_H(j)dy \right] \quad (2.41)$$

$$= \sum_{i=0}^{M-1} \sum_{j>i} \left[\int_y \min \{f_{Y|H}(y|i)P_H(i), f_{Y|H}(y|j)P_H(j)\} dy \right] \quad (2.42)$$

(Hint: Apply the Union of Events Bound to equation (2.39) and then group the terms corresponding to \mathcal{B}_{ij} and \mathcal{B}_{ji} . For proving the last part, go back to the definition of \mathcal{B}_{ij} .)

(iii) Hence show that:

$$\Pr\{e\} \leq \sum_{i=0}^{M-1} \sum_{j>i} \left[\left(\frac{P_H(i) + P_H(j)}{2} \right) \int_y \sqrt{f_{Y|H}(y|i)f_{Y|H}(y|j)} dy \right] \quad (2.43)$$

(Hint: For $a, b \geq 0$, $\min(a, b) \leq \sqrt{ab} \leq \frac{a+b}{2}$.)

As an application of the above bound, consider the following binary hypothesis testing problem:

$$H = 0 : Y \sim \mathcal{N}(-a, \sigma^2) \quad (2.44)$$

$$H = 1 : Y \sim \mathcal{N}(+a, \sigma^2) \quad (2.45)$$

where the two hypotheses are equiprobable. Use the above bound to show that:

$$\Pr\{e\} = \Pr\{e | H = 0\} \quad (2.46)$$

$$\leq \frac{1}{2} \exp \left\{ -\frac{a^2}{2\sigma^2} \right\} \quad (2.47)$$

But $\Pr\{e\} = Q\left(\frac{a}{\sigma}\right)$. Hence we have re-derived the bound (see lecture 1):

$$Q(x) \leq \frac{1}{2} \exp \left\{ -\frac{x^2}{2} \right\}. \quad (2.48)$$

PROBLEM 21. (An Application to the Tight Bhattacharyya Bound) As an application of the tight Bhattacharyya bound, consider the following binary hypothesis testing problem

$$H = 0 : Y \sim \mathcal{N}(-a, \sigma^2)$$

$$H = 1 : Y \sim \mathcal{N}(+a, \sigma^2)$$

where the two hypotheses are equiprobable.

(i) Use the Tight Bhattacharyya Bound to derive a bound on $\Pr\{e\}$.

(ii) We know that the probability of error for this binary hypothesis testing problem is $Q\left(\frac{a}{\sigma}\right) \leq \frac{1}{2} \exp \left\{ -\frac{a^2}{2\sigma^2} \right\}$, where we have used the result $Q(x) \leq \frac{1}{2} \exp \left\{ -\frac{x^2}{2} \right\}$ derived in lecture 1. How do the two bounds compare? Are you surprised (and why)?

PROBLEM 22. (Bhattacharyya Bound for DMCs) Consider a Discrete Memoryless Channel (DMC). This is a channel model described by an input alphabet \mathcal{X} , an output alphabet \mathcal{Y} and a transition probability⁵ $P(y|x)$. When we use this channel to transmit an n -tuple $\mathbf{x} \in \mathcal{X}^n$, the transition probability is

$$P(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^n P(y_i|x_i).$$

So far we have come across two DMCs, namely the BSC (Binary Symmetric Channel) and the BEC (Binary Erasure Channel). The purpose of this problem is to realize that for DMCs, the Bhattacharyya Bound takes on a simple form, in particular when the channel input alphabet \mathcal{X} contains only two letters.

(i) Consider a source that sends \mathbf{s}_0 when $H = 0$ and \mathbf{s}_1 when $H = 1$. Justify the following chain of inequalities.

$$\begin{aligned} Pr\{e\} &\stackrel{(a)}{\leq} \frac{1}{2} \sum_{\mathbf{y}} \sqrt{P(\mathbf{y}|\mathbf{s}_0)P(\mathbf{y}|\mathbf{s}_1)} \\ &\stackrel{(b)}{\leq} \sum_{\mathbf{y}} \sqrt{\prod_{i=1}^n P(y_i|s_{0i})P(y_i|s_{1i})} \\ &\stackrel{(c)}{=} \sum_{y_1, \dots, y_n} \prod_{i=1}^n \sqrt{P(y_i|s_{0i})P(y_i|s_{1i})} \\ &\stackrel{(d)}{=} \left[\sum_{y_1} \sqrt{P(y_1|s_{01})P(y_1|s_{11})} \right] \cdots \left[\sum_{y_n} \sqrt{P(y_n|s_{0n})P(y_n|s_{1n})} \right] \\ &\stackrel{(e)}{=} \prod_{i=1}^n \sum_y \sqrt{P(y|s_{0i})P(y|s_{1i})} \\ &\stackrel{(f)}{=} \prod_{a \in \mathcal{X}, b \in \mathcal{X}, a \neq b} \left(\sum_y \sqrt{P(y|s_{0i})P(y|s_{1i})} \right)^{n(a,b)}. \end{aligned}$$

where $n(a, b)$ is the number of positions i in which $s_{0i} = a$ and $s_{1i} = b$.

(ii) The Hamming distance $d_H(\mathbf{s}_0, \mathbf{s}_1)$ is defined as the number of positions in which \mathbf{s}_0 and \mathbf{s}_1 differ. Show that for a binary input channel, i.e., when $\mathcal{X} = \{a, b\}$, the Bhattacharyya Bound becomes

$$Pr\{e\} \leq z^{d_H(\mathbf{s}_0, \mathbf{s}_1)},$$

where

$$z = \sum_y \sqrt{P(y|a)P(y|b)}.$$

⁵Here we are assuming that the output alphabet is discrete. Otherwise we need to deal with densities instead of probabilities.

Notice that z depends only on the channel whereas its exponent depends only on s_0 and s_1 .

(iii) What is z for:

(a) The binary input Gaussian channel described by the densities

$$\begin{aligned} f_{Y|X}(y|0) &= \mathcal{N}(-\sqrt{E}, \sigma^2) \\ f_{Y|X}(y|1) &= \mathcal{N}(\sqrt{E}, \sigma^2). \end{aligned}$$

(b) The Binary Symmetric Channel (BSC) with the transition probabilities described by

$$p_{Y|X}(y|x) = \begin{cases} 1 - \delta, & \text{if } y = x, \\ \delta, & \text{otherwise.} \end{cases}$$

Verify your result with that of homework 4 problem 1.

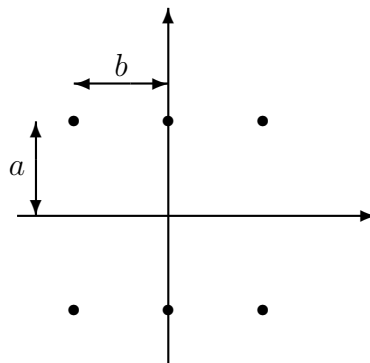
(c) The Binary Erasure Channel (BEC) with the transition probabilities given by

$$p_{Y|X}(y|x) = \begin{cases} 1 - \delta, & \text{if } y = x, \\ \delta, & \text{if } y = E \\ 0, & \text{otherwise.} \end{cases}$$

Verify your result with the one obtained in class.

(iv) Extra question for the curious ones: Assume that the BSC has been obtained from the binary-input Gaussian channel via a one-bit quantizer applied at the channel output like in homework 4 problem (i). Plot the z of the original and the quantized channel as a function of the input power. By how much do we need to increase the input power of the quantized channel to match the z of the unquantized channel?

PROBLEM 23. (Signal Constellation) The following signal constellation with six signals is used in additive white Gaussian noise of variance σ^2 :



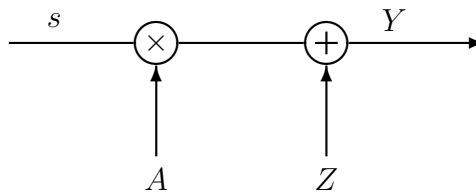
Assume that the six signals are used with equal probabilities.

- (i) Draw the boundaries of the decision regions into the above figure.
- (ii) Compute the average probability of error, $Pr\{e\}$, for this signal constellation.
- (iii) Compute the average energy per symbol for this signal constellation.

PROBLEM 24. (Application of Hypothesis Testing to Fading) Consider the following communication problem:

There are two equiprobable hypotheses. When $H = 0$, we transmit $s = -b$, where b is an arbitrary but fixed positive number. When $H = 1$, we transmit $s = b$.

The channel is as shown in the figure below, where $Z \sim \mathcal{N}(0, \sigma^2)$ represents the noise, $A \in \{0, 1\}$ represents a random attenuation (fading) with $P_A(0) = \frac{1}{2}$, and Y is the channel output. The random variables H , A and Z are independent.



- (i) Find the decision rule that the receiver should implement to minimize the probability of error. Sketch the decision regions.
- (ii) Calculate the probability of error $Pr\{e\}$, based on the above decision rule.

PROBLEM 25. (Dice Tossing)

You have two dices, one fair and one loaded (truqué). A friend told you that the loaded dice produces a 6 with probability $\frac{1}{4}$, and the other values with uniform probabilities. You do not know a priori which one is fair or which one is loaded. You pick with uniform probabilities one of the two dices, and perform N consecutive tosses (lancés) with the dice you have chosen. Let

$$Y = (Y_1, \dots, Y_N)$$

be the sequence of numbers observed.

- (a) Based on the sequence of observations Y , find the decision rule to determine whether the dice you have chosen is loaded. Your decision rule should maximize the probability of correct decision.

- (b) Identify a compact sufficient statistic for this hypothesis testing problem, call it S . Justify your answer. [Hint: $S \in \mathbb{N}$.]
- (c) Find the Bhattacharyya bound on the probability of error. You can either work with the observation (Y_1, \dots, Y_N) or with (Z_1, \dots, Z_N) , where Z_i indicates whether the i th observation is a six or not, or you can work with S . In some cases you may find it useful to know that $\sum_{i=0}^N \binom{N}{i} x^i = (1+x)^N$ for $N \in \mathbb{N}$. In other cases the following may be useful: $\sum_{Y_1, Y_2, \dots, Y_N} \prod_{i=1}^N f(Y_i) = \left(\sum_{Y_1} f(Y_1) \right)^N$.

PROBLEM 26. (Who Wants to Be a Millionaire)

Assume you are at a quiz show. You are shown three boxes which look identical from the outside, except they have labels 0, 1, and 2, respectively. Exactly one of them contains one million Swiss francs, the other two contain nothing. A computer randomly chooses a box with uniform probability. Let A be the random variable which denotes his choice, $A \in \{0, 1, 2\}$. The quizmaster now eliminates from the remaining two boxes one that does not contain the prize. This means that if neither of the two remaining boxes contain the prize then the quizmaster eliminates one with uniform probability. Otherwise, he simply eliminates the one which does not contain the prize. Let B denote the random variable corresponding to the box eliminated by the quizmaster, $B \in \{0, 1, 2\}$, and let C denote the remaining box. You are asked to choose one of the three boxes knowing A and B .

- (a) Formulate this as a hypotheses testing problem. What is the set of hypotheses, what are the observations, and what are the priors?
- (b) Write down the general rule for the optimal decision. Assume that $A = 0$ and $B = 1$. What is the optimal decision?
- (c) What is the optimal decision in the general case?

PROBLEM 27. (Playing Darts)

Assume that you are throwing darts at a target. We assume that the target is one-dimensional, i.e., that the darts all end up on a line. The “bulls eye” is in the center of the line, and we give it the coordinate 0. The position of a dart on the target can then be measured with respect to 0.

We assume that the position X_1 of a dart that lands on the target is a random variable that has a Gaussian distribution with variance σ_1^2 and mean 0.

Assume now that there is a second target, which is further away. If you throw dart to that target, the position X_2 has a Gaussian distribution with variance σ_2^2 (where $\sigma_2^2 > \sigma_1^2$) and mean 0.

You play the following game: You toss a coin which gives you “head” with probability p and “tail” with probability $1 - p$ for some fixed $p \in [0, 1]$. We can model a coin as a Bernoulli random variable Z . If $Z = 1$, you throw a dart onto the first target. If $Z = 0$, you aim the second target instead. Let X be the relative position of the dart with respect to the center of the target that you have chosen.

- (i) Write down X in terms of X_1 , X_2 and Z .
- (ii) Compute the variance of X .

Bonus question: Is the distribution of X a Gaussian (Note that X is not a linear combination of X_1 and X_2)? Explain.

- (iii) Let $S = |X|$ be the score, which is given by the distance of the dart to the center of the target (that you picked using the coin). Compute the average score $\mathbb{E}[S]$.

PROBLEM 28. (Properties of the Q Function)

Prove properties (a) through (d) of the Q function defined in the lecture notes, Section 2.3.

You can use the following hint to help you with property (d):

Hint: Define $\Phi(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$. Then, integrate $\int_x^\infty \Phi(t)\frac{1}{t^2}dt$ by parts.

The resulting equality can be used to prove both inequalities in (d). Start by proving the right-hand inequality.

PROBLEM 29. (Uncorrelation vs. Independence)

Let X and Y be two random variables.

- (i) When are X and Y uncorrelated? When are they independent? Write down the definitions.
- (ii) Show that if X and Y are independent, they are also uncorrelated.
- (iii) We first define two new random variables U and V :

$$U = \begin{cases} 0 & \text{with prob. } \frac{1}{2} \\ 1 & \text{with prob. } \frac{1}{2}, \end{cases}$$

$$V = \begin{cases} 0 & \text{with prob. } \frac{1}{2} \\ 1 & \text{with prob. } \frac{1}{2}, \end{cases}$$

where U and V are independent. Now, assume that X and Y are defined as follows: $X = U + V$ and $Y = |U - V|$. Are X and Y independent? Compute the covariance of X and Y . What do you conclude?

PROBLEM 30. (Bhattacharyya Bound and Laplacian Noise) When $Y \in \mathbb{R}$ is a continuous random variable, the Bhattacharyya bound states that

$$Pr\{Y \in \mathcal{B}_{i,j} | H = i\} \leq \sqrt{\frac{P_H(j)}{P_H(i)}} \int_{y \in \mathbb{R}} \sqrt{f_{Y|H}(y|i) f_{Y|H}(y|j)} dy,$$

where i, j are two possible hypotheses and $\mathcal{B}_{i,j} = \{y \in \mathbb{R} : P_H(i) f_{Y|H}(y|i) \leq P_H(j) f_{Y|H}(y|j)\}$. In this problem $\mathcal{H} = \{0, 1\}$ and $P_H(0) = P_H(1) = 0.5$.

- Write a sentence that expresses the meaning of $Pr\{Y \in \mathcal{B}_{0,1} | H = 0\}$. Use words that have operational meaning.
- Do the same but for $Pr\{Y \in \mathcal{B}_{0,1} | H = 1\}$. (Note that we have written $\mathcal{B}_{0,1}$ and not $\mathcal{B}_{1,0}$.)
- Evaluate the right hand side of the Bhattacharyya bound for the special case $f_{Y|H}(y|0) = f_{Y|H}(y|1)$.
- Evaluate the Bhattacharyya bound for the following (Laplacian noise) setting:

$$\begin{aligned} H = 0 : & \quad Y = -a + Z \\ H = 1 : & \quad Y = a + Z, \end{aligned}$$

where $a \in \mathbb{R}_+$ is a constant and $f_Z(z) = \frac{1}{2} \exp(-|z|)$, $z \in \mathbb{R}$. Hint: it does not matter if you evaluate the bound for $H = 0$ or $H = 1$.

- For which value of a should the bound give the result obtained in (c)? Verify that it does. Check your previous calculations if it does not.

PROBLEM 31. (Irrelevance and the Markov Chain) Assume that H is a random variable that corresponds to a hypothesis. Let Y_s and Y_i be two more random variables, that are observations.

Definition: We say that $H \rightarrow Y_s \rightarrow Y_i$ forms a Markov chain if

$$f_{Y_i | H, Y_s}(y_i | i, y_s) = f_{Y_i | Y_s}(y_i | y_s)$$

for all possible values of y_i , i and y_s .

- (a) Show that $H \rightarrow Y_s \rightarrow Y_i$ forms a Markov chain if and only if $Y_i \rightarrow Y_s \rightarrow H$ forms a Markov chain, i.e., if and only if

$$P_{H|Y_s, Y_i}(i|y_s, y_i) = P_{H|Y_s}(i|y_s)$$

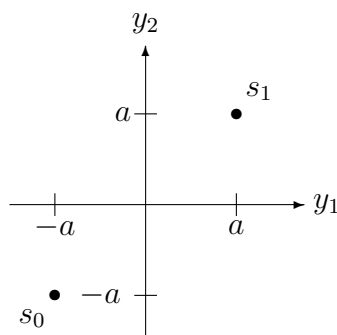
holds for all values of i , y_s and y_i .

- (b) We know from the homework that if $f_{Y_i|H, Y_s}(y_i|i, y_s)$ does not depend on i (in other words, if $H \rightarrow Y_s \rightarrow Y_i$ forms a Markov chain), then Y_s is a sufficient statistic and Y_i is irrelevant. Using part (a), this tells us that if $P_{H|Y_s, Y_i}(i|y_s, y_i) = P_{H|Y_s}(i|y_s)$ for all values of i , y_s and y_i , then Y_s is a sufficient statistic (and Y_i is irrelevant).

Is this intuitive? Explain why / why not.

PROBLEM 32. (Antipodal Signaling)

Consider the following signal constellation:



Assume that \mathbf{s}_1 and \mathbf{s}_0 are used for communication over the Gaussian vector channel. More precisely:

$$\begin{aligned} H = 0 : \mathbf{Y} &= \mathbf{s}_0 + \mathbf{Z}, \\ H = 1 : \mathbf{Y} &= \mathbf{s}_1 + \mathbf{Z}, \end{aligned}$$

where $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_2)$. Hence, \mathbf{Y} is a vector with two components $\mathbf{Y} = (Y_1, Y_2)$.

- (a) Argue why Y_1 is not a sufficient statistic.
- (b) Give a different signal constellation with two signals $\tilde{\mathbf{s}}_0$ and $\tilde{\mathbf{s}}_1$ such that, when using them in the above communication setting, Y_1 is a sufficient statistic.

PROBLEM 33. (Hypothesis Testing on a Uniform Vector Channel) Consider a binary hypothesis testing problem in which the hypotheses $H = 0$ and $H = 1$ occur with probability $P_H(0)$ and $P_H(1) = 1 - P_H(0)$, respectively. The observation \mathbf{Y} is a sequence of zeros and ones of length $2k$, where k is a fixed integer. When $H = 0$, each component of \mathbf{Y} is 0 or a 1 with probability $\frac{1}{2}$ and components are independent. When $H = 1$, \mathbf{Y} is chosen uniformly at random from the set of all sequences of length $2k$ that have an equal number of ones and zeros. There are $\binom{2k}{k}$ such sequences.

- What is $P_{\mathbf{Y}|H}(\mathbf{y}|0)$? What is $P_{\mathbf{Y}|H}(\mathbf{y}|1)$?
- Find a maximum likelihood decision rule. What is the single number you need to know about \mathbf{y} to implement this decision rule?
- Find a decision rule that minimizes the error probability.
- Are there values of $P_H(0)$ and $P_H(1)$ such that the decision rule that minimizes the error probability always decides for only one of the alternatives? If yes, what are these values, and what is the decision?

PROBLEM 34. (SIMO Channel with Laplacian Noise)

One of the two signals $s_0 = -1, s_1 = 1$ is transmitted over the channel shown on the left of Figure 2.9. The two noise random variables Z_1 and Z_2 are statistically independent of the transmitted signal and of each other.

Their density functions are

$$f_{Z_1}(\alpha) = f_{Z_2}(\alpha) = \frac{1}{2} e^{-|\alpha|}.$$

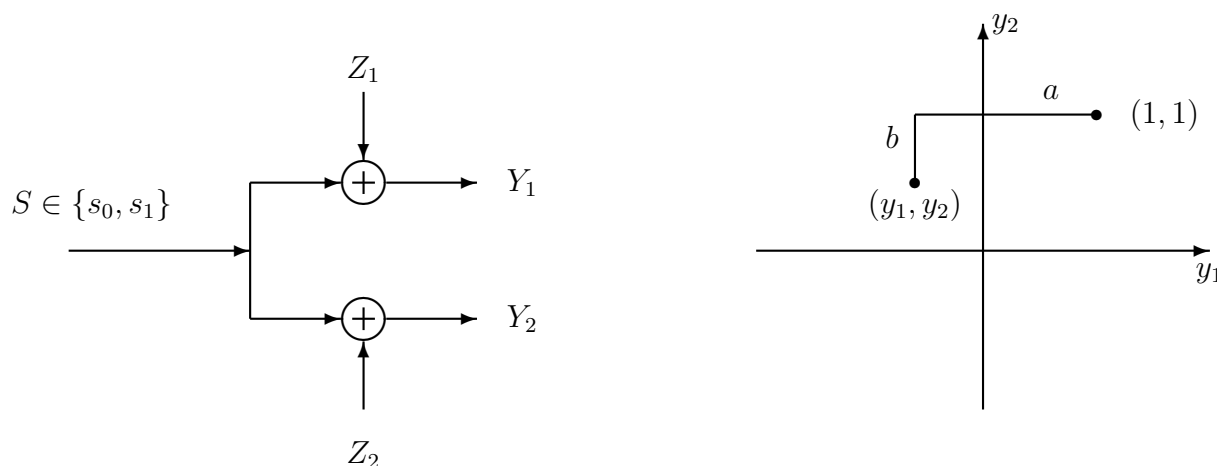


Figure 2.9: The channel (on the left) and a figure explaining the hint.

- (a) Derive a maximum likelihood decision rule.
- (b) Describe the maximum likelihood decision regions in the (y_1, y_2) plane. Try to describe the "Either Choice" regions, i.e., the regions in which it does not matter if you decide for s_0 or for s_1 . Hint: Use geometric reasoning and the fact that for a point (y_1, y_2) as shown on the right of Figure 1, $|y_1 - 1| + |y_2 - 1| = a + b$.
- (c) A receiver decides that s_1 was transmitted if and only if $(y_1 + y_2) > 0$. Does this receiver minimize the error probability for equally likely messages?
- (d) What is the error probability for the receiver in (c)? Hint: if $W = Z_1 + Z_2$ then $f_W(\omega) = \frac{e^{-\omega}}{4}(1 + \omega)$ for $w > 0$.
- (e) Could you have derived f_W as in (d)? If yes, say how but omit detailed calculations.

PROBLEM 35. (ML Receiver and UBB for Orthogonal Signaling)

Let $H \in \{1, \dots, m\}$ be uniformly distributed and consider the communication problem described by:

$$H = i : \quad \mathbf{Y} = s_i + Z, \quad Z \sim \mathcal{N}(0, \sigma^2 I_m),$$

where s_1, \dots, s_m , $s_i \in \mathbb{R}^m$, is a set of constant-energy orthogonal signals. Without loss of generality we assume

$$s_i = \sqrt{\mathcal{E}} e_i,$$

where e_i is the i th unit vector in \mathbb{R}^m , i.e., the vector that contains 1 at position i and 0 elsewhere, and \mathcal{E} is some positive constant.

- (a) Describe the maximum likelihood decision rule. (Make use of the fact that $s_i = \sqrt{\mathcal{E}} e_i$.)
- (b) Find the distance $\|s_i - s_j\|$.
- (c) Upper-bound the error probability $Pr\{e|H = i\}$ using the union bound and the Q function.

PROBLEM 36. (Data Storage Channel)

The process of storing and retrieving binary data on a thin-film disk may be modeled as transmitting binary symbols across an additive white Gaussian noise channel where the noise Z has a variance that depends on the transmitted (stored) binary symbol S . The noise has the following input-dependent density:

$$f_Z(z) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{z^2}{2\sigma_1^2}} & \text{if } S = 1 \\ \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{z^2}{2\sigma_0^2}} & \text{if } S = 0, \end{cases}$$

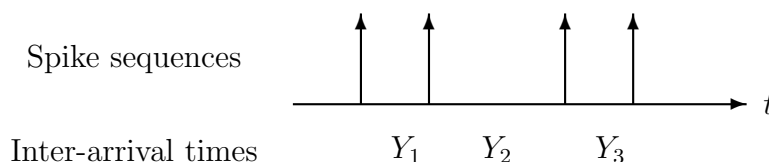
where $\sigma_1 > \sigma_0$. The channel inputs are equally likely.

- (a) On the same graph, plot the two possible output probability density functions. Indicate, qualitatively, the decision regions.
- (b) Determine the optimal receiver in terms of σ_1 and σ_0 . (Use the back of the previous page for details.)
- (c) Write an expression for the error probability P_e as a function of σ_0 and σ_1 .

PROBLEM 37. (Lie Detector)

You are asked to develop a “lie detector” and analyze its performance. Based on the observation of brain cell activity, your detector has to decide if a person is telling the truth or is lying.

For the purpose of this problem, the brain cell produces a sequence of spikes as shown in the figure. For your decision you may use only a sequence of n consecutive inter-arrival times Y_1, Y_2, \dots, Y_n . Hence Y_1 is the time elapsed between the first and second spike, Y_2 the time between the second and third, etc.



We assume that, a priori, a person lies with some known probability p . When the person is telling the truth, Y_1, \dots, Y_n is an i.i.d. sequence of exponentially distributed random variables with intensity α , ($\alpha > 0$), i.e.

$$f_{Y_i}(y) = \alpha e^{-\alpha y}, \quad y \geq 0.$$

When the person lies, Y_1, \dots, Y_n is i.i.d. exponentially distributed with intensity β , ($\alpha < \beta$).

- (a) Describe the decision rule of your lie detector for the special case $n = 1$. Your detector shall be designed so as to minimize the probability of error.
- (b) What is the probability $P_{L/T}$ that your lie detector says that the person is lying when the person is telling the truth?
- (c) What is the probability $P_{T/L}$ that your test says that the person is telling the truth when the person is lying.
- (d) Repeat (a) and (b) for a general n . Hint: There is no need to repeat every step of your previous derivations.

PROBLEM 38. (Fault Detector)

As an engineer, you are required to design the test performed by a fault-detector for a “black-box” that produces a sequence of i.i.d. binary random variables $\dots, X_1, X_2, X_3, \dots$. Previous experience shows that this “black box” has an a priori failure probability of $\frac{1}{1025}$. When the “black box” works properly, $p_{X_i}(1) = p$. When it fails, the output symbols are equally likely to be 0 or 1.

Your detector has to decide based on the observation of the past 16 symbols, i.e., at time k the decision will be based on X_{k-16}, \dots, X_{k-1} .

- (a) Describe your test.
- (b) What does your test decide if it observes the output sequence 0101010101010101? Assume that $p = 1/4$.

PROBLEM 39. (A Simple Multiple-Access Scheme)

Consider the following very simple model of a multiple-access scheme. There are two users. Each user has two hypotheses. Let $\mathcal{H}^1 = \mathcal{H}^2 = \{0, 1\}$ denote the respective set of hypotheses and assume that both users employ a uniform prior. Further, let X^1 and X^2 be the respective signals sent by user one and two. Assume that the transmissions of both users are independent and that $X^1 \in \{\pm 1\}$ and $X^2 \in \{\pm 2\}$ where X^1 and X^2 are positive if their respective hypothesis is zero and negative otherwise. Assume that the receiver observes the signal $Y = X^1 + X^2 + Z$, where Z is a zero mean Gaussian random variable with variance σ^2 and is independent of the transmitted signal.

- (a) Assume that the receiver observes Y and wants to estimate both transmitted signals, i.e., the receiver forms the estimate $\hat{H} = (\hat{H}^1, \hat{H}^2)$. Starting from first principles, what is the generic form of the optimal decision rule?
- (b) For the specific set of signals given, what is the set of possible observations assuming that $\sigma^2 = 0$? Label these signals by the corresponding (joint) hypotheses.
- (c) Assuming now that $\sigma^2 > 0$, draw the optimal decision regions.
- (d) What is the resulting probability of correct decision? i.e., determine the probability $P\{\hat{H}^1 = H^1, \hat{H}^2 = H^2\}$.
- (e) Finally, assume that we are only interested in the transmission of user two. What is $P\{\hat{H}^2 = H^2\}$?

PROBLEM 40. (Uncoded Transmission)

Consider the following transmission scheme. We have two possible sequences $\{X_j^1\}$ and $\{X_j^2\}$ taking values in $\{-1, +1\}$, for $j = 0, 1, 2, \dots, k-1$. The transmitter chooses one of the two sequences and sends it directly over an additive white Gaussian noise channel. Thus, the received value is $Y_j = X_j^i + Z_j$, where $i = 1, 2$ depending of the transmitted

sequence, and $\{Z_j\}$ is a sequence of i.i.d. zero-mean Gaussian random variables with variance σ^2 .

- (a) Using basic principles, write down the optimal decision rule that the receiver should implement to distinguish between the two possible sequences. Simplify this rule to express it as a function of inner products of vectors.
- (b) Let d be the number of positions in which $\{X_j^1\}$ and $\{X_j^2\}$ differ. Assuming that the transmitter sends the first sequences $\{X_j^1\}$, find the probability of error (the probability that the receiver decides on $\{X_j^2\}$), in terms of the Q function and d .

PROBLEM 41. (Data Dependent Noise)

Consider the following binary Gaussian hypothesis testing problem with data dependent noise.

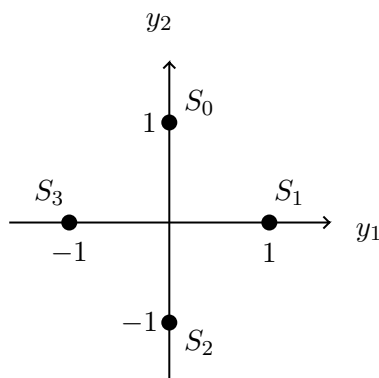
Under hypothesis H_0 the transmitted signal is $s_0 = -1$ and the received signal is $Y = s_0 + Z_0$, where Z_0 is zero-mean Gaussian with variance one.

Under hypothesis H_1 on the other hand, the transmitted signal is $s_1 = 1$ and the received signal is $Y = s_1 + Z_1$, where Z_1 is zero-mean Gaussian with variance σ^2 . Assume that the prior is uniform.

- (a) Starting from first principles write down the optimal decision rule as a function of the parameter σ^2 and the received signal Y .
- (b) For the value $\sigma^2 = \exp(4)$ compute the decision regions.
- (c) Give as simple expressions as possible for the error probabilities $Pr\{e|0\}$ and $Pr\{e|1\}$.

Hint: It might be handy to recall that the solutions of the quadratic equation $ax^2+bx+c = 0$ are given by $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$.

PROBLEM 42. (Decision Problem) Consider the following decision problem. For the hypothesis $H = i$, $i \in \{0, 1, 2, 3\}$, we send the point S_i , as shown in the figure below.



$$S_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, S_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, S_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, S_3 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

The receiver observes the vector $Y = S_i + Z$, where Z is a zero-mean Gaussian random vector whose covariance matrix is $\Sigma_Z = \begin{pmatrix} 4 & 2 \\ 2 & 5 \end{pmatrix}$.

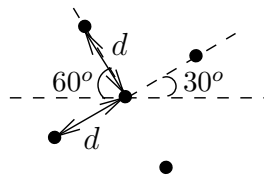
- (a) In order to simplify the decision problem, we transform Y into $\hat{Y} = BY$, where B is a 2-by-2 matrix, and use \hat{Y} to take our decision. What is the appropriated matrix B to choose?

Hint: If $A = \frac{1}{4} \begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix}$, then $A\Sigma_Z A^T = I$, with $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

- (b) What are the new transmitted points \hat{S}_i ? Draw the resulting transmitted points and the decision regions associated to them.
- (c) Give an upper bound to the error probability in this decision problem.

PROBLEM 43. (Signal Constellation under AWGN)

Consider the following 5-ary hypothesis testing problems. We assume that, the prior is uniform, the signal constellation consists of elements of \mathbb{R}^2 , and the noise is additive Gaussian with independent components of variance σ^2 in each dimension.



- (a) Draw the boundaries of the decision regions into the above figure.
- (b) Give the exact probability of error for the middle point of the constellation.
- (c) Find an upper bound on the probability of error, $\Pr\{e\}$, using the union bounding technique.
- (d) Consider the following set of parameters:

$$\sigma^2 = 1, d = 1 \quad (1)$$

$$\sigma^2 = 2, d = 2 \quad (2)$$

$$\sigma^2 = 4, d = 2 \quad (3)$$

Which one do you prefer and why?

PROBLEM 44. (Football)

Consider four teams A, B, C, D playing in a football tournament. There are two rounds in the competition. In the first round there are two matches and the winners progress to play in the final. In the first round A plays against one of the other three teams with equal probability $\frac{1}{3}$ and the remaining two teams play against each other. The probability of A winning against any team depends on the number of red cards “ r ” A gets in the previous match. The probabilities of winning for A against B, C, D denoted by p_b, p_c, p_d are $p_b = \frac{0.5}{(1+r)}, p_c = p_d = \frac{0.6}{1+r}$. In a match against B , A will get 1 red card and in a match against C or D , A will get 2 red cards. Assuming that initially A has 0 red cards and the other teams receive no red cards in the entire tournament and among B, C, D each team has equal chances to win against each other.

Is betting on team A as the winner a good choice ?

PROBLEM 45. (Fourier Transform)

- (a) Prove that if $x(t)$ is a real-valued signal, then its Fourier transform $X(f)$ satisfies the symmetric property

$$X(f) = X^*(-f) \quad (\text{Symmetry Property})$$

where X^* is the complex conjugate of X .

- (b) Prove that if $x(t)$ is a purely imaginary-valued signal, then its Fourier transform $X(f)$ satisfies the anti-symmetry property

$$X(f) = -X^*(-f) \quad (\text{Anti-Symmetry Property})$$

PROBLEM 46. (Rayleigh distribution)

Let X and Y be two independent, zero-mean, unit-variance, Gaussian random variables: $(X, Y) = \mathcal{N}(0, \mathbf{I}_2)$. Let R and Θ be the corresponding polar coordinates, i.e., $X = R \cos \Theta$ and $Y = R \sin \Theta$. Find the probability density functions $f_{R, \Theta}$, f_R , and f_Θ .

Hint: Try to solve this problem without looking up formulas. First write down the expression of $f_{R, \Theta}$ as a function of $f_{X, Y}$ assuming that you have a linear transformation of the kind $(X, Y)^T = A(R, \Theta)^T$ for some 2×2 invertible matrix A . See the Appendix of the lecture notes, Chapter 2, if you don't know how to do this from memory using the hints given in class. Recall that $\det A$ equals $\frac{1}{\det A^{-1}}$, which means that you may work with the determinant of A or with that of A^{-1} , whichever is more convenient. Next, instead of A use the Jacobian J of the transformation that maps R, Θ into X, Y . The Jacobian is the matrix that maps $(dR, d\Theta)^T$ into $(dX, dY)^T$.

Appendix 2.A Facts About Matrices

We now review a few definitions and results that will be useful throughout. Hereafter H^\dagger is the conjugate transpose of H also called the *Hermitian adjoint* of H .

DEFINITION 9. A matrix $U \in \mathbb{C}^{n \times n}$ is said to be unitary if $U^\dagger U = I$. If, in addition, $U \in \mathbb{R}^{n \times n}$, U is said to be orthogonal. \square

The following theorem lists a number of handy facts about unitary matrices. Most of them are straightforward. For a proof see [1, page 67].

THEOREM 10. if $U \in \mathbb{C}^{n \times n}$, the following are equivalent:

- (a) U is unitary;
- (b) U is nonsingular and $U^\dagger = U^{-1}$;
- (c) $UU^\dagger = I$;
- (d) U^\dagger is unitary
- (e) The columns of U form an orthonormal set;
- (f) The rows of U form an orthonormal set; and
- (g) For all $\mathbf{x} \in \mathbb{C}^n$ the Euclidean length of $\mathbf{y} = U\mathbf{x}$ is the same as that of \mathbf{x} ; that is, $\mathbf{y}^\dagger \mathbf{y} = \mathbf{x}^\dagger \mathbf{x}$.

THEOREM 11. (Schur) Any square matrix A can be written as

$$A = URU^\dagger$$

where U is unitary and R is an upper-triangular matrix whose diagonal entries are the eigenvalues of A .

Proof. Let us use induction on the size n of the matrix. The theorem is clearly true for $n = 1$. Let us now show that if it is true for $n - 1$ it follows that it is true for n . Given A of size n , let \mathbf{v} be an eigenvector of unit norm, and λ the corresponding eigenvalue. Let V be a unitary matrix whose first column is \mathbf{v} . Then, consider the matrix

$$V^\dagger AV.$$

Now, the first column of this matrix is given by

$$V^\dagger A\mathbf{v} = \lambda V^\dagger \mathbf{v} = \lambda \mathbf{e}_1$$

where \mathbf{e}_1 is the unit vector along the first coordinate. Thus

$$V^\dagger AV = \begin{pmatrix} \lambda & * \\ 0 & B \end{pmatrix},$$

where B is square and of dimension $n - 1$. By the induction hypothesis $B = WSW^\dagger$, where W is unitary and S is upper triangular. Thus,

$$V^\dagger AV = \begin{pmatrix} \lambda & * & 0 & WSW^\dagger \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & W \end{pmatrix} \begin{pmatrix} \lambda & * \\ 0 & S \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & W^\dagger \end{pmatrix} \quad (2.49)$$

and putting

$$U = V \begin{pmatrix} 1 & 0 \\ 0 & W \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} \lambda & * \\ 0 & S \end{pmatrix},$$

we see that U is unitary, R is upper-triangular and $A = URU^\dagger$, completing the induction step. To see that the diagonal entries of R are indeed the eigenvalues of A it suffices to bring the characteristic polynomial of A in the following form: $\det(\lambda I - A) = \det[U^\dagger(\lambda I - R)U] = \det(\lambda I - R) = \prod_i (\lambda - r_{ii})$. \square

DEFINITION 12. A matrix $H \in \mathbb{C}^{n \times n}$ is said to be Hermitian if $H = H^\dagger$. It is said to be Skew-Hermitian if $H = -H^\dagger$.

Recall that an $n \times n$ matrix has exactly n eigenvalues in \mathbb{C} .

LEMMA 13. A Hermitian matrix $H \in \mathbb{C}^{n \times n}$ can be written as

$$H = U\Lambda U^\dagger = \sum_i \lambda_i \mathbf{u}_i \mathbf{u}_i^\dagger$$

where U is unitary and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ is a diagonal that consists of the eigenvalues of H . Moreover, the eigenvalues are real and the i th column of U is an eigenvector associated to λ_i .

Proof. By Theorem 11 (Schur) we can write $H = URU^\dagger$ where U is unitary and R is upper triangular with the diagonal elements consisting of the eigenvalues of A . From $R = U^\dagger H U$ we immediately see that R is Hermitian. Hence it is diagonal and the diagonal elements must be real.

If \mathbf{u}_i is the i th column of U , then

$$H\mathbf{u}_i = U\Lambda U^\dagger \mathbf{u}_i = U\Lambda \mathbf{e}_i = U\lambda_i \mathbf{e}_i = \lambda_i \mathbf{u}_i$$

showing that it is indeed an eigenvector associated to the i th eigenvalue λ_i . \square

The reader interested in properties of Hermitian matrices is referred to [1, Section 4.1].

EXERCISE 14. Show that if $H \in \mathbb{C}^{n \times n}$ is Hermitian, then $\mathbf{u}^\dagger H \mathbf{u}$ is real for all $\mathbf{u} \in \mathbb{C}^n$.

A class of Hermitian matrices with a special positivity property arises naturally in many applications, including communication theory. They provide a generalization to matrices of the notion of positive numbers.

DEFINITION 15. An Hermitian matrix $H \in \mathbb{C}^{n \times n}$ is said to be positive definite if

$$\mathbf{u}^\dagger H \mathbf{u} > 0 \quad \text{for all non zero } \mathbf{u} \in \mathbb{C}^n.$$

If the above strict inequality is weakened to $\mathbf{u}^\dagger H \mathbf{u} \geq 0$, then A is said to be positive semidefinite. Implicit in these defining inequalities is the observation that if H is Hermitian, the left hand side is always a real number.

EXAMPLE 16. Show that a non-singular covariance matrix is always positive definite. \square

THEOREM 17. (SVD) Any matrix $A \in \mathbb{C}^{m \times n}$ can be written as a product

$$A = UDV^\dagger,$$

where U and V are unitary (of dimension $m \times m$ and $n \times n$, respectively) and $D \in \mathbb{R}^{m \times n}$ is non-negative and diagonal. This is called the singular value decomposition (SVD) of A . Moreover, letting k be the rank of A , the following statements are true:

(i) The columns of V are the eigenvectors of $A^\dagger A$. The last $n - k$ columns span the null space of A .

(ii) The columns of U are eigenvectors of AA^\dagger . The first k columns span the range of A .

(iii) If $m \geq n$ then

$$D = \begin{pmatrix} \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) \\ \dots\dots\dots \\ \mathbf{0}_{m-n} \end{pmatrix},$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > \lambda_{k+1} = \dots = \lambda_n = 0$ are the eigenvalues of $A^\dagger A \in \mathbb{C}^{n \times n}$ which are non-negative since $A^\dagger A$ is Hermitian. If $m \leq n$ then

$$D = (\text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_m}) : \mathbf{0}_{n-m}),$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > \lambda_{k+1} = \dots = \lambda_m = 0$ are the eigenvalues of AA^\dagger .

Note 1: Recall that the nonzero eigenvalues of AB equals the nonzero eigenvalues of BA , see e.g. Horn and Johnson, Theorem 1.3.29. Hence the nonzero eigenvalues in (iii) are the same for both cases.

Note 2: To remember that V is associated to $H^\dagger H$ (as opposed to being associated to HH^\dagger) it suffices to look at the dimensions: $V \in \mathbb{R}^n$ and $H^\dagger H \in \mathbb{R}^{n \times n}$.

Proof. It is sufficient to consider the case with $m \geq n$ since if $m < n$ we can apply the result to $A^\dagger = UDV^\dagger$ and obtain $A = VD^\dagger U^\dagger$.

Hence let $m \geq n$, and consider the matrix $A^\dagger A \in \mathbb{C}^{n \times n}$. This matrix is Hermitian. Hence its eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \lambda_n \geq 0$ are real and non-negative and one can choose the eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ to form an orthonormal basis for \mathbb{C}^n . Let $V = (\mathbf{v}_1, \dots, \mathbf{v}_n)$. Let k be the number of positive eigenvectors and choose.

$$\mathbf{u}_i = \frac{1}{\sqrt{\lambda_i}} A \mathbf{v}_i, \quad i = 1, 2, \dots, k. \quad (2.50)$$

Observe that

$$\mathbf{u}_i^\dagger \mathbf{u}_j = \frac{1}{\sqrt{\lambda_i \lambda_j}} \mathbf{v}_i^\dagger A^\dagger A \mathbf{v}_j = \sqrt{\frac{\lambda_j}{\lambda_i}} \mathbf{v}_i^\dagger \mathbf{v}_j = \delta_{ij}, \quad 0 \leq i, j \leq k.$$

Hence $\{\mathbf{u}_i : i = 1, \dots, k\}$ form an orthonormal set in \mathbb{C}^m . Complete this set to an orthonormal basis for \mathbb{C}^m by choosing $\{\mathbf{u}_i : i = k+1, \dots, m\}$ and let $U = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m)$. Note that (2.50) implies

$$\mathbf{u}_i \sqrt{\lambda_i} = A \mathbf{v}_i, \quad i = 1, 2, \dots, k, k+1, \dots, n,$$

where for $i = k+1, \dots, n$ the above relationship holds since $\lambda_i = 0$ and \mathbf{v}_i is a corresponding eigenvector. Using matrix notation we obtain

$$U \begin{pmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_n} \\ \dots & & \dots \\ & & \mathbf{0}_{m-n} \end{pmatrix} = AV, \quad (2.51)$$

i.e., $A = UDV^\dagger$. For $i = 1, 2, \dots, m$,

$$\begin{aligned} AA^\dagger \mathbf{u}_i &= UDV^\dagger V^\dagger D^\dagger U^\dagger \mathbf{u}_i \\ &= UDD^\dagger U^\dagger \mathbf{u}_i = \mathbf{u}_i \lambda_i, \end{aligned}$$

where the last equality follows from the fact that $U^\dagger \mathbf{u}_i$ has a 1 at position i and is zero otherwise and $DD^\dagger = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k, 0, \dots, 0)$. This shows that λ_i is also an eigenvalue of AA^\dagger . We have also shown that $\{\mathbf{v}_i : i = k+1, \dots, n\}$ spans the null space of A and from (2.51) we see that $\{\mathbf{u}_i : i = 1, \dots, k\}$ spans the range of A . \square

The following key result is a simple application of the SVD.

LEMMA 18. *The linear transformation described by a matrix $A \in \mathbb{R}^{n \times n}$ maps the unit cube into a parallelepiped of volume $|\det A|$.*

Proof. (Question to the students: do we need to review what a unit cube is, that the linear transformation maps \mathbf{e}_i into the vector \mathbf{a}_i that forms the i -th column of A , and that the volume of an n -dimensional object (set) \mathcal{A} is $\int_{\mathcal{A}} d\mathbf{x}$?) From the singular value decomposition, $A = UDV^\dagger$, where D is diagonal and U and V are orthogonal

matrices. The linear transformation associated to A is the same as that associated to $U^\dagger AV = D$. (We are just changing the coordinate system). But D maps the unit vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ into $\lambda_1 \mathbf{e}_1, \lambda_2 \mathbf{e}_2, \dots, \lambda_n \mathbf{e}_n$. Hence, the unit cube is mapped into a rectangle of sides $\lambda_1, \lambda_2, \dots, \lambda_n$. Its volume is $|\prod \lambda_i| = |\det D| = |\det A|$. \square

Appendix 2.B Densities After Linear Transformations

The previous result leads to the following fundamental result.

THEOREM 19. *Let $\mathbf{X} \in \mathbb{R}^n$ be a random vector of given pdf $f_{\mathbf{X}}(\mathbf{x})$, $A \in \mathbb{R}^{n \times n}$ a non-singular matrix, and \mathbf{Y} be defined through the linear transformation $\mathbf{Y} = A\mathbf{X}$. The pdf of \mathbf{Y} is given by*

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{f_{\mathbf{X}}(A^{-1}\mathbf{y})}{|\det A|}.$$

Outline of the proof: The probability that \mathbf{X} is inside an infinitesimally small cube $\delta\mathcal{A}$ (in n -dimensions) is $f_{\mathbf{X}}(\mathbf{x}^*)\delta\mathcal{A}$ (plus terms that become negligible as the volume of $\delta\mathcal{A}$ goes to zero), where \mathbf{x}^* is any point inside $\delta\mathcal{A}$. Now \mathbf{x}^* maps into $\mathbf{y}^* = A\mathbf{x}^*$ and $\delta\mathcal{A}$ into some $\delta\mathcal{B}$ of volume $\text{Vol}(\delta\mathcal{B}) = \text{Vol}(\delta\mathcal{A})|\det A|$. Since the probability that \mathbf{Y} is inside $\delta\mathcal{B}$ is the same as the probability that \mathbf{X} is inside $\delta\mathcal{A}$ we have (in the limit):

$$\text{Vol}(\delta\mathcal{B})f_{\mathbf{Y}}(\mathbf{y}^*) = \text{Vol}(\delta\mathcal{A})f_{\mathbf{X}}(\mathbf{x}^*).$$

Solving for $f_{\mathbf{Y}}(\mathbf{y}^*)$ yields the desired result. \square

Appendix 2.C Gaussian Random Vectors

We now study Gaussian random vectors. A Gaussian random vector is nothing else than a collection of jointly Gaussian random variables. We learn to use vector notation since this will simplify matters significantly.

Recall that a random variable W is a mapping $W : \Omega \rightarrow \mathbb{R}$ from the sample space Ω to the reals \mathbb{R} . W is a Gaussian random variable with mean m and variance σ^2 if and only if (iff) its probability density function (pdf) is

$$f_W(w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(w-m)^2}{2\sigma^2} \right\}.$$

Since a Gaussian random variable is completely specified by its mean m and variance σ^2 , we use the short-hand notation $\mathcal{N}(m, \sigma^2)$ to denote its pdf. Hence $W \sim \mathcal{N}(0, \sigma^2)$.

An n -dimensional random vector (n -rv) \mathbf{X} is a mapping $\mathbf{X} : \Omega \rightarrow \mathbb{R}^n$. It can be seen as a collection $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ of n random variables. The pdf of \mathbf{X} is the joint pdf

of X_1, X_2, \dots, X_n . The expected value of \mathbf{X} , denoted by $E\mathbf{X}$ or by $\bar{\mathbf{X}}$, is the n -tuple $(EX_1, EX_2, \dots, EX_n)^T$. The *covariance matrix* of \mathbf{X} is $K_{\mathbf{X}} = E[(\mathbf{X} - \bar{\mathbf{X}})(\mathbf{X} - \bar{\mathbf{X}})^T]$. Notice that $\mathbf{X}\mathbf{X}^T$ is an $n \times n$ random matrix, i.e., a matrix of random variables, and the expected value of such a matrix is, by definition, the matrix whose components are the expected values of those random variables. Notice that a covariance matrix is always Hermitian.

The pdf of a vector $\mathbf{W} = (W_1, W_2, \dots, W_n)^T$ that consists of independent and identically distributed (iid) $\sim \mathcal{N}(0, \sigma^2)$ components is

$$f_{\mathbf{W}}(\mathbf{w}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{w_i^2}{2\sigma^2}\right) \quad (2.52)$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\mathbf{w}^T \mathbf{w}}{2\sigma^2}\right). \quad (2.53)$$

The following is one of several possible ways to define a Gaussian random vector.

DEFINITION 20. *The random vector $\mathbf{Y} \in \mathbb{R}^m$ is a zero-mean Gaussian random vector and Y_1, Y_2, \dots, Y_n are zero-mean jointly Gaussian random variables, iff there exists a matrix $A \in \mathbb{R}^{m \times n}$ such that \mathbf{Y} can be expressed as*

$$\mathbf{Y} = A\mathbf{W} \quad (2.54)$$

where \mathbf{W} is a random vector of iid $\sim \mathcal{N}(0, 1)$ components.

NOTE 21. *From the above definition it follows immediately that linear combination of zero-mean jointly Gaussian random variables are zero-mean jointly Gaussian random variables. Indeed, $\mathbf{Z} = B\mathbf{Y} = BAW$. \square*

Recall that if $\mathbf{Y} = A\mathbf{W}$ for some nonsingular matrix $A \in \mathbb{R}^{n \times n}$, then

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{f_{\mathbf{W}}(A^{-1}\mathbf{y})}{|\det A|}.$$

When \mathbf{W} has iid $\sim \mathcal{N}(0, 1)$ components,

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\exp\left(-\frac{(A^{-1}\mathbf{y})^T(A^{-1}\mathbf{y})}{2}\right)}{(2\pi)^{n/2}|\det A|}.$$

The above expression can be simplified and brought to the standard expression

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^n \det K_{\mathbf{Y}}}} \exp\left(-\frac{1}{2}\mathbf{y}^T K_{\mathbf{Y}}^{-1} \mathbf{y}\right) \quad (2.55)$$

using $K_{\mathbf{Y}} = EAW(AW)^T = EAWW^T A^T = AI_n A^T = AA^T$ to obtain

$$\begin{aligned} (A^{-1}\mathbf{y})^T(A^{-1}\mathbf{y}) &= \mathbf{y}^T(A^{-1})^T A^{-1} \mathbf{y} \\ &= \mathbf{y}^T(AA^T)^{-1} \mathbf{y} \\ &= \mathbf{y}^T K_{\mathbf{Y}}^{-1} \mathbf{y} \end{aligned}$$

and

$$\sqrt{\det K_{\mathbf{Y}}} = \sqrt{\det AA^T} = \sqrt{\det A \det A} = |\det A|.$$

FACT 22. Let \mathbf{Y} be a zero-mean random vector with arbitrary covariance matrix $K_{\mathbf{Y}}$ and pdf as in (2.55). Since a covariance matrix is Hermitian, we can write (see Appendix 2.A)

$$K_{\mathbf{Y}} = U\Lambda U^\dagger \quad (2.56)$$

where U is unitary and Λ is diagonal. It is immediate to verify that $U\sqrt{\Lambda}\mathbf{W}$ has covariance $K_{\mathbf{Y}}$. This shows that an arbitrary zero-mean random vector \mathbf{Y} with pdf as in (2.55) can always be written in the form $\mathbf{Y} = A\mathbf{W}$ where \mathbf{W} has iid $\sim \mathcal{N}(0,1)$ components.

The contrary is not true in degenerated cases. We have already seen that (2.55) follows from (2.54) when A is a non-singular squared matrix. The derivation extends to any non-rectangular matrix A , provided that it has linearly independent rows. This result is derived as a homework exercise. In that exercise we also see that it is indeed necessary that the rows of A be linearly independent since otherwise $K_{\mathbf{Y}}$ is singular and $K_{\mathbf{Y}}^{-1}$ is not defined. Then (2.55) is not defined either. An example will show how to handle such degenerated cases.

It should be pointed out that many authors use (2.55) to define a Gaussian random vector. We favor (2.54) because it is more general, but also since it makes it straightforward to prove a number of key results associated to Gaussian random vectors. Some of these are dealt with in the examples below.

In any case, a zero-mean Gaussian random vector is completely characterized by its covariance matrix. Hence the short-hand notation $\mathbf{Y} \sim \mathcal{N}(0, K_{\mathbf{Y}})$.

NOTE 23. (Degenerate case) Let $W \sim \mathcal{N}(0,1)$, $A = (1,1)^T$, and $Y = AW$. By our definition, Y is a Gaussian random vector. However, A is a matrix of linearly dependent rows implying that \mathbf{Y} has linearly dependent components. Indeed $Y_1 = Y_2$. This also implies that $K_{\mathbf{Y}}$ is singular: it is a 2×2 matrix with 1 in each component. As already pointed out, we can't use (2.55) to describe the pdf of \mathbf{Y} . This immediately raises the question: how do we compute the probability of events involving \mathbf{Y} if we don't know its pdf? The answer is easy. Any event involving \mathbf{Y} can be rewritten as an event involving Y_1 only (or equivalently involving Y_2 only). For instance, the event $\{Y_1 \in [3, 5]\} \cap \{Y_2 \in [4, 6]\}$ occurs iff $\{Y_1 \in [4, 5]\}$. Hence

$$Pr \{Y_1 \in [3, 5]\} \cap \{Y_2 \in [4, 6]\} = Pr \{Y_1 \in [4, 5]\} = Q(4) - Q(5).$$

□

EXERCISE 24. Show that the i th component Y_i of a Gaussian random vector \mathbf{Y} is a Gaussian random variable.

Solution: $Y_i = A\mathbf{Y}$ when $A = \mathbf{e}_i^T$ is the unit row vector with 1 in the i -th component and 0 elsewhere. Hence Y_i is a Gaussian random variable. To appreciate the convenience of working with (2.54) instead of (2.55), compare this answer with the tedious derivation consisting of integrating over $f_{\mathbf{Y}}$ to obtain f_{Y_i} (see Problem 7). \square

EXERCISE 25. Let U be an orthogonal matrix. Determine the pdf of $\mathbf{Y} = U\mathbf{W}$.

Solution: \mathbf{Y} is zero-mean and Gaussian. Its covariance matrix is $K_{\mathbf{Y}} = UK_{\mathbf{W}}U^T = U\sigma^2I_nU^T = \sigma^2UU^T = \sigma^2I_n$, where I_n denotes the $n \times n$ identity matrix. Hence, when an n -dimensional Gaussian random vector with iid $\sim \mathcal{N}(0, \sigma^2)$ components is projected onto n orthonormal vectors, we obtain n iid $\sim \mathcal{N}(0, \sigma^2)$ random variables. This fact will be used often. \square

EXERCISE 26. (Gaussian random variables are not necessarily jointly Gaussian) Let $Y_1 \sim \mathcal{N}(0, 1)$, let $X \in \{\pm 1\}$ be uniformly distributed, and let $Y_2 = Y_1X$. Notice that Y_2 has the same pdf as Y_1 . This follows from the fact that the pdf of Y_1 is an even function. Hence Y_1 and Y_2 are both Gaussian. However, they are not jointly Gaussian. We come to this conclusion by observing that $Z = Y_1 + Y_2 = Y_1(1 + X)$ is 0 with probability 1/2. Hence Z can't be Gaussian. \square

EXERCISE 27. Is it true that uncorrelated Gaussian random variables are always independent? If you think it is ... think twice. The construction above labeled "Gaussian random variables are not necessarily jointly Gaussian" provides a counter example (you should be able to verify without much effort). However, the statement is true if the random variables under consideration are jointly Gaussian (the emphasis is on "jointly"). You should be able to provide an easy proof using (2.55). The contrary is always true: random variables (not necessarily Gaussian) that are independent are always uncorrelated. Again, you should be able to provide the straightforward proof. (You are strongly encouraged to brainstorm this and similar exercises with other students. Hopefully this will create healthy discussions. Let us know if you can't clear every doubt this way ... we are very much interested in knowing where the difficulties are.) \square

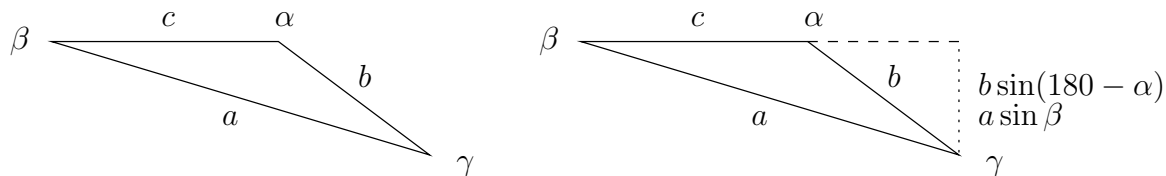
DEFINITION 28. The random vector \mathbf{Y} is a Gaussian random vector (and Y_1, \dots, Y_n are jointly Gaussian random variables) iff $\mathbf{Y} - m$ is a zero mean Gaussian random vector as defined above, where $m = E\mathbf{Y}$. If the covariance $K_{\mathbf{Y}}$ is non-singular (which implies that no component of \mathbf{Y} is determined by a linear combination of other components), then its pdf is

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^n \det K_{\mathbf{Y}}}} \exp\left(-\frac{1}{2}(\mathbf{y} - E\mathbf{y})^T K_{\mathbf{Y}}^{-1}(\mathbf{y} - E\mathbf{y})\right).$$

\square

Appendix 2.D A Fact About Triangles

To determine an exact expression of the probability of error, in Example 7 we use the following fact about triangles.



For a triangle with edges a , b , c and angles α , β , γ (see the figure), the following relationship holds:

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta} = \frac{c}{\sin \gamma}. \quad (2.57)$$

To prove the equality relating a and b we project the common vertex γ onto the extension of the segment connecting the other two edges (α and β). This projection gives rise to two triangles that share a common edge whose length can be written as $a \sin \beta$ and as $b \sin(180 - \alpha)$ (see right figure). Using $b \sin(180 - \alpha) = b \sin \alpha$ leads to $a \sin \beta = b \sin \alpha$. The second equality is proved similarly. \square

Appendix 2.E Inner Product Spaces

Vector Space

We assume that you are familiar with vector spaces. In this Chapter 2 we will be dealing with the vector space of n -tuples over \mathbb{R} but later we will need both the vector space of n -tuples over \mathbb{C} and the vector space of finite-energy complex-valued functions. So to be as general as needed we assume that the vector space is over the field of complex numbers, in which case it is called a *complex vector space*. When the scalar field is \mathbb{R} , the vector space is called a *real vector space*.

Inner Product Space

Given a vector space and nothing more, one can introduce the notion of a basis for the vector space, but one does not have the tool needed to define an orthonormal basis. Indeed the axioms of a vector space say nothing about geometric ideas such as “length” or “angle.” To remedy, one endows the vector space with the notion of inner product.

DEFINITION 29. Let V be a vector space over \mathbb{C} . An inner product on V is a function that assigns to each ordered pair of vectors α, β in V a scalar $\langle \alpha, \beta \rangle$ in \mathbb{C} in such a way

that for all α, β, γ in \mathcal{V} and all scalars c in \mathbb{C}

- (a) $\langle \alpha + \beta, \gamma \rangle = \langle \alpha, \gamma \rangle + \langle \beta, \gamma \rangle$
 $\langle c\alpha, \beta \rangle = c\langle \alpha, \beta \rangle;$
- (b) $\langle \beta, \alpha \rangle = \langle \alpha, \beta \rangle^*;$ (Hermitian Symmetry)
- (c) $\langle \alpha, \alpha \rangle \geq 0$ with equality iff $\alpha = 0$.

It is implicit in (c) that $\langle \alpha, \alpha \rangle$ is real for all $\alpha \in \mathcal{V}$. From (a) and (b), we obtain an additional property

- (d) $\langle \alpha, \beta + \gamma \rangle = \langle \alpha, \beta \rangle + \langle \alpha, \gamma \rangle$
 $\langle \alpha, c\beta \rangle = c^*\langle \alpha, \beta \rangle.$

Notice that the above definition is also valid for a vector space over the field of real numbers but in this case the complex conjugates appearing in (b) and (d) are superfluous; however, over the field of complex numbers they are necessary for the consistency of the conditions. Without these complex conjugates, for any $\alpha \neq 0$ we would have the contradiction:

$$0 < \langle i\alpha, i\alpha \rangle = -1\langle \alpha, \alpha \rangle < 0,$$

where the first inequality follows from condition (c) and the fact that $i\alpha$ is a valid vector, and the equality follows from (a) and (d) (without the complex conjugate).

On \mathbb{C}^n there is an inner product that is sometimes called the *standard inner product*. It is defined on $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{b} = (b_1, \dots, b_n)$ by

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_j a_j b_j^*.$$

On \mathbb{R}^n , the standard inner product is often called the dot or scalar product and denoted by $\mathbf{a} \cdot \mathbf{b}$. Unless explicitly stated otherwise, over \mathbb{R}^n and over \mathbb{C}^n we will always assume the standard inner product.

An *inner product space* is a real or complex vector space, together with a specified inner product on that space. We will use the letter \mathcal{V} to denote a generic inner product space.

EXAMPLE 30. The vector space \mathbb{R}^n equipped with the dot product is an inner product space and so is the vector space \mathbb{C}^n equipped with the standard inner product. \square

By means of the inner product we introduce the notion of length, called *norm*, of a vector α , via

$$\|\alpha\| = \sqrt{\langle \alpha, \alpha \rangle}.$$

Using linearity, we immediately obtain that the *squared norm* satisfies

$$\|\alpha \pm \beta\|^2 = \langle \alpha \pm \beta, \alpha \pm \beta \rangle = \|\alpha\|^2 + \|\beta\|^2 \pm 2\operatorname{Re}\{\langle \alpha, \beta \rangle\}. \quad (2.58)$$

The above generalizes $(a \pm b)^2 = a^2 + b^2 \pm 2ab$, $a, b \in \mathbb{R}$, and $|a \pm b|^2 = |a|^2 + |b|^2 \pm 2\operatorname{Re}\{ab\}$, $a, b \in \mathbb{C}$.

EXAMPLE 31. Consider the vector space V spanned by a finite collection of complex-valued finite-energy signals, where addition of vectors and multiplication of a vector with a scalar (in \mathbb{C}) are defined in the obvious way. You should verify that the axioms of a vector space are fulfilled. This includes showing that the sum of two finite-energy signals is a finite-energy signal. The standard inner product for this vectors space is defined as

$$\langle \alpha, \beta \rangle = \int \alpha(t)\beta^*(t)dt$$

which implies the norm

$$\|\alpha\| = \sqrt{\int |\alpha(t)|^2 dt}.$$

□

EXAMPLE 32. The previous example extends to the inner product space \mathcal{L}_2 of all complex-valued finite-energy functions. This is an infinite dimensional inner product space and to be careful one has to deal with some technicalities that we will just mention here. (You may skip the rest of this example if you wish without losing anything important for the sequel). If α and β are two finite-energy functions that are identical except on a countable number of points, then $\langle \alpha - \beta, \alpha - \beta \rangle = 0$ (the integral is over a function that vanishes except for a countable number of points). The definition of inner product requires that $\alpha - \beta$ be the zero vector. This seems to be in contradiction with the fact that $\alpha - \beta$ is non-zero on a countable number of point. To deal with this apparent contradiction one can define vectors to be equivalence classes of finite-energy functions. In other words, if the norm of $\alpha - \beta$ vanishes then α and β are considered to be the same vector and $\alpha - \beta$ is seen as a zero vector. This equivalence may seem artificial at first but it is actually consistent with the reality that if $\alpha - \beta$ has zero energy then no instrument will be able to distinguish between α and β . The signal captured by the antenna of a receiver is finite energy, thus in \mathcal{L}_2 . It is for this reason that we are interested in \mathcal{L}_2 . However, as we will see, the receiver may obtain a sufficient statistics by projecting the received signal on a finite-dimension subspace of \mathcal{L}_2 . During our brief exposures with \mathcal{L}_2 we will not be confronted with the subtle issues we have just mentioned. □

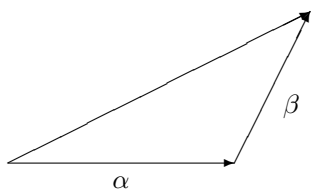
THEOREM 33. If \mathcal{V} is an inner product space, then for any vectors α, β in \mathcal{V} and any scalar c ,

- (a) $\|c\alpha\| = |c|\|\alpha\|$
- (b) $\|\alpha\| \geq 0$ with equality iff $\alpha = 0$
- (c) $|\langle \alpha, \beta \rangle| \leq \|\alpha\|\|\beta\|$ with equality iff $\alpha = c\beta$ for some c .
(Cauchy-Schwarz inequality)
- (d) $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$ with equality iff $\alpha = c\beta$ for some non-negative $c \in \mathbb{R}$.
(Triangle inequality)

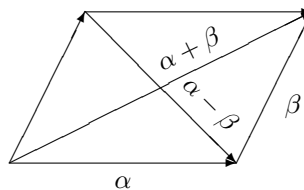
$$(e) \quad \|\alpha + \beta\|^2 + \|\alpha - \beta\|^2 = 2(\|\alpha\|^2 + \|\beta\|^2)$$

(Parallelogram equality)

Proof. Statements (a) and (b) follow immediately from the definitions. We postpone the proof of the Cauchy-Schwarz inequality to Example 35 since it will be more insightful once we have defined the concept of a projection. To prove the triangle inequality we use (2.58) and the Cauchy-Schwarz inequality applied to $\operatorname{Re}\{\langle\alpha, \beta\rangle\} \leq |\langle\alpha, \beta\rangle|$ to prove that $\|\alpha + \beta\|^2 \leq (\|\alpha\| + \|\beta\|)^2$. You should verify that $\operatorname{Re}\{\langle\alpha, \beta\rangle\} \leq |\langle\alpha, \beta\rangle|$ holds with equality iff $\alpha = c\beta$ for some non-negative $c \in \mathbb{R}$. Hence this condition is necessary for the triangle inequality to hold with equality. It is also sufficient since then also the Cauchy-Schwarz inequality holds with equality. The parallelogram equality follows immediately from (2.58) used twice, once with each sign. \square



Triangle inequality



Parallelogram equality

At this point we could use the inner product and the norm to define the angle between two vectors but we don't have any use for that. Instead, we will make frequent use of the notion of orthogonality. Two vectors α and β are defined to be *orthogonal* if $\langle\alpha, \beta\rangle = 0$.

THEOREM 34. (Pythagorean Theorem) *If α and β are orthogonal vectors in \mathcal{V} , then*

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2.$$

Proof. The Pythagorean theorem follows immediately from the equality $\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2 + 2\operatorname{Re}\{\langle\alpha, \beta\rangle\}$ and the fact that $\langle\alpha, \beta\rangle = 0$ by definition of orthogonality. \square

Given two vectors $\alpha, \beta \in \mathcal{V}$, $\beta \neq 0$, we define the *projection of α on β* as the vector $\alpha_{|\beta}$ collinear to β (i.e. of the form $c\beta$ for some scalar c) such that $\alpha_{\perp\beta} = \alpha - \alpha_{|\beta}$ is orthogonal to β . Using the definition of orthogonality, what we want is

$$0 = \langle\alpha_{\perp\beta}, \beta\rangle = \langle\alpha - c\beta, \beta\rangle = \langle\alpha, \beta\rangle - c\|\beta\|^2.$$

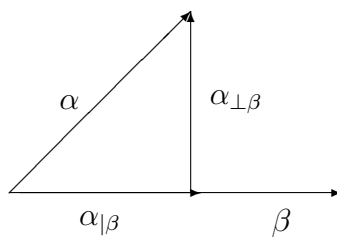
Solving for c we obtain $c = \frac{\langle\alpha, \beta\rangle}{\|\beta\|^2}$. Hence

$$\alpha_{|\beta} = \frac{\langle\alpha, \beta\rangle}{\|\beta\|^2}\beta \quad \text{and} \quad \alpha_{\perp\beta} = \alpha - \alpha_{|\beta}.$$

The projection of α on β does not depend on the norm of β . To see this let $\beta = b\psi$ for some $b \in \mathbb{C}$. Then

$$\alpha_{|\beta} = \langle\alpha, \psi\rangle\psi = \alpha_{|\psi},$$

regardless of b . The norm of the projection is $\langle\alpha, \psi\rangle = \langle\alpha, \beta\rangle/\|\beta\|$.

Projection of α on β

Any non-zero vector β defines a *hyperplane* by the relationship

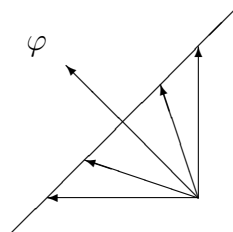
$$\{\alpha \in \mathcal{V} : \langle \alpha, \beta \rangle = 0\}.$$

It is the set of vectors that are orthogonal to β . A hyperplane always contains the zero vector.

An *affine space*, defined by a vector β and a scalar c , is an object of the form

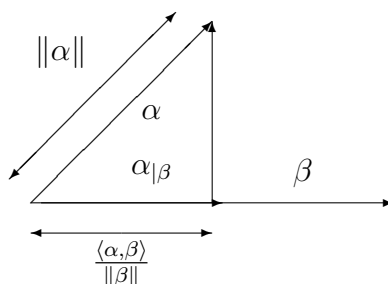
$$\{\alpha \in \mathcal{V} : \langle \alpha, \beta \rangle = c\}.$$

The defining vector and scalar are not unique, unless we agree that we use only normalized vectors to define hyperplanes. By letting $\varphi = \frac{\beta}{\|\beta\|}$, the above definition of affine plane may equivalently be written as $\{\alpha \in \mathcal{V} : \langle \alpha, \varphi \rangle = \frac{c}{\|\beta\|}\}$ or even as $\{\alpha \in \mathcal{V} : \langle \alpha - \frac{c}{\|\beta\|}\varphi, \varphi \rangle = 0\}$. The first shows that an affine plane is the set of vectors that have the same projection $\frac{c}{\|\beta\|}\varphi$ on φ . The second form shows that the affine plane is a hyperplane translated by the vector $\frac{c}{\|\beta\|}\varphi$. Some authors make no distinction between affine planes and hyperplanes. In that case both are called hyperplane.

Affine plane defined by φ .

Now it is time to prove the Cauchy-Schwarz inequality stated in Theorem 33. We do it as an application of a projection.

EXAMPLE 35. (Proof of the Cauchy-Schwarz Inequality). *The Cauchy-Schwarz inequality says states that for any $\alpha, \beta \in \mathcal{V}$, $|\langle \alpha, \beta \rangle| \leq \|\alpha\|\|\beta\|$ with equality iff $\alpha = c\beta$ for some scalar $c \in \mathbb{C}$. The statement is obviously true if $\beta = 0$. Assume $\beta \neq 0$ and write $\alpha = \alpha_{\parallel\beta} + \alpha_{\perp\beta}$. The Pythagorean theorem states that $\|\alpha\|^2 = \|\alpha_{\parallel\beta}\|^2 + \|\alpha_{\perp\beta}\|^2$. If we drop the second term, which is always nonnegative, we obtain $\|\alpha\|^2 \geq \|\alpha_{\parallel\beta}\|^2$ with equality iff α and β are collinear. From the definition of projection, $\|\alpha_{\parallel\beta}\|^2 = \frac{|\langle \alpha, \beta \rangle|^2}{\|\beta\|^2}$. Hence $\|\alpha\|^2 \geq \frac{|\langle \alpha, \beta \rangle|^2}{\|\beta\|^2}$ with equality iff α and β are collinear. This is the Cauchy-Schwarz inequality. \square*



The Cauchy-Schwarz inequality

Every finite-dimensional vector space has a basis. If $\beta_1, \beta_2, \dots, \beta_n$ is a basis for the inner product space \mathcal{V} and $\alpha \in \mathcal{V}$ is an arbitrary vector, then there are scalars a_1, \dots, a_n such that $\alpha = \sum a_i \beta_i$ but finding them may be difficult. However, finding the coefficients of a vector is particularly easy when the basis is orthonormal.

A basis $\varphi_1, \varphi_2, \dots, \varphi_n$ for an inner product space \mathcal{V} is orthonormal if

$$\langle \varphi_i, \varphi_j \rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

Finding the i -th coefficient a_i of an orthonormal expansion $\alpha = \sum a_i \psi_i$ is immediate. It suffices to observe that all but the i th term of $\sum a_i \psi_i$ are orthogonal to ψ_i and that the inner product of the i th term with ψ_i yields a_i . Hence if $\alpha = \sum a_i \psi_i$ then

$$a_i = \langle \alpha, \psi_i \rangle.$$

Observe that a_i is the norm of the projection of α on ψ_i . This should not be surprising given that the i th term of the orthonormal expansion of α is collinear to ψ_i and the sum of all the other terms are orthogonal to ψ_i .

There is another major advantage of working with an orthonormal basis. If \mathbf{a} and \mathbf{b} are the n -tuples of coefficients of the expansion of α and β with respect to the same orthonormal basis then

$$\langle \alpha, \beta \rangle = \langle \mathbf{a}, \mathbf{b} \rangle$$

where the right hand side inner product is with respect to the standard inner product. Indeed

$$\begin{aligned} \langle \alpha, \beta \rangle &= \left\langle \sum a_i \psi_i, \sum b_j \psi_j \right\rangle \\ &= \sum a_i \langle \psi_i, \sum b_j \psi_j \rangle \\ &= \sum a_i \langle \psi_i, b_i \varphi_i \rangle \\ &= \sum a_i b_i^* \\ &= \langle \mathbf{a}, \mathbf{b} \rangle. \end{aligned}$$

Letting $\beta = \alpha$ the above implies also

$$\|\alpha\| = \|\mathbf{a}\|,$$

where the right hand side is the standard norm $\|\mathbf{a}\| = \sum |a_i|^2$.

An orthonormal set of vectors ψ_1, \dots, ψ_n of an inner product space \mathcal{V} is a linearly independent set. Indeed $0 = \sum a_i \psi_i$ implies $a_i = \langle 0, \psi_i \rangle = 0$. By normalizing the vectors and recomputing the coefficients one can easily extend this reasoning to a set of orthogonal (but not necessarily orthonormal) vectors $\alpha_1, \dots, \alpha_n$. They too must be linearly independent.

The idea of a projection on a vector generalizes to a projection on a subspace. If \mathcal{W} is a subspace of an inner product space \mathcal{V} , and $\alpha \in \mathcal{V}$, the projection of α on \mathcal{W} is defined to be a vector $\alpha|_{\mathcal{W}} \in \mathcal{W}$ such that $\alpha - \alpha|_{\mathcal{W}}$ is orthogonal to all vectors in \mathcal{W} . If ψ_1, \dots, ψ_m is an orthonormal basis for \mathcal{W} then the condition that $\alpha - \alpha|_{\mathcal{W}}$ is orthogonal to all vectors of \mathcal{W} implies $0 = \langle \alpha - \alpha|_{\mathcal{W}}, \psi_i \rangle = \langle \alpha, \psi_i \rangle - \langle \alpha|_{\mathcal{W}}, \psi_i \rangle$. This shows that $\langle \alpha, \psi_i \rangle = \langle \alpha|_{\mathcal{W}}, \psi_i \rangle$. The right side of this equality is the i -th coefficient of the orthonormal expansion of $\alpha|_{\mathcal{W}}$ with respect to the orthonormal basis. This proves that

$$\alpha|_{\mathcal{W}} = \sum_{i=1}^m \langle \alpha, \psi_i \rangle \psi_i$$

is the unique projection of α on \mathcal{W} .

THEOREM 36. *Let \mathcal{V} be an inner product space and let β_1, \dots, β_n be any collection of linearly independent vectors in \mathcal{V} . Then one may construct orthogonal vectors $\alpha_1, \dots, \alpha_n$ in \mathcal{V} such that they form a basis for the subspace spanned by β_1, \dots, β_n .*

Proof. The proof is constructive via a procedure known as the Gram-Schmidt orthogonalization procedure. First let $\alpha_1 = \beta_1$. The other vectors are constructed inductively as follows. Suppose $\alpha_1, \dots, \alpha_m$ have been chosen so that they form an orthogonal basis for the subspace \mathcal{W}_m spanned by β_1, \dots, β_m . We choose the next vector as

$$\alpha_{m+1} = \beta_{m+1} - \beta_{m+1}|_{\mathcal{W}_m}, \quad (2.59)$$

where $\beta_{m+1}|_{\mathcal{W}_m}$ is the projection of β_{m+1} on \mathcal{W}_m . By definition, α_{m+1} is orthogonal to every vector in \mathcal{W}_m , including $\alpha_1, \dots, \alpha_m$. Also, $\alpha_{m+1} \neq 0$ for otherwise β_{m+1} contradicts the hypothesis that it is linearly independent of β_1, \dots, β_m . Therefore $\alpha_1, \dots, \alpha_{m+1}$ is an orthogonal collection of nonzero vectors in the subspace \mathcal{W}_{m+1} spanned by $\beta_1, \dots, \beta_{m+1}$. Therefore it must be a basis for \mathcal{W}_{m+1} . Thus the vectors $\alpha_1, \dots, \alpha_n$ may be constructed one after the other according to (2.59). \square

COROLLARY 37. *Every finite-dimensional vector space has an orthonormal basis.*

Proof. Let β_1, \dots, β_n be a basis for the finite-dimensional inner product space \mathcal{V} . Apply the Gram-Schmidt procedure to find an orthogonal basis $\alpha_1, \dots, \alpha_n$. Then ψ_1, \dots, ψ_n , where $\psi_i = \frac{\alpha_i}{\|\alpha_i\|}$, is an orthonormal basis. \square

Gram-Schmidt Orthonormalization Procedure

We summarize the Gram-Schmidt procedure, modified so as to produce orthonormal vectors. If β_1, \dots, β_n is a linearly independent collection of vectors in the inner product space \mathcal{V} then we may construct a collection ψ_1, \dots, ψ_n that forms an orthonormal basis for the subspace spanned by β_1, \dots, β_n as follows: we let $\psi_1 = \frac{\beta_1}{\|\beta_1\|}$ and for $i = 2, \dots, n$ we choose

$$\alpha_i = \beta_i - \sum_{j=1}^{i-1} \langle \beta_i, \psi_j \rangle \psi_j$$

$$\psi_i = \frac{\alpha_i}{\|\alpha_i\|}.$$

The following table gives an example of the Gram-Schmidt procedure.

i	β_i	$\langle \beta_i, \psi_j \rangle$ $j < i$	$\beta_i _{\mathcal{W}_{i-1}}$	$\alpha_i = \beta_i - \beta_i _{\mathcal{W}_{i-1}}$	$\ \alpha_i\ $	ψ_i	β_i
1		-	-		2		$\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$
2		1			1		$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$
3		0, 1			4		$\begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}$

Table 2.1: Application of the Gram-Schmidt orthonormalization procedure. Axes are marked with unit length intervals.

Chapter 3

Communication Across the Waveform AWGN Channel

3.1 Introduction

In the previous chapter we have learned how to communicate across the Vector AWGN (Additive White Gaussian Noise) channel. Given a transmitter for that channel, we now know what a receiver that minimizes the error probability should do and how to evaluate or bound the resulting error probability. In this chapter we will deal with a channel model which is closer to reality, namely the Waveform AWGN channel. Apart from the channel model, the main objectives of this and the previous chapters are the same: understand what a receiver should do to minimize the error probability and learn techniques to evaluate the receiver performance. We will also learn that the transmitter and the receiver for the waveform channel may be obtained as natural extensions from the transmitter and receiver for the vector channel studied in the previous chapter. The extension of the transmitter is the Waveform Generator and that of the receiver is the Baseband Front-End, both shown in Figure 1.2. No new technique will be needed to evaluate the error probability.

The starting point for this chapter is the system model shown in Figure 3.1. As usual, we assume that the channel is given (which means that we know the power spectral density N_0 of the white Gaussian noise), and that we have to design the Transmitter (TX) and the Receiver (RX).

The operation of the Waveform Transmitter is similar to that of the Vector Transmitter of the previous chapter except that the output $S(t)$ is now an element of a set of m finite-energy waveforms

$$S(t) \in \{s_0(t), \dots, s_{m-1}(t)\} \subset \mathcal{L}_2.$$

The task of the receiver is to implement a ML decision rule for the following hypothesis

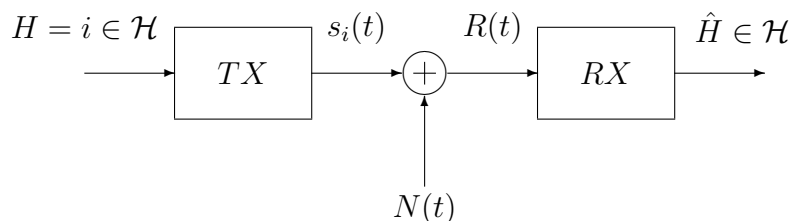


Figure 3.1: Communication across the AWGN channel.

testing problem:

$$H = i : \quad R(t) = s_i(t) + N(t),$$

where $N(t)$ is a zero-mean white Gaussian noise process with (two-sided) power spectral density $N_0/2$. (N_0 is the one-sided power spectral density, namely what one would measure with an instrument.) The source picks the index i according to some probability $P_H(i)$ (typically uniform) over the message set \mathcal{H} .

We will see that, without loss of generality, we may (and should) think of the transmitter as consisting of a part that maps the message H into an n -tuple, as in the previous chapter, followed by a *waveform generator* that maps the n -tuple into a waveform. Similarly, we will see that the receiver may consist of a *front-end* that takes the channel output and produces an n -tuple that is a sufficient statistic. From the waveform generator input to the receiver front-end output, we see a *vector channel* of the kind studied in the previous chapter. Hence, we know already what an optimal receiver should do with the sufficient statistic produced by the receiver front end.

In this chapter we assume familiarity with the linear space \mathcal{L}_2 of finite energy functions and with the concept of white Gaussian noise (WGN) process. Throughout the chapter we will assume that the set $\{s_i(t) : i \in \mathcal{H}\}$ is given. The problem of choosing this set conveniently will be studied in subsequent chapters.

3.2 The Binary Equiprobable Case

We start with the binary hypothesis case since it allows us to focus on the essential. Generalizing to M hypothesis will be straightforward.

There are two hypotheses. When $H = i$, $i \in \mathcal{H} = \{0, 1\}$, we send the waveform $s_i(t)$. To avoid distractions, we assume $P_H(1) = 1/2$. The receiver observes $R(t) = s_i(t) + N(t)$, where $N(t)$ is white Gaussian noise of constant power spectral density $N_0/2$. We are interested in the receiver that minimizes the probability of error.

3.2.1 Sufficient Statistics via Projections

The strategy is to reduce the new (waveform) hypothesis testing problem to the familiar problem where we observe n -tuples.

The key idea is that waveforms of an inner product space can be represented by n -tuples. Which inner product space should we work with? We would like to use the smallest possible one, i.e., the one spanned by s_0 and s_1 . Let us call this space \mathcal{W} . There is a potential problem though: the noise is not in \mathcal{W} , hence $R = s_i + N$ is not in \mathcal{W} either.

If we project the received waveform onto \mathcal{W} , we obtain a waveform $Y = R_{|\mathcal{W}}$ which consists of R minus the portion of the noise which is orthogonal to \mathcal{W} (see Figure 3.2). The intuition is that, in so doing, we remove just noise from the received waveform. More formally, a MAP decision rule based on $R_{|\mathcal{W}}$ results in the same probability of error as one based on R iff $R_{|\mathcal{W}}$ is a *sufficient statistic* or, equivalently, if the portion N_{\perp} of the noise being removed is *irrelevant*. This is the case if, conditioned on $R_{|\mathcal{W}}$ and H , the pdf of N_{\perp} does not depend on H . But this is indeed the case since any finite collection of samples from N_{\perp} is independent of both H and $R_{|\mathcal{W}}$.

Based on the above argument, we consider $R_{|\mathcal{W}}$ as the observable. To be consistent with our notation we will use Y for $R_{|\mathcal{W}}$ and Z for $N_{|\mathcal{W}}$.

We can now restate our hypothesis testing problem:

$$H = i : \quad Y = s_i + Z.$$

We hope that this is a progress since Y , s_i , and Z are all in \mathcal{W} . After choosing a basis $\{\psi_1, \psi_2\}$ for \mathcal{W} , e.g. via the Gram Schmidt procedure on s_0 and s_1 , the corresponding n -tuples \mathbf{Y} , \mathbf{s}_i , and \mathbf{Z} are well defined. Specifically:

$$\begin{aligned} \mathbf{Y} &= (Y_1, Y_2)^T \text{ where} \\ Y_i &= \langle R, \psi_i \rangle, \quad i = 1, 2, \end{aligned}$$

$$\begin{aligned} \mathbf{s}_k &= (s_{k1}, s_{k2})^T \text{ where} \\ s_{ki} &= \langle s_k, \psi_i \rangle, \quad i = 1, 2, \end{aligned}$$

and

$$\begin{aligned} \mathbf{Z} &= (Z_1, Z_2)^T \text{ where} \\ Z_i &= \langle N, \psi_i \rangle, \quad i = 1, 2. \end{aligned}$$

Hereafter we will use the following convention. We use lowercase fonts for deterministic vectors in \mathcal{L}_2 such as s_i . For random vectors in \mathcal{L}_2 we use capital letters, such as S . The corresponding n tuples will be denoted with bold letters such as \mathbf{s}_i and \mathbf{S} .

The hypothesis testing problem based on \mathbf{Y} is now the familiar

$$H = i : \quad \mathbf{Y} = \mathbf{s}_i + \mathbf{Z} \quad \mathbf{Z} \sim \mathcal{N}\left(0, \frac{N_0}{2} I_2\right).$$

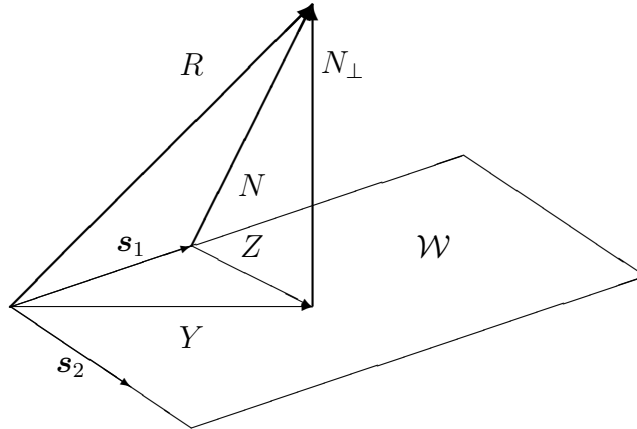


Figure 3.2: Projection of R onto \mathcal{W} . It is assumed that $H = i$. Thinner vectors are in \mathcal{W}

3.2.2 Optimal Test

The test that minimizes the error probability is the ML decision rule:

$$\begin{aligned} \hat{H} &= 1 \\ \|\mathbf{y} - \mathbf{s}_0\|^2 &\geq \|\mathbf{y} - \mathbf{s}_1\|^2 \\ \hat{H} &= 0 \end{aligned}$$

As usual, ties may be resolved either way.

3.2.3 Receiver Structures

In this section we deal with receiver structures. There are various ways to implement the receiver since :

- (a) the ML test can be rewritten in various ways.
- (b) there are two basic ways to implement a projection;

Hereafter is a list of equivalent ML tests. Each of them either suggests a viable implementation or provides valuable insight. The tests are:

$$\begin{aligned} \hat{H} &= 1 \\ \|\mathbf{y} - \mathbf{s}_0\| &\begin{array}{l} \geq \\ < \end{array} \|\mathbf{y} - \mathbf{s}_1\| \quad (\text{T1}) \end{aligned}$$

$$\begin{aligned} \hat{H} &= 0 \\ \hat{H} &= 1 \\ \langle \mathbf{y}, \mathbf{s}_1 \rangle - \frac{\|\mathbf{s}_1\|^2}{2} &\begin{array}{l} \geq \\ < \end{array} \langle \mathbf{y}, \mathbf{s}_0 \rangle - \frac{\|\mathbf{s}_0\|^2}{2} \quad (\text{T2}) \end{aligned}$$

$$\begin{aligned} \hat{H} &= 0 \\ \hat{H} &= 1 \\ \langle R, \mathbf{s}_1 \rangle - \frac{\|\mathbf{s}_1\|^2}{2} &\begin{array}{l} \geq \\ < \end{array} \langle R, \mathbf{s}_0 \rangle - \frac{\|\mathbf{s}_0\|^2}{2} \quad (\text{T3}) \\ \hat{H} &= 0 \end{aligned}$$

Test (T1) is the test derived in the previous section after taking the square root on both sides. Since the square root of a nonnegative number is a monotonic operation, the test outcome remains unchanged. Test (T1) is useful to visualize decoding regions and to compute the probability of error. It says that the decoding region of \mathbf{s}_0 is the set of \mathbf{y} that are closer to \mathbf{s}_0 than to \mathbf{s}_1 .

Figure 3.3 shows the block diagram of a receiver inspired by (T1). The receiver front-end maps R into $\mathbf{Y} = (Y_1, Y_2)$. This part of the receiver deals with waveforms and in the past it has been implemented via analog circuitry. The slicer implements the test (T1). We will refer often to the slicer. It is a conceptual device that knows the decoding regions and checks which decoding region contains \mathbf{y} . The slicer shown in the Figure 3.3 assumes *antipodal signals*, i.e., $\mathbf{s}_0 = -\mathbf{s}_1$, and $\psi_1 = \mathbf{s}_1/\|\mathbf{s}_1\|$. In this case the signal space is one-dimensional and Y_2 is irrelevant.

A slicer for a 2-dimensional signal space spanned by orthogonal signals \mathbf{s}_0 and \mathbf{s}_1 is shown in Figure 3.4, where we defined $\psi_1 = \mathbf{s}_0/\|\mathbf{s}_0\|$ and $\psi_2 = \mathbf{s}_1/\|\mathbf{s}_1\|$.

Once we have a description of the decoding regions, it is conceptually easy to compute the probability of error under each hypothesis. For instance, the probability of error given that $H = 0$ is the probability that an iid Gaussian random vector \mathbf{Z} of variance $\frac{N_0}{2}$ in each component centered at \mathbf{s}_0 ends up in the decoding region of \mathbf{s}_1 . This is the integral over the decoding region of \mathbf{s}_1 of a Gaussian p.d.f. centered at \mathbf{s}_0 .

Depending on the shape of the decoding region of \mathbf{s}_1 , carrying out the integration may or may not be easy. It is easy in general when the decoding region is bounded by perpendicular half-planes as in Figure 3.3 and 3.4. In this case the result can simply be expressed in terms of Q -functions. Explicit examples will be given later.

Perhaps the biggest advantage of test (T1) is the geometrical insight it gives as shown by the slicer. It is, however, not the most economical test in terms of number of steps

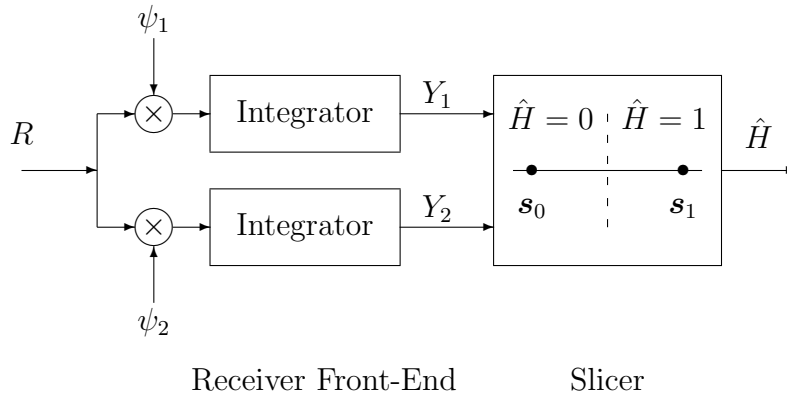


Figure 3.3: Implementation of test (T1). The front-end is based on correlators.

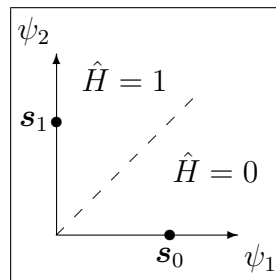


Figure 3.4: Slicer for two orthogonal signals

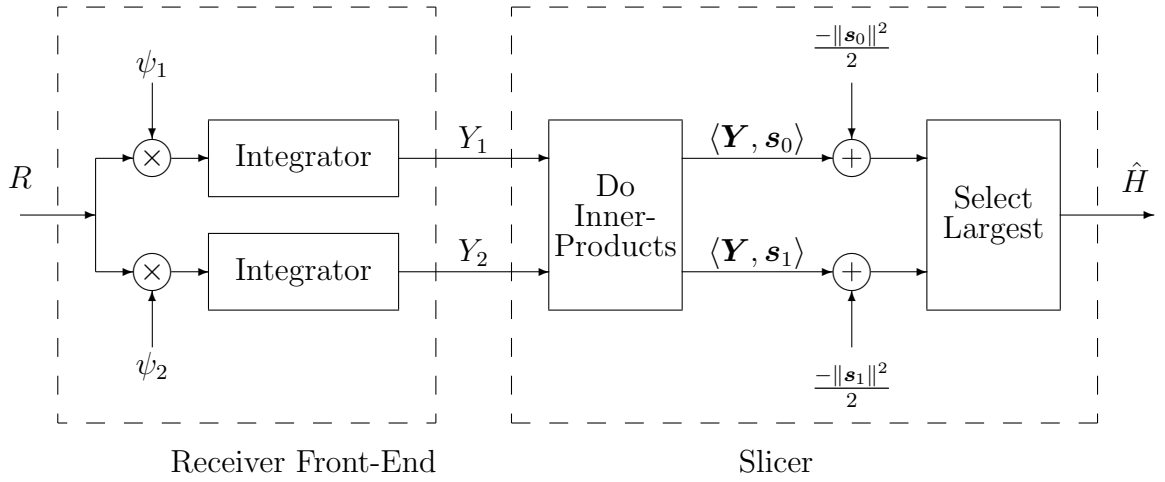


Figure 3.5: Receiver implementation following (T2)

needed to implement it verbatim.

Test (T2) is obtained from (T1) using the relationship

$$\begin{aligned} \|\mathbf{y} - \mathbf{s}_i\|^2 &= \langle \mathbf{y} - \mathbf{s}_i, \mathbf{y} - \mathbf{s}_i \rangle \\ &= \|\mathbf{y}\|^2 - 2\text{Re}\{\langle \mathbf{y}, \mathbf{s}_i \rangle\} + \|\mathbf{s}_i\|^2, \end{aligned}$$

after canceling out common terms, multiplying each side by $-1/2$, and using the fact that $a > b$ iff $-a < -b$. Test (T2) is implemented by the block diagram of Figure 3.5.

The added value of the slicer in Figure 3.5 is that its operation is completely specified in terms of easy-to-implement operations. However, the slicer in Figure 3.3 gives more geometrical insight.

Test (T3) is obtained from (T2) via Parseval's relationship and a bit more to account for the fact that projecting R onto \mathbf{s}_i is the same as projecting Y . Specifically, for $i = 1, 2$,

$$\begin{aligned} \langle \mathbf{y}, \mathbf{s}_i \rangle &= \langle Y, \mathbf{s}_i \rangle \\ &= \langle Y + N_{\perp}, \mathbf{s}_i \rangle \\ &= \langle R, \mathbf{s}_i \rangle. \end{aligned}$$

Test (T3) is implemented by the block diagram in Figure 3.6.

Even though test (T2) and (T3) look similar, they differ fundamentally and practically. First of all (T3) does not require finding a basis for the signal space spanned by $\mathbf{s}_i, i = 1, 2$. As a side benefit this proves that the receiver performance does not depend on the basis used to perform (T2) (or (T1) for that matter).

Second, Test (T2) requires an extra layer of computation, namely that needed to perform the inner products $\langle \mathbf{y}, \mathbf{s}_i \rangle$. This step comes for free in (T3).

However, the number of integrators needed in Figure 3.6 equals the number m of hypotheses (2 in our case), whereas that in Figure 3.5 equals to dimensionality n of the

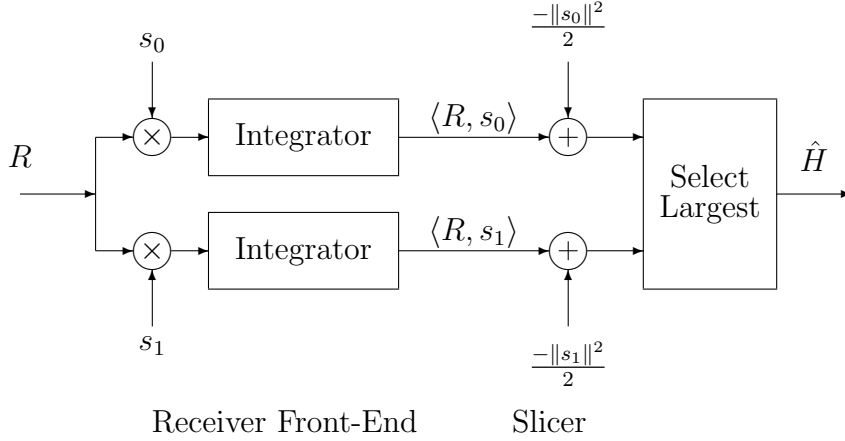


Figure 3.6: Receiver implementation following (T3)

signal space W . We know that $n \leq m$ and one can easily construct examples where equality holds or where $n \ll m$. In the latter case it is preferable to implement test (T2).

Each of the tests (T1), (T2), and (T3) can be implemented in two ways. One way is shown in Figs. 3.3, 3.5 and 3.6, respectively. The other way makes use of the fact that a projection

$$\langle R, s \rangle = \int R(t) s^*(t) dt$$

can always be implemented by means of a filter of impulse response $h(t) = s^*(T - t)$ as shown in Figure 3.7 (b), where T is an arbitrary delay selected in such a way as to make h a causal impulse response.

To verify that the implementation of Figure (3.7)(b) also leads to $\langle R, s \rangle$, we proceed as follows. Let y be the filter output when the input is R . If $h(t) = s^*(T - t)$, $t \in \mathbb{R}$, is the filter impulse response, then

$$y(t) = \int R(\alpha) h(t - \alpha) d\alpha = \int R(\alpha) s^*(T + \alpha - t) d\alpha.$$

At $t = T$ the output is

$$y(T) = \int R(\alpha) s^*(\alpha) d\alpha,$$

which is indeed $\langle R, s \rangle$ (by definition). The implementation of Figure 3.7(b) is referred to as *matched-filter implementation* of the receiver front-end.

In each of the receiver front ends shown in Figs. 3.3, 3.5 and 3.6, we can substitute matched filters for correlators.

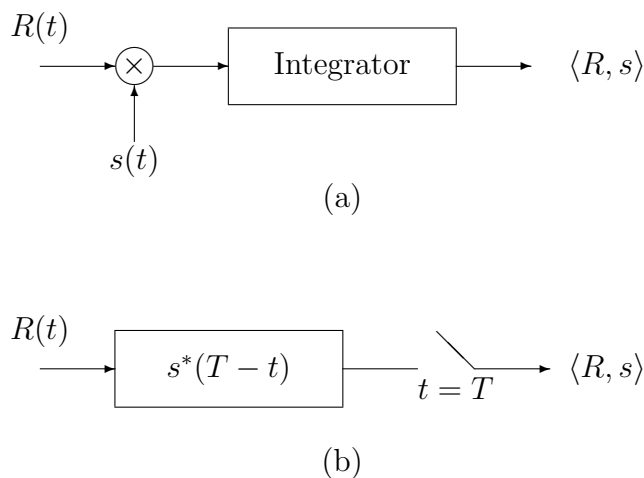


Figure 3.7: Two ways to implement the projection $\langle R, s \rangle$, namely via a “correlator” (a) and via a “matched filter” (b).

3.2.4 Probability of Error

Computing the probability of error is straightforward when we have only two hypotheses. From test (T1) we see that when $H = 0$ we make an error if \mathbf{Y} is closer to \mathbf{s}_1 than to \mathbf{s}_0 . This happens if the projection of the noise N in direction $\mathbf{s}_1 - \mathbf{s}_0$ exceeds $\frac{\|\mathbf{s}_1 - \mathbf{s}_0\|}{2}$. This event has probability.

$$P_e(0) = Q\left(\frac{\|\mathbf{s}_1 - \mathbf{s}_0\|}{2\sigma}\right)$$

where $\sigma^2 = \frac{N_0}{2}$ is the variance of the projection of the noise in any direction.

By symmetry, $P_e(1) = P_e(0)$. Hence

$$P_e = \frac{1}{2}P_e(1) + \frac{1}{2}P_e(0) = Q\left(\frac{\|\mathbf{s}_1 - \mathbf{s}_0\|}{\sqrt{2N_0}}\right),$$

where

$$\|\mathbf{s}_1 - \mathbf{s}_0\| = \|s_1 - s_0\| = \sqrt{\int [s_1(t) - s_0(t)]^2 dt}.$$

It is interesting to observe that the probability of error depends only on the distance $\|\mathbf{s}_1 - \mathbf{s}_0\|$ and not on the particular shape of the waveforms s_0 and s_1 .

In the following example we represent a rectangular pulse by an indicator function.

EXAMPLE 38. Consider the following signal choices and verify that, in all cases, the signal space representation is $\mathbf{s}_0 = (\sqrt{\mathcal{E}}, 0)^T$ and $\mathbf{s}_1 = (0, \sqrt{\mathcal{E}})^T$. To reach this conclusion, it is

enough to verify that $\langle \mathbf{s}_i, \mathbf{s}_j \rangle = \mathcal{E} \delta_{ij}$, where δ_{ij} equals 1 if $i = j$ and 0 otherwise. This means that, in each case, \mathbf{s}_0 and \mathbf{s}_2 are orthogonal.

Choice 1 (Rectangular Pulse Position Modulation) :

$$s_0(t) = \sqrt{\frac{\mathcal{E}}{T}} 1_{[0,T]}(t)$$

$$s_1(t) = \sqrt{\frac{\mathcal{E}}{T}} 1_{[T,2T]}(t).$$

Rectangular pulses can easily be generated, e.g. by a switch. They are used to communicate binary symbols within a circuit. A drawback of rectangular pulses is that they have infinite support in the frequency domain.

Choice 2 (Frequency Shift Keying):

$$s_0(t) = \sqrt{\frac{2\mathcal{E}}{T}} \sin\left(\pi k \frac{t}{T}\right) 1_{[0,T]}(t)$$

$$s_1(t) = \sqrt{\frac{2\mathcal{E}}{T}} \sin\left(\pi l \frac{t}{T}\right) 1_{[0,T]}(t),$$

where k and l are positive integers, $k \neq l$. With a large value of k and l , these signals could be used for wireless communication. As they are they also have infinite support in the frequency domain. Using the trigonometric identity $\sin(\alpha) \sin(\beta) = \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)]$, it is straightforward to verify that the signals are orthogonal.

Choice 3 (Sinc Pulse Position Modulation):

$$s_0(t) = \sqrt{\frac{\mathcal{E}}{T}} \operatorname{sinc}\left(\frac{t}{T}\right)$$

$$s_1(t) = \sqrt{\frac{\mathcal{E}}{T}} \operatorname{sinc}\left(\frac{t - T}{T}\right)$$

The biggest advantage of sinc pulses is that they have finite support in the frequency domain. This means that they have infinite support in the time domain. In practice one uses a truncated version of the time domain signal.

Choice 4 (Spread Spectrum):

$$s_0(t) = \sqrt{\frac{1}{T}} \sum_{j=1}^n s_{0j} 1_{[0, \frac{T}{n}]} \left(t - j \frac{T}{n}\right)$$

$$s_1(t) = \sqrt{\frac{1}{T}} \sum_{j=1}^n s_{1j} 1_{[0, \frac{T}{n}]} \left(t - j \frac{T}{n}\right)$$

where $\underline{s}_0 = (s_{01}, \dots, s_{0n})^T$ and $\underline{s}_1 = (s_{11}, \dots, s_{1n})^T$ are orthogonal and have square norm \mathcal{E} . This signaling method is called spread spectrum. It uses much bandwidth but it has an inherent robustness with respect to interfering (non-white) signals.

As a function of time, the above signal choices vary from choice to choice quite significantly. Nevertheless, as you should be able to convince yourself quickly, on an AWGN channel they all lead to the same probability of error.

3.3 The m -ary Case And The Vector Channel

The solution to the binary hypothesis testing problem derived thus far easily generalizes to the m -ary hypothesis testing problem that we re-formulate for convenience:

$$H = j : \quad R = s_j + N, \quad j \in \mathcal{H},$$

where $\mathcal{H} = \{1, 2, \dots, m\}$, $\mathcal{S} = \{s_1, s_2, \dots, s_m\}$ is the signal constellation which is assumed to be known to the receiver, N is white Gaussian noise, and $P_H(i)$ is the probability that hypothesis H_j , $j \in \mathcal{H}$, is selected. Here R , s_j , and N are functions of time.

We summarize the steps leading to the optimal receivers, leaving out details since they are handled as in the binary case.

First we assume that we have selected a basis $\{\psi_1, \psi_2, \dots, \psi_n\}$ for the vector space spanned by s_1, s_2, \dots, s_m , denoted $\mathcal{W} = \mathcal{W}\{s_1, s_2, \dots, s_m\}$. Like for the binary case, it will turn out that an optimal receiver can be implemented without going through the step of finding a basis.

At the receiver we obtain a sufficient statistic by projecting the received signal R onto each of the basis vector. The result is:

$$\begin{aligned} \mathbf{Y} &= (Y_1, Y_2, \dots, Y_n)^T \text{ where} \\ Y_i &= \langle R, \psi_i \rangle, \quad i = 1, \dots, n. \end{aligned}$$

The signal $Y = \sum Y_i \psi_i$ differs from the received signal R by the component of the noise which is orthogonal to the signal space \mathcal{W} . Let $\mathbf{s}_i = (s_{i1}, \dots, s_{in})^T$ be the n -tuple whose components are the coefficient of s_i with respect to the selected basis. Then

$$s_i = \sum_{j=1}^n s_{ij} \psi_j.$$

It is instructive to visualize the transmitter and the receiver as shown in Figure 3.8. As indicated in this figure, we may think of the cascade of the waveform generator, waveform channel, and receiver front end as of a *vector*¹ *channel*. The vector channel is a suitable channel model to describe (in mathematical terms) the statistical behavior of the channel output given the input. It is also suitable to derive the receiver that minimizes the probability of error. It is also the channel model of choice in information theory to derive the channel capacity. The channel capacity is the maximal rate at which it is possible to

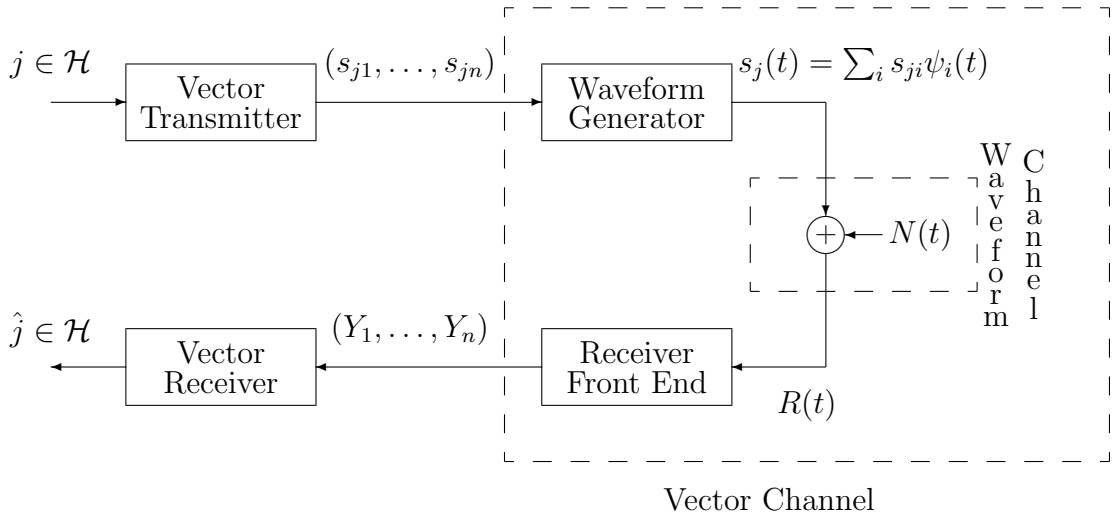


Figure 3.8: The waveform generator and receiver front-end transform the waveform channel into the vector channel.

transmit information in a reliable way across a physical channel by a suitable choice of the transmitter/receiver pair.

The Vector Receiver “sees” the vector hypothesis testing problem

$$H = j : \quad \mathbf{Y} = \mathbf{s}_j + \mathbf{Z} \sim \mathcal{N}(\mathbf{s}_j, \frac{N_0}{2} I_n)$$

studied in Chapter 2.

The receiver observes \mathbf{y} and decides for $\hat{H} = j$ only if

$$P_H(j) f_{\mathbf{Y}}^{(j)}(\mathbf{y}) = \max_k \{P_H(k) f_{\mathbf{Y}}^{(k)}(\mathbf{y})\}$$

where we have introduced the notation $f_{\mathbf{Y}}^{(k)}(\mathbf{y})$ for $f_{\mathbf{Y}|H}(\mathbf{y}|k)$.

Any receiver that satisfies this decision rule minimizes the probability of error. If the maximum is not unique, the receiver may declare any of the hypotheses that achieves the maximum.

For the additive white Gaussian channel under consideration

$$f_{\mathbf{Y}}^{(j)}(\mathbf{y}) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{\|\mathbf{y} - \mathbf{s}_j\|^2}{2\sigma^2}\right)$$

where $\sigma^2 = \frac{N_0}{2}$.

¹More appropriately, this channel should be called *n-tuple channel*.

Plugging into the above decoding rule, taking the log which is a monotonic function, multiplying by minus N_0 , and canceling terms that do not depend on j , we obtain that a MAP decoder decides for one of the $j \in \mathcal{H}$ that minimizes

$$-N_0 \ln P_H(j) + \|\mathbf{y} - \mathbf{s}_j\|^2.$$

The expression should be compared to test (T1) of the previous section. The manipulations of $\|\mathbf{y} - \mathbf{s}_j\|^2$ that have led to test (T2) and (T3) are valid also here. In particular, the equivalent of (T2) consists of maximizing

$$\langle \mathbf{y}, \mathbf{s}_j \rangle + c_j$$

where $c_j = \frac{1}{2}(N_0 \ln P_H(j) - \|\mathbf{s}_j\|^2)$.

Finally, we can use Parseval's relationship to substitute $\langle R, s_j \rangle$ for $\langle \mathbf{Y}, \mathbf{s}_j \rangle$ and get rid of the need to find an orthonormal basis. This leads to the generalization of (T3), namely

$$\langle R, s_j \rangle + c_j.$$

Figure 3.9 shows three MAP receivers where the receiver front end is implemented via a bank of matched filters. Three alternative forms are obtained by using correlators instead of matched filters.

In the first figure, the slicer partitions \mathbb{C}^n into decoding regions. The decoding region for $H = j$ is the set of points $\mathbf{y} \in \mathbb{C}^n$ for which

$$-N_0 \ln P_H(k) + \|\mathbf{y} - \mathbf{s}_k\|^2$$

is minimized when $k = j$.

Notice that in the first two implementations there are n matched filters, where n is the dimension of the signal space \mathcal{W} spanned by \mathcal{S} , whereas in the third implementation the number of matched filters equals the number m of signals in \mathcal{S} . In general, $n \leq m$. Sometimes $n = m$. In this case the third implementation is preferable to the second since it does not require the weighing matrix and does not require finding a basis for \mathcal{W} . Sometimes $n = 1$ whereas m is large. Then the second implementation is preferable since it requires fewer filters.

3.4 Problems

PROBLEM 1. (Matched Filter Implementation.)

In this problem, we consider the implementation of matched filter receivers. In particular, we consider Frequency Shift Keying (FSK) with the following signals:

$$s_j(t) = \begin{cases} \sqrt{\frac{2}{T}} \cos 2\pi \frac{n_j}{T} t, & \text{for } 0 \leq t \leq T, \\ 0, & \text{otherwise,} \end{cases} \quad (3.1)$$

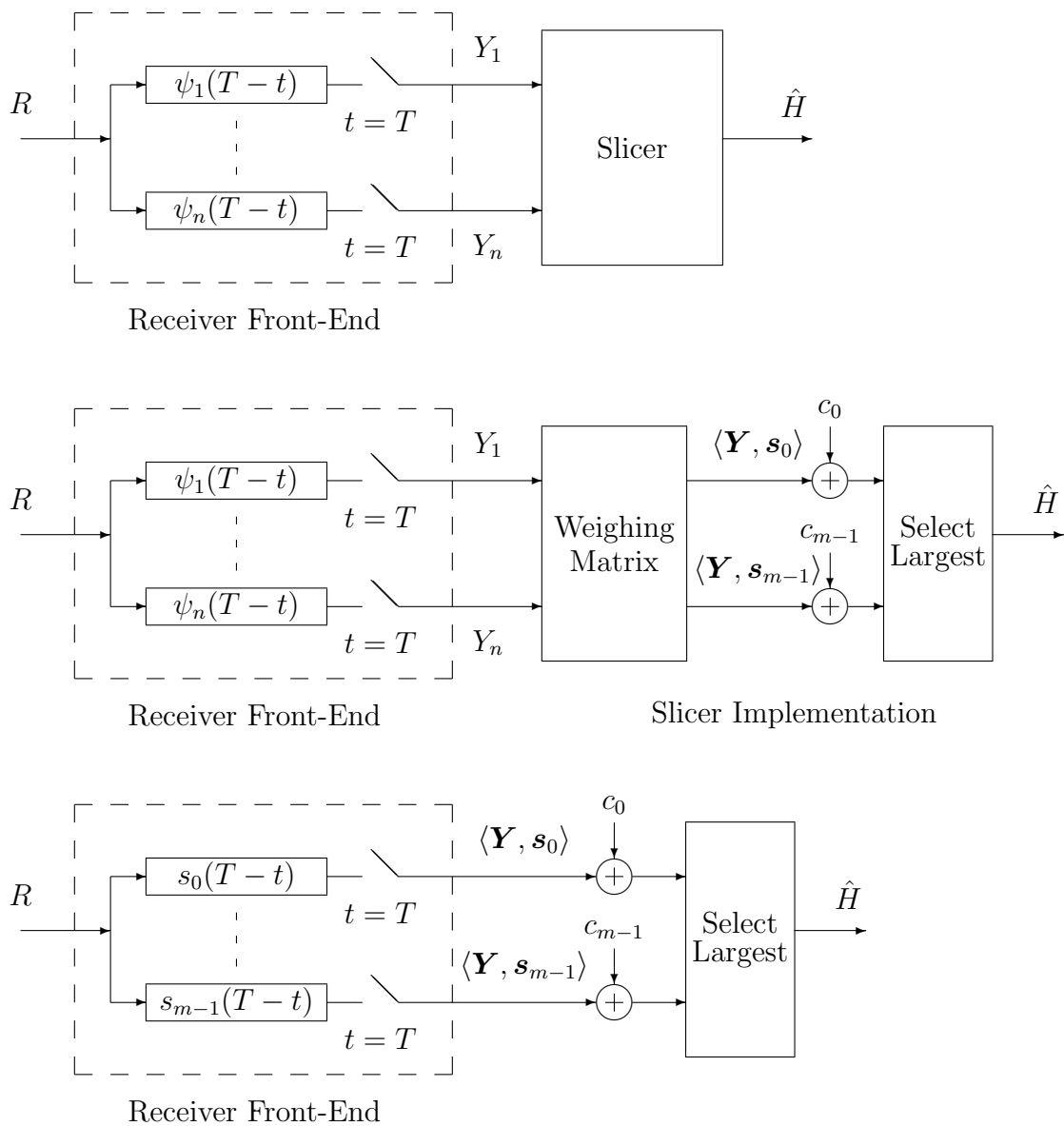
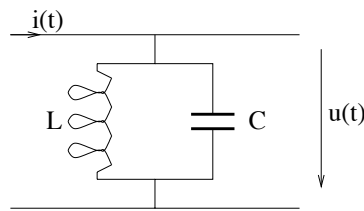


Figure 3.9: Three block diagrams of an optimal receiver. Each receiver front end may alternatively be implemented via correlators.

where $n_j \in \mathbb{Z}$ and $0 \leq j \leq m - 1$. Thus, the communications scheme consists of m signals $s_j(t)$ of different frequencies $\frac{n_j}{T}$

- (i) Determine the impulse response $h_j(t)$ of the matched filter for the signal $s_j(t)$. Plot $h_j(t)$.
- (ii) Sketch the matched filter receiver. How many matched filters are needed?
- (iii) For $-T \leq t \leq 3T$, sketch the output of the matched filter with impulse response $h_j(t)$ when the input is $s_j(t)$. (Hint: We recommend you to use Matlab.)
- (iv) Consider the following ideal resonance circuit:



For this circuit, the voltage response to a unit impulse of current is

$$h(t) = \frac{1}{C} \cos \frac{t}{\sqrt{LC}}. \quad (3.2)$$

Show how this can be used to implement the matched filter for signal $s_j(t)$. Determine how L and C should be chosen. Hint: Suppose that $i(t) = s_j(t)$. In that case, what is $u(t)$?

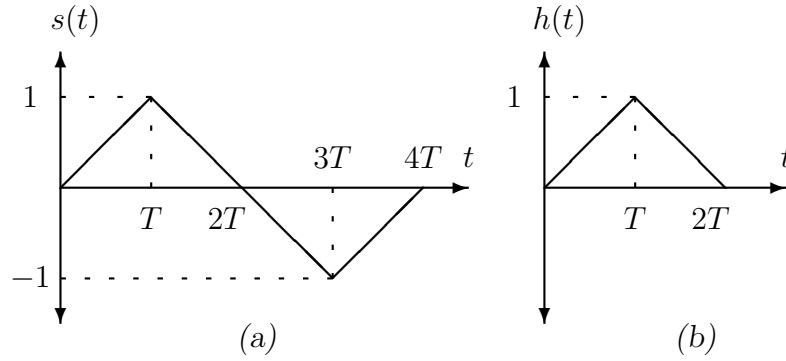
PROBLEM 2. (On-Off Signaling)

Consider the following equiprobable binary hypothesis testing problem specified by:

$$\begin{aligned} H = 0 & : Y(t) = s(t) + N(t) \\ H = 1 & : Y(t) = N(t) \end{aligned}$$

where $N(t)$ is AWGN (Additive White Gaussian Noise) of power spectral density $N_0/2$ and $s(t)$ is the signal shown in the Figure (a) below.

- (a) First consider a receiver that only observes $Y(t_0)$ for some fixed t_0 . Does it make sense to choose \hat{H} based on $Y(t_0)$? Explain.
- (b) Describe the maximum-likelihood receiver for the observable $Y(t)$, $t \in \mathcal{R}$.
- (c) Determine the error probability for the receiver you described in (b).
- (d) Can you realize your receiver of part (b) using a filter with impulse response $h(t)$ shown in Figure (b)?



PROBLEM 3. (Matched Filter Basics)

Consider a communication system that uses antipodal signals $S_i \in \{-1, 1\}$. Using a fixed function $h(t)$, the transmitted waveform $S(t)$ is

$$S(t) = \sum_{k=1}^K S_k h(t - kT).$$

The function $h(t)$ and its shifts by multiples of T form an orthonormal set, i.e.,

$$\int_{-\infty}^{\infty} h(t)h(t - kT)dt = \begin{cases} 0, & k \neq 0 \\ 1, & k = 0. \end{cases}$$

Hint: You don't need Parts (a) and (b) to solve Part (c).

(a) Suppose $S(t)$ is filtered at the receiver by the matched filter with impulse response $h(-t)$. That is, the filtered waveform is $R(t) = \int_{-\infty}^{\infty} S(\tau)h(\tau - t)d\tau$. Show that the samples of this waveform at multiples of T are $R(mT) = S_m$, for $1 \leq m \leq K$.

(b) Now suppose that the channel has an echo in it and behaves like a filter of impulse response $f(t) = \delta(t) + \rho\delta(t - T)$, where ρ is some constant between -1 and 1 . Assume that the transmitted waveform $S(t)$ is filtered by $f(t)$, then filtered at the receiver by $h(-t)$. The resulting waveform $\tilde{R}(t)$ is again sampled at multiples of T . Determine the samples $\tilde{R}(mT)$, for $1 \leq m \leq K$.

(c) Suppose that the k th received sample is $Y_k = S_k + \alpha S_{k-1} + Z_k$, where $Z_k \sim \mathcal{N}(0, \sigma^2)$ and $0 \leq \alpha < 1$ is a constant. S_k and S_{k-1} are independent random variables that take on the values 1 and -1 with equal probability. Suppose that the detector decides $\hat{S}_k = 1$ if $Y_k > 0$, and decides $\hat{S}_k = -1$ otherwise. Find the probability of error for this receiver.

PROBLEM 4. (Matched Filter Intuition.)

In this problem, we develop some further intuition about matched filters.

We have seen that an optimal receiver front end for the signal set $\{s_j(t)\}_{j=0}^{m-1}$ reduces the received (noisy) signal $R(t)$ to the m real numbers $\langle R, s_j \rangle$, $j = 0, \dots, m-1$. We gain additional intuition about the operation $\langle R, s_j \rangle$ by considering

$$R(t) = s(t) + N(t), \quad (3.3)$$

where $N(t)$ is additive white Gaussian noise of power spectral density $N_0/2$ and $s(t)$ is an arbitrary but fixed signal. Let $h(t)$ be an arbitrary waveform, and consider the receiver operation

$$Y = \langle R, h \rangle = \langle s, h \rangle + \langle N, h \rangle. \quad (3.4)$$

The signal-to-noise ratio (SNR) is thus

$$SNR = \frac{|\langle s, h \rangle|^2}{E[|\langle N, h \rangle|^2]}. \quad (3.5)$$

Notice that the SNR is not changed when $h(t)$ is multiplied by a constant. Therefore, we assume that $h(t)$ is a unit energy signal and denote it by $\phi(t)$. Then,

$$E[|\langle N, \phi \rangle|^2] = \frac{N_0}{2}. \quad (3.6)$$

(i) Use Cauchy-Schwarz inequality to give an upper bound on the SNR. What is the condition for equality in the Cauchy-Schwarz inequality? Find the $\phi(t)$ that maximizes the SNR. What is the relationship between the maximizing $\phi(t)$ and the signal $s(t)$?

(ii) To further illustrate this point, take ϕ and s to be two-dimensional vectors and use a picture to discuss why your result in (i) makes sense.

(iii) Take $\phi = (\phi_1, \phi_2)^T$ and $s = (s_1, s_2)^T$ and show how a high school student (without knowing about Cauchy-Schwarz inequality) would have found the matched filter. Hint: You have to maximize $\langle s, \phi \rangle$ subject to the constraint that ϕ has unit energy.

(iv) Hence to maximize the SNR, for each value of t we have to weigh (multiply) $R(t)$ with $s(t)$ and then integrate. Verify with a picture (convolution) that the output at time T of a filter with input $s(t)$ and impulse response $h(t) = s(T-t)$ is indeed $\int_0^T s^2(t) dt$.

(v) We may also look at the situation in terms of Fourier transforms. Write out the filter operation in the frequency domain. Express in terms of $S(f) = \mathcal{F}\{s(t)\}$.

PROBLEM 5. (Optimal receiver for signaling in non-white Gaussian noise.)

We consider the receiver design problem for signals used in non-white additive Gaussian noise. That is, we are given a set of signals $\{s_j(t)\}_{j=0}^{m-1}$ as usual, but the noise added to

those signals is no longer white; rather, it is a Gaussian stochastic process with a given power spectral density

$$S_N(f) = G^2(f), \quad (3.7)$$

where we assume that $G(f) \neq 0$ inside the bandwidth of the signal set $\{s_j(t)\}_{j=0}^{m-1}$. The problem is to design the receiver that minimizes the probability of error.

(i) Find a way to transform the above problem into one that you can solve, and derive the optimum receiver.

(ii) Suppose there is an interval $[f_0, f_0 + \Delta]$ inside the bandwidth of the signal set $\{s_j(t)\}_{j=0}^{m-1}$ for which $G(f) = 0$. What do you do? Describe in words.

PROBLEM 6. (Antipodal signaling in non-white Gaussian noise.)

In this problem, antipodal signaling (i.e. $s_0(t) = -s_1(t)$) is to be used in non-white additive Gaussian noise of power spectral density

$$S_N(f) = G^2(f), \quad (3.8)$$

where we assume that $G(f) \neq 0$ inside the bandwidth of the signal $s(t)$.

How should the signal $s(t)$ be chosen (as a function of $G(f)$) such as to minimize the probability of error? Hint: For ML decoding of antipodal signaling in AWGN (of fixed variance), the $\Pr\{e\}$ depends only on the signal energy.

PROBLEM 7. (Mismatched Receiver.)

Let the received waveform $Y(t)$ be given by

$$Y(t) = cX s(t) + N(t), \quad (3.9)$$

where $c > 0$ is some deterministic constant, X is a random variable that takes on the values $\{3, 1, -1, -3\}$ equiprobably, $s(t)$ is the deterministic waveform

$$s(t) = \begin{cases} 1, & \text{if } 0 \leq t < 1 \\ 0, & \text{otherwise,} \end{cases} \quad (3.10)$$

and $N(t)$ is white Gaussian noise of spectral density $\frac{N_0}{2}$.

(a) Describe the receiver that, based on the received waveform $Y(t)$, decides on the value of X with least probability of error. Be sure to indicate precisely when your decision rule would declare “+3”, “+1”, “−1”, and “−3”.

(b) Find the probability of error of the detector you have found in Part (a).

(c) Suppose now that you still use the detector you have found in Part (a), but that the received waveform is actually

$$Y(t) = \frac{3}{4} c X s(t) + N(t), \quad (3.11)$$

i.e., you were mis-informed about the signal amplitude. What is the probability of error now?

(d) Suppose now that you still use the detector you have found in Part (a) and that $Y(t)$ is according to Equation (3.9), but that the noise is colored. In fact, $N(t)$ is a zero-mean stationary Gaussian noise process of auto-covariance function

$$K_N(\tau) = E[N(t)N(t+\tau)] = \frac{1}{4\alpha} e^{-|\tau|/\alpha}, \quad (3.12)$$

where $0 < \alpha < \infty$ is some deterministic real parameter. What is the probability of error now?

PROBLEM 8. (QAM receiver)

Consider a transmitter which transmits waveforms of the form,

$$s(t) = \begin{cases} s_1 \sqrt{\frac{2}{T}} \cos 2\pi f_c t + s_2 \sqrt{\frac{2}{T}} \sin 2\pi f_c t, & \text{for } 0 \leq t \leq T, \\ 0, & \text{otherwise,} \end{cases} \quad (3.13)$$

where $2f_c T \in \mathbb{Z}$. $(s_1, s_2) \in \{(\sqrt{E}, \sqrt{E}), (-\sqrt{E}, \sqrt{E}), (-\sqrt{E}, -\sqrt{E}), (\sqrt{E}, -\sqrt{E})\}$ with equal probability. The signal received at the receiver is corrupted by AWGN of power spectral density $\frac{N_0}{2}$.

- (a) Specify the receiver for this transmission scheme.
- (b) Draw the decoding regions and find the probability of error.

PROBLEM 9. Consider the following functions $S_0(t)$, $S_1(t)$ and $S_2(t)$.

(Gram-Schmidt for Three Signals)

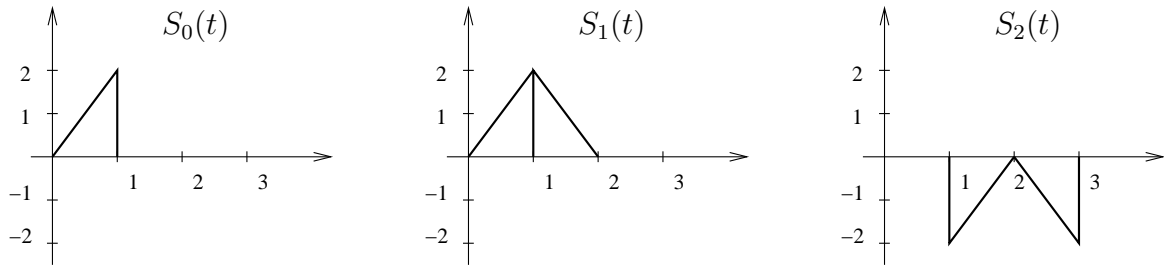
(i) Using the Gram-Schmidt procedure, determine a basis of the space spanned by $\{s_0(t), s_1(t), s_2(t)\}$. Denote the basis functions by $\phi_0(t)$, $\phi_1(t)$ and $\phi_2(t)$.

(ii) Let

$$V_1 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \quad \text{and} \quad V_2 = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$$

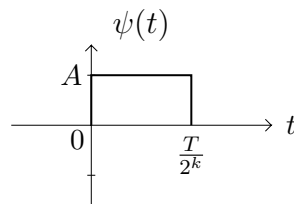
be two points in the space spanned by $\{\phi_0(t), \phi_1(t), \phi_2(t)\}$. What is their corresponding signal, $V_1(t)$ and $V_2(t)$? (You can simply draw a detailed graph.)

(iii) Compute $\int V_1(t)V_2(t)dt$.



PROBLEM 10. Consider the following communication chain. We have 2^k possible hypotheses with $k \in \mathbb{N}$ to convey through a waveform channel. When hypothesis i is selected, the transmitted signal is $s_i(t)$ and the received signal is given by $R(t) = s_i(t) + N(t)$, where $N(t)$ denotes a white Gaussian noise with double-sided power spectral density $\frac{N_0}{2}$. Assume that the transmitter uses the position of a pulse $\psi(t)$ in an interval $[0, T]$, in order to convey the desired hypothesis, i.e., to send hypothesis i , the transmitter sends the signal $\psi_i(t) = \psi(t - \frac{iT}{2^k})$.

(i) If the pulse is given by the waveform $\psi(t)$ depicted below. What is the value of A that gives us signals of energy equal to one as a function of k and T ?



(ii) We want to transmit the hypothesis $i = 3$ followed by the hypothesis $j = 2^k - 1$. Plot the waveform you will see at the output of the transmitter, using the pulse given in the previous question.

(iii) Sketch the optimal receiver.

What is the minimum number of filters you need for the optimal receiver? Explain.

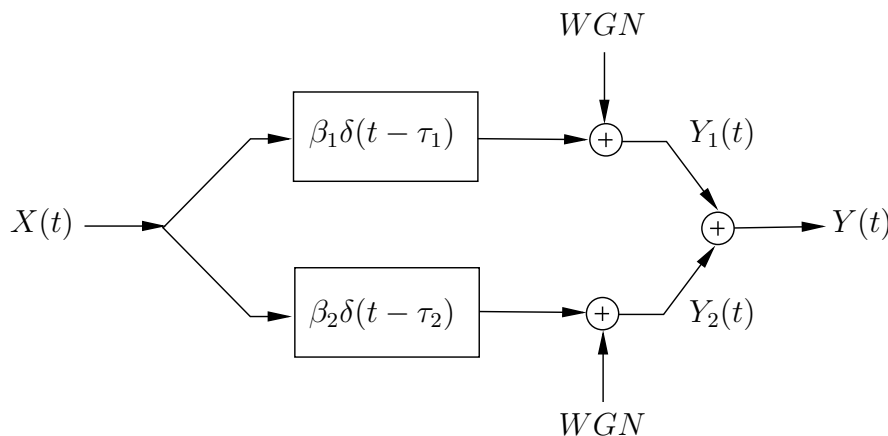
(iv) What is the major drawback of this signaling scheme? Explain.

PROBLEM 11. (Communication Chain with two receive antennas)

Consider the following communication chain, where we have two possible hypotheses H_0 and H_1 . Assume that $P_H(H_0) = P_H(H_1) = \frac{1}{2}$. The transmitter uses antipodal signaling. To transmit H_0 , the transmitter sends a unit energy pulse $p(t)$, and to transmit H_1 , it sends $-p(t)$. That is, the transmitted signal is $X(t) = \pm p(t)$. The observation consists of $Y_1(t)$ and $Y_2(t)$ as shown below. The signal along each "path" is an attenuated and delayed version of the transmitted signal $X(t)$. The noise is additive white Gaussian with double sided power spectral density $N_0/2$. Also, the noise added to the two observations

is independent and independent of the data. The goal of the receiver is to decide which hypothesis was transmitted, based on its observation.

We will look at two different scenarios: either the receiver has access to each individual signal $Y_1(t)$ and $Y_2(t)$, or the receiver has only access to the combined observation $Y(t) = Y_1(t) + Y_2(t)$.



- a. The case where the receiver has only access to the combined output $Y(t)$.
 1. In this case, observe that we can write the received waveform as $\pm g(t) + Z(t)$. What are $g(t)$ and $Z(t)$ and what are the statistical properties of $Z(t)$?
Hint: Recall that $\int \delta(\tau - \tau_1)p(t - \tau)d\tau = p(t - \tau_1)$.
 2. What is the optimal receiver for this case? Your answer can be in the form of a block diagram that shows how to process $Y(t)$ or in the form of equations. In either case, specify how the decision is made between H_0 or H_1 .
 3. Assume that $\int p(t - \tau_1)p(t - \tau_2)dt = \gamma$, where $-1 \leq \gamma \leq 1$. Find the probability of error for this optimal receiver, express it in terms of the Q function, β_1 , β_2 , γ and $N_0/2$.
- b. The case where the receiver has access to the individual observations $Y_1(t)$ and $Y_2(t)$.
 1. Argue that the performance of the optimal receiver for this case can be no worse than that of the optimal receiver for part (a).
 2. Compute the sufficient statistics (Y_1, Y_2) , where $Y_1 = \int Y_1(t)p(t - \tau_1)dt$ and $Y_2 = \int Y_2(t)p(t - \tau_2)dt$. Show that this sufficient statistic (Y_1, Y_2) has the form $(Y_1, Y_2) = (\beta_1 + Z_1, \beta_2 + Z_2)$ under H_0 , and $(-\beta_1 + Z_1, -\beta_2 + Z_2)$ under H_1 , where Z_1 and Z_2 are independent zero-mean Gaussian random variables of variance $N_0/2$.
 3. Using the LLR (Log-Likelihood Ratio), find the optimum decision rule for this case.

Hint: It may help to draw the two hypotheses as points in \mathbb{R}^2 . If we let $V = (V_1, V_2)$ be a Gaussian random vector of mean $m = (m_1, m_2)$ and covariance matrix $\Sigma = \sigma^2 I$, then its pdf is $p_V(v_1, v_2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(v_1-m_1)^2}{2\sigma^2} - \frac{(v_2-m_2)^2}{2\sigma^2}\right)$.

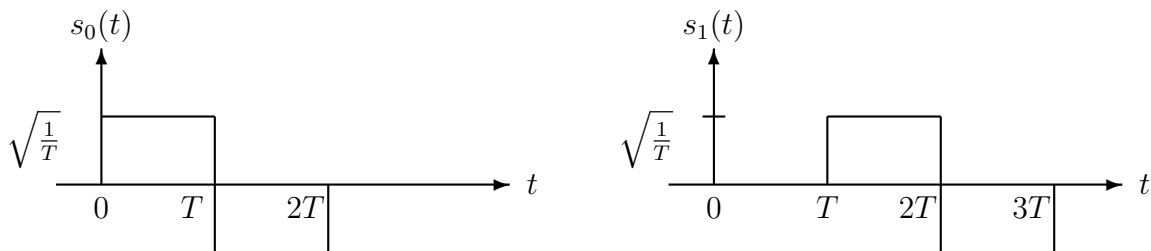
4. What is the optimal receiver for this case? Your answer can be in the form of a block diagram that shows how to process $Y_1(t)$ and $Y_2(t)$ or in the form of equations. In either case, specify how the decision is made between H_0 or H_1 .
5. Find the probability of error for this optimal receiver, express it in terms of the Q function, β_1 , β_2 and N_0 .

c. Comparison of the two cases

1. In the case of $\beta_2 = 0$, that is the second observation is solely noise, give the probability of error for both cases (a) and (b). What is the difference between them? Explain why.

PROBLEM 12. (Delayed Signals)

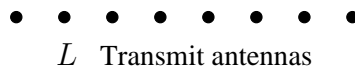
One of two signals shown in the figure below is transmitted over the additive white Gaussian noise channel. There is no bandwidth constraint and either signal is selected with probability $1/2$.



- (a) Draw a block diagram of a maximum likelihood receiver. Be as specific as you can. Try to use the smallest possible number of filters and/or correlators.
- (b) Determine the error probability in terms of the Q -function, assuming that the power spectral density of the noise is $\frac{N_0}{2} = 5 \left[\frac{W}{Hz}\right]$.

PROBLEM 13. (Antenna Array)

Consider an L -element antenna array as shown in the figure below.



Let $u(t)\beta_i$ be the (complex-valued baseband equivalent) signal transmitted at antenna element i , $i = 1, 2, \dots, L$ (according to some indexing which is irrelevant here) and let

$$v(t) = \sum_{i=1}^L u(t - \tau_D)\beta_i\alpha_i$$

(plus noise) be the sum-signal at the receiver antenna, where α_i is the path strength for the signal transmitted at antenna element i and τ_D is the (common) path delay.

- (a) Choose the vector $\beta = (\beta_1, \beta_2, \dots, \beta_L)^T$ that maximizes the signal energy at the receiver, subject to the constraint $\|\beta\| = 1$. The signal energy is defined as $E_v = \int |v(t)|^2 dt$. Hint Use the Cauchy-Schwarz inequality: for any two vectors \mathbf{a} and \mathbf{b} in \mathbb{C}^n , $|\langle \mathbf{a}, \mathbf{b} \rangle|^2 \leq \|\mathbf{a}\|^2 \|\mathbf{b}\|^2$ with equality iff \mathbf{a} and \mathbf{b} are linearly dependent.
- (b) Let $u(t) = \sqrt{E_u}\phi(t)$ where $\phi(t)$ has unit energy. Determine the received signal power as a function of L when β is selected as in (a) and $\alpha = (\alpha, \alpha, \dots, \alpha)^T$ for some complex number α .
- (c) In the above problem the received energy grows monotonically with L while the transmit energy is constant. Does this violate energy conservation or some other fundamental law of physics? Hint: an antenna array is not an isotropic antenna (i.e. an antenna that sends the same energy in all directions).

PROBLEM 14. (Cioffi)

The signal set

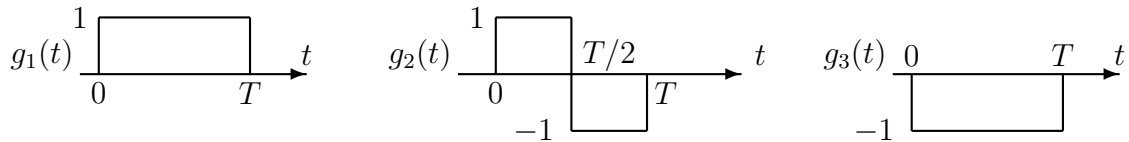
$$\begin{aligned} s_0(t) &= \text{sinc}^2(t) \\ s_1(t) &= \sqrt{2}\text{sinc}^2(t) \cos(4\pi t) \end{aligned}$$

is used to communicate across an AWGN channel of power spectral density $\frac{N_0}{2}$.

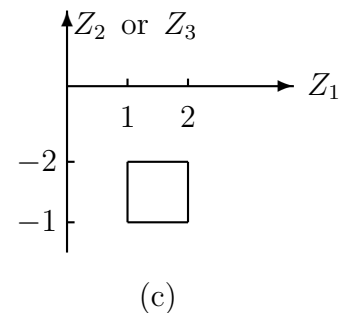
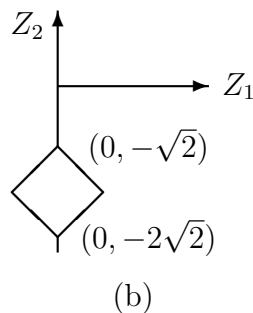
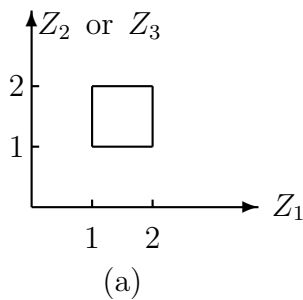
- (a) Find the Fourier transforms of the above signals and plot them.
- (b) Sketch a block diagram of a ML receiver for the above signal set.
- (c) Determine its error probability of your receiver assuming that $s_0(t)$ and $s_1(t)$ are equally likely.
- (d) If you keep the same receiver, but use $s_0(t)$ with probability $\frac{1}{3}$ and $s_1(t)$ with probability $\frac{2}{3}$, does the error probability increase, decrease, or remain the same? Justify your answer.

PROBLEM 15. (Probabilities in Different Signal Spaces)

Let $N(t)$ be a zero-mean white Gaussian process of power spectral density $\frac{N_0}{2}$. Let $g_1(t)$, $g_2(t)$, and $g_3(t)$ be waveforms as shown in the following figure.



- (a) Determine the norm $\|g_i\|$, $i = 1, 2, 3$.
- (b) Let Z_i be the projection of $N(t)$ onto $g_i(t)$. Write down the mathematical expression that describes this projection, i.e. how you obtain Z_i from $N(t)$ and $g_i(t)$.
- (c) Describe the object Z_1 , i.e. tell us everything you can say about it. Be as concise as you can.
- (d) Are Z_1 and Z_2 independent? Justify your answer.
- (e) (i) Describe the object $\mathbf{Z} = (Z_1, Z_2)$. (We are interested in what it is, not on how it is obtained.)
 (ii) Find the probability P_a that \mathbf{Z} lies in the square labeled (a) in the figure below.
 (iii) Find the probability P_b that \mathbf{Z} lies in the square (b) of the same figure. Justify your answer.
- (f) (i) Describe the object $\mathbf{W} = (Z_1, Z_3)$.
 (ii) Find the probability Q_a that \mathbf{W} lies in the square (a).
 (iii) Find the probability Q_c that \mathbf{W} lies in the square (c).



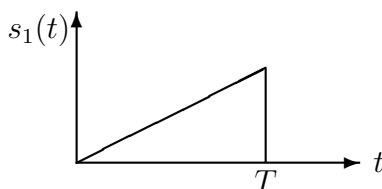
PROBLEM 16. (Gram Schmidt for Two Signals)

- (a) Use Gram Schmidt procedure to find an orthonormal basis for the vector space spanned by the functions shown below. Clearly indicate every step of the procedure. Make sure that \mathbf{s}_1 , \mathbf{s}_2 , and the orthonormal basis are clearly visible.
- (b) Let $s(t) = \beta \text{sinc}(\alpha t)$. Plot $s(t)$ (qualitatively but label your plot appropriately) and determine the area $A = \int_{-\infty}^{\infty} s(t) dt$.



PROBLEM 17. (ML Receiver with a Single Causal Filter)

You want to design a Maximum Likelihood (ML) receiver for a system that communicates an equiprobable binary hypothesis by means of the signals $s_1(t)$ and $s_2(t) = s_1(t - T_d)$, where $s_1(t)$ is shown in the figure and T_d is a fixed number assumed to be known at the receiver.



The channel is the usual AWGN channel with noise power spectral density $N_0/2$. At the receiver front end you are allowed to use a single causal filter of impulse response $h(t)$ (A causal filter is one whose impulse response is 0 for $t < 0$).

- Describe the $h(t)$ that you chose for your receiver.
- Sketch a block diagram of your receiver. Be specific about the sampling times.
- Assuming that $T_d > T$, determine the error probability for the receiver as a function of N_0 and E_s ($E_s = \|s_1(t)\|^2$).

PROBLEM 18. (Signal Space Distance)

Let there be an orthonormal basis $\{\phi_m(t)\}_{m=0}^{N-1}$ for a given space \mathbb{S} of functions. Then, every function $s(t) \in \mathbb{S}$ can be written as

$$s(t) = \sum_{m=0}^{N-1} s_m \phi_m(t), \quad (3.14)$$

with $s_m = \langle s(t), \phi_m(t) \rangle$.

Show that the squared L_2 distance $d_{i,j}^2$ between two waveforms $s_i(t)$ and $s_j(t)$, which is defined as

$$d_{i,j}^2 = \int (s_i(t) - s_j(t))^2 dt, \quad (3.15)$$

is equal to the signal space distance

$$d_{i,j}^2 = \sum_{m=0}^{N-1} (s_{i,m} - s_{j,m})^2. \quad (3.16)$$

Note: In communications, it is often preferable to consider the vector of coefficients $s = (s_0, \dots, s_{N-1})^T$ rather than the waveform $s(t)$ itself. We call s the signal space representation of the waveform $s(t)$.

PROBLEM 19. (Waveform Receiver)

Consider the signals $s_0(t)$ and $s_1(t)$ shown in the figure.

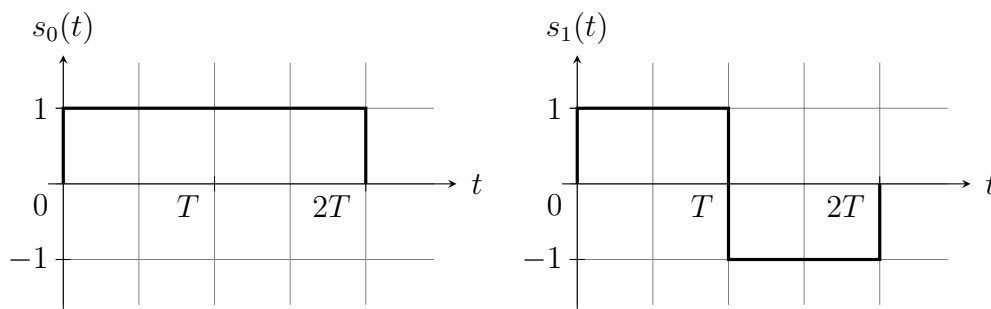


Figure 3.10: Signal waveforms

- (a) Determine an orthonormal basis $\{\psi_0(t), \psi_1(t)\}$ for the space spanned by $\{s_0(t), s_1(t)\}$ and find the n -tuples of coefficients \mathbf{s}_0 and \mathbf{s}_1 that correspond to $s_0(t)$ and $s_1(t)$, respectively.
- (b) Let X be a uniformly distributed binary random variable that takes values in $\{0, 1\}$. We want to communicate the value of X over an additive white Gaussian noise channel. When $X = 0$, we send $S(t) = s_0(t)$, and when $X = 1$, we send $S(t) = s_1(t)$. The received signal at the destination is

$$Y(t) = S(t) + Z(t),$$

where $Z(t)$ is AWGN of power spectral density $\frac{N_0}{2}$.

- (i) On the back of this page, draw an optimal matched filter receiver for this case. Specifically say how the decision is made.
- (ii) What is the output of the matched filter(s) when $X = 0$ and the noise variance is zero ($\frac{N_0}{2} = 0$)?

-
- (iii) Describe the output of the matched filter when $S(t) = 0$ and the noise variance is $\frac{N_0}{2} > 0$.
- (c) Plot the \mathbf{s}_0 and \mathbf{s}_1 that you have found in part (a), and determine the error probability P_e of this scheme as a function of T and N_0 .
- (d) Find a suitable waveform $v(t)$, such that the new signals $\hat{s}_0(t) = s_0(t) - v(t)$ and $\hat{s}_1(t) = s_1(t) - v(t)$ have minimal energy and plot the resulting $\hat{s}_0(t)$ and $\hat{s}_1(t)$. Hint: you may first want to find \mathbf{v} , the n -tuple of coefficients that corresponds to $v(t)$.
- (e) Compare $\hat{s}_0(t)$ and $\hat{s}_1(t)$ to $s_0(t)$ and $s_1(t)$, respectively, and comment on the part $v(t)$ that has been removed.

Appendix 3.A Rectangle and Sinc as Fourier Transform Pairs

The Fourier transform of a rectangular pulses is a sinc pulse. Often one has to go back and forth between such Fourier pairs. The purpose of this appendix is to make it easier to figure out the details.

First of all let us recall that a function g and its Fourier transform $g_{\mathcal{F}}$ are related by

$$g(u) = \int g_{\mathcal{F}}(\alpha) \exp(j2\pi u\alpha) d\alpha$$

$$g_{\mathcal{F}}(v) = \int g(\alpha) \exp(-j2\pi v\alpha) d\alpha.$$

Notice that $g_{\mathcal{F}}(0)$ is the area under g and $g(0)$ is the area under $g_{\mathcal{F}}$.

Next let us recall that $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$ is the function that equals 1 at $x = 0$ and equals 0 at all other integer values of x . Hence if $a, b \in \mathbb{R}$ are arbitrary constants, $a\text{sinc}(bx)$ equals a at $x = 0$ and equals 0 at nonzero multiples of $1/b$.

If you could remember that the area under $a\text{sinc}(bx)$ is a/b then, from the two facts above, you could conclude that its Fourier transform, which you know is a rectangle, has height equals a/b and area a . Hence the width of this rectangle must be b .

It is actually easy to remember that the area under $a\text{sinc}(bx)$ is a/b : it is the area of the triangle described by the main lobe of $a\text{sinc}(bx)$, namely the area of the triangle with coordinates $(-1/b, 0)$, $(0, a)$, $(1/b, 0)$.

Appendix 3.B White Gaussian Noise

We assume that you are familiar with the concept of White Gaussian Noise. The purpose of this appendix is just to write down what you absolutely need to remember, for the purpose of this Chapter, about White Gaussian Noise.

If $N(t)$ is White Gaussian Noise of double-sided spectral density $\frac{N_0}{2}$ then:

- Its covariance function is

$$K_N(\tau) \triangleq \frac{N_0}{2} \delta(\tau), \quad \forall \tau.$$

- Its spectrum (the Fourier transform of the covariance function) is

$$S_N(f) = \frac{N_0}{2}.$$

- If

$$Z_i = \int_{-\infty}^{\infty} N(t) g_i(t) dt, \quad i = 1, \dots, K,$$

then (Z_1, \dots, Z_N) is a zero-mean Gaussian random vector and for any $1 \leq i, j \leq K$,

$$\begin{aligned}
 E[Z_i Z_j] &= E \left[\int_{-\infty}^{\infty} N(t) g_i(t) \int_{-\infty}^{\infty} N(\xi) g_j(\xi) d\xi \right] \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E[N(t) N(\xi)] g_i(t) g_j(\xi) dt d\xi \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{N_0}{2} \delta(t - \xi) g_i(t) g_j(\xi) dt d\xi \\
 &= \int_{-\infty}^{\infty} \frac{N_0}{2} g_i(t) g_j(t) dt
 \end{aligned}$$

In particular, if $g_1(t), \dots, g_k(t)$ are an orthonormal set then Z_1, \dots, Z_K are iid $\mathcal{N}(0, \frac{N_0}{2})$.

Chapter 4

Signal Design Trade-Offs

4.1 Introduction

It is time to shift our focus to the transmitter and take a look at some of the options we have in terms of choosing the signal constellation. The goal is to build up some intuition about the impact that those options have on fundamental performance measures such as transmission rate, bandwidth, power, and error probability. Throughout this section we assume that the channel is the AWGN channel and that the receiver implements a ML decision rule.

To put things into perspective, we mention from the outset that the problem of choosing a convenient signal constellation is not as clean-cut as the receiver design problem that has kept us busy until now. The reason is that the receiver design problem has a clear objective, namely to minimize the error probability, and an essentially unique solution, a MAP decision rule. In contrast, choosing a good signal constellation is making a tradeoff among conflicting objectives. For instance, if we could, we would choose a signal constellation that contains a very large number m of signals of very small duration T and very small bandwidth B . If we could choose these parameters at will, we could also achieve any desired rate $\frac{\log_2 m}{TB}$ (expressed in bits per second per Hz). In addition, we would choose our signals so that they use very little energy and result in a very small error probability. These are conflicting goals.

Besides the quantities already mentioned, other quantities that will come up in our discussion are the number $k = \log_2 m$ of bits associated to each signal, the average time $T_b = T/k$ it takes to transmit one bit, the dimensionality n of the signal space, the energy per bit \mathcal{E}_b , the block error probability P_e and the bit error probability P_b .

4.2 Transformations That Do Not Affect P_e

Two sets of waveforms can look very different yet lead to the same probability of error. In this section we take a look at some of the transformations that change the signal constellation without affecting the error probability. There are at least two obvious reasons why we may want to evoke such a transformation: (i) we may save ourselves some time if we recognize that the probability of error associated to the constellation we are using is the same as that of another constellation for which we have already determined the error probability or for which we know an easy way to determine it; (ii) given a signal constellation that has the desired probability of error, we may be able to transform it into one that has the same probability of error and uses less energy, or less bandwidth, or less time.

4.2.1 Isometric Transformations In \mathbb{R}^n

An isometry in \mathbb{R}^n is a distance-preserving transformation $a : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Hence for any two points $\mathbf{p}, \mathbf{q} \in \mathbb{R}^n$, the distance from \mathbf{p} to \mathbf{q} equals the distance from $a(\mathbf{p})$ to $a(\mathbf{q})$.

If we apply the same isometry to every point of a signal constellation and to every decoding region, the probability of error (for the AWGN channel) remains the same. This intuitive fact can be verified mathematically as follows. Let

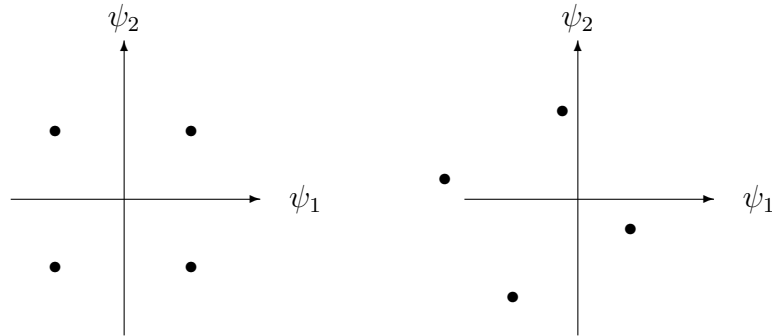
$$g(\gamma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\gamma^2}{2\sigma^2}\right), \quad \gamma \in \mathbb{R}$$

so that for $\mathbf{Z} \sim \mathcal{N}(0, \sigma^2 I_n)$ we can write $f_{\mathbf{Z}}(\mathbf{z}) = g(\|\mathbf{z}\|^2)$. Then for any isometry $a : \mathbb{R}^n \rightarrow \mathbb{R}^n$ we have

$$\begin{aligned} P_c(i) &= Pr\{\mathbf{Y} \in \mathcal{R}_i | \mathbf{S} = \mathbf{s}_i\} \\ &= \int_{\mathbf{y} \in \mathcal{R}_i} g(\|\mathbf{y} - \mathbf{s}_i\|) d\mathbf{y} \\ &\stackrel{(a)}{=} \int_{\mathbf{y} \in \mathcal{R}_i} g(\|a(\mathbf{y}) - a(\mathbf{s}_i)\|) d\mathbf{y} \\ &\stackrel{(b)}{=} \int_{a(\mathbf{y}) \in a(\mathcal{R}_i)} g(\|a(\mathbf{y}) - a(\mathbf{s}_i)\|) d\mathbf{y} \\ &\stackrel{(c)}{=} \int_{\boldsymbol{\alpha} \in a(\mathcal{R}_i)} g(\|\boldsymbol{\alpha} - a(\mathbf{s}_i)\|) d\boldsymbol{\alpha} = Pr\{\mathbf{Y} \in a(\mathcal{R}_i) | \mathbf{S} = a(\mathbf{s}_i)\}, \end{aligned}$$

where in (a) we used the distance preserving property of an isometry, in (b) we used the fact that $y \in \mathcal{R}_i$ iff $a(\mathbf{y}) \in a(\mathcal{R}_i)$, and in (c) we made the change of variable $\boldsymbol{\alpha} = a(\mathbf{y})$ and used the fact that the Jacobian of an isometry is ± 1 . The last line is the probability of decoding correctly when the transmitter sends $a(\mathbf{s}_i)$ and the corresponding decoding region is $a(\mathcal{R}_i)$.

EXAMPLE 39. The composition of a translation and a rotation is an isometry. The figure below shows an original signal set and a translated and rotated copy. Both have the same probability of error but not the same energy.



□

Given $\{\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_{m-1}\}$, we are interested in finding the translating vector \mathbf{a} so that the average energy of $\{\mathbf{s}'_0, \mathbf{s}'_1, \dots, \mathbf{s}'_{m-1}\}$, where $\mathbf{s}'_i = \mathbf{s}_i - \mathbf{a}$, is minimized. The appropriate choice of \mathbf{a} is (see Problem 1)

$$\mathbf{a} = \sum_i P_H(i) \mathbf{s}_i.$$

From now on we will use \mathcal{E} to denote the average energy of the signal constellation at hand. Sometimes we will use \mathcal{E}_b to denote the average energy per bit. Hence for a signal set $\{s_0, s_1, \dots, s_{m-1}\}$, where signal s_i is used with probability $P_H(i)$, we have

$$\mathcal{E} = \sum_i P_H(i) \|s_i\|^2$$

$$\mathcal{E}_b = \frac{\mathcal{E}}{\log m},$$

where $\log m$ is the number of bits of information that we convey when we communicate one of m possible choices.

When we make an isometric transformation as defined in this subsection, the signal space in which we are living does not change (the basis is the same). In the next section we consider isometric transformations that carry the signal space from one subspace of \mathcal{L}_2 to another.

4.2.2 Changing the Orthonormal Basis

The error probability depends solely on the position of the signals *in the signal space*. We may think of constructing various sets of time-domain signals in the following way.

We first choose a signal set in the signal space. From this set we construct a set of time-domain waveforms by selecting an orthonormal basis. We then construct a second set of time-domain waveforms by selecting a second orthonormal basis. The procedure may be repeated indefinitely. The resulting sets of waveforms may look very different. For instance one set may be of signals that have finite support (i.e. they vanish when t is outside a specified time interval), whereas another set may have infinite support. Nevertheless the associated average probability of error is identical for all signal sets obtained as described. Notice that changing the basis constitutes an isometric transformation in \mathcal{L}_2 (as opposed to an isometry in \mathbb{R}^n as in the previous section).

EXAMPLE 40. Let the signals in the signal space be $s_0 = (\sqrt{\mathcal{E}}, 0)^T$ and $s_1 = (0, \sqrt{\mathcal{E}})^T$. This choice completely determines the error probability and the average energy. It does not say anything, however, about the signals $s_0(t)$ and $s_1(t)$, except that they are orthogonal to one another. Example 38 shows four possible choices for $s_0(t)$ and $s_1(t)$. (There are infinitively many other possibilities).

4.3 Time, Bandwidth, and Dimensionality

The bandwidth plays an important role in practice and should be included in our discussion on signal constellations. As a start we will focus on *baseband* signals, i.e. signals that have their spectral components centered around the origin.

One is tempted to define the bandwidth of a baseband signal $s(t)$ to be B if the support of $s_{\mathcal{F}}(t)$ is $[-\frac{B}{2}, \frac{B}{2}]$. This definition is not useful in practice since all man-made signals $s(t)$ have finite support (in the time domain) and thus $s_{\mathcal{F}}(f)$ has infinite support.¹

A better definition (but not the only one that makes sense) is to fix a number $\eta \in (0, 1)$ and say that the baseband signal $s(t)$ has bandwidth B if

$$\int_{-B}^B |s_{\mathcal{F}}(f)|^2 df = \|s\|^2(1 - \eta).$$

In words, the signal has bandwidth B if B is the smallest number such that the interval $(-B, B)$ contains $100(1 - \eta)\%$ of the signal power. The bandwidth changes if we change η . Reasonable values for η are $\eta = 0.1$ and $\eta = 0.01$.

This definition allows us to relate time, bandwidth, and dimensionality. If we let $\eta = \frac{1}{12}$ and define

$$\mathcal{L}_2(T, B) = \left\{ s(t) \in \mathcal{L}_2 : s(t) = 0, t \notin \left[-\frac{T}{2}, \frac{T}{2}\right] \text{ and } \int_{-B}^B |s_{\mathcal{F}}(f)|^2 df \geq \|s\|^2(1 - \eta) \right\}$$

then one can show that the dimensionality of $\mathcal{L}_2(T, B)$ is $n = \lfloor 2TB + 1 \rfloor$ (see Wozencraft & Jacobs for more on this). As T goes to infinity, we see that the number of dimensions

¹The support of a real or complex valued signal $x(t)$, $t \in \mathbb{R}$ is the set of those t for which $x(t) \neq 0$.

per second, $\frac{n}{T}$, goes to $2B$. Moreover, if one changes the value of η , then the essentially linear relationship between $\frac{n}{T}$ and B remains (but the constant in front of B will be different than 2).

4.4 Examples

The aim of this section is to sharpen our intuition by looking at some examples.

EXAMPLE 41. (PAM) *In this example we consider Pulse Amplitude Modulation (PAM). For $i \in \{0, 1, \dots, (m-1)\}$, let \mathbf{s}_i be a unique element of the signal space constellation $\{\pm\sqrt{E_w}, \pm 3\sqrt{E_w}, \pm 5\sqrt{E_w}, \dots, \pm(m-1)\sqrt{E_w}\}$ and let $s_i(t) = \mathbf{s}_i\psi(t)$, where ψ is an arbitrary unit-energy waveform.² Figure 4.1 shows the signal space constellation for $m = 6$.*

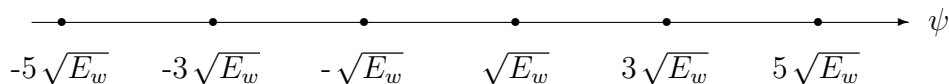


Figure 4.1: Signal Space Constellation for 6-ary PAM.

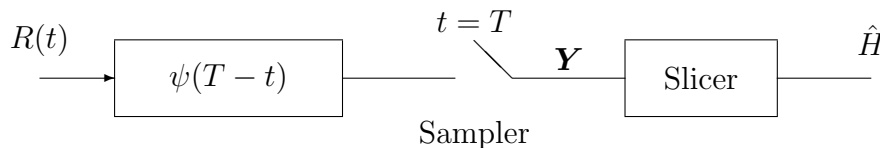


Figure 4.2: PAM Receiver

Naturally enough, in the receiver block diagram of Figure 4.2 projects the received waveform onto the unique elements of the orthonormal basis. The alternative of projecting onto each one of the m possible waveform signals by means of m filters would have been non-sense. The slicer finds one of the i for which \mathbf{s}_i is as close to \mathbf{y} as to any other element of the signal space constellation.

Determining the probability of error is straightforward. If i corresponds to one of the two end points, $P_e(i) = Q(\frac{d}{2\sigma})$, where $\frac{d}{2} = \sqrt{E_w}$ and $\sigma = \sqrt{\frac{N_0}{2}}$. For the remaining $m-2$ signal points, the probability of error is twice that of the end points. Taking the average, we obtain

$$P_e = \left(2 - \frac{2}{m}\right) Q\left(\sqrt{\frac{2E_w}{N_0}}\right). \quad (4.1)$$

²We are following our convention to use bold symbols such as \mathbf{s}_i for elements of the signal space, even though in this case the signal space has dimension 1.

Another quantity of interest is the average energy \mathcal{E} . A simple approximation to determine the average energy is obtained by computing the second moment of a random variable X that, instead of being uniformly distributed over the signal space constellation $\{\pm\sqrt{E_w}, \pm 3\sqrt{E_w}, \pm 5\sqrt{E_w}, \dots, \pm(m-1)\sqrt{E_w}\}$, is uniformly distributed over the interval $[-m\sqrt{E_w}, m\sqrt{E_w}]$. The second moment of this random variable is

$$E[X^2] = \frac{1}{m\sqrt{E_w}} \int_0^{m\sqrt{E_w}} s^2 ds = \frac{m^2 E_w}{3}.$$

This approximation becomes better as the elements of the signal space constellation become denser in the interval used for the approximation, which is the case if we let the number k of bits grow and keep the average energy per bit equal to some constant \mathcal{E}_b . In fact, the number of elements of the signal space constellation will grow exponentially in k and the number $\sqrt{E_w}$ that determines the spacing between them will decrease exponentially in k . The latter claim is quickly verified by letting $\mathcal{E} = k\mathcal{E}_b$ and choosing $\sqrt{E_w}$ so that $\mathcal{E} = EX^2$. A quick calculation shows that $EX^2 = \frac{m^2 E_w}{3}$, which yields $\sqrt{E_w} = \frac{\sqrt{3\mathcal{E}}}{m} = \frac{\sqrt{3k\mathcal{E}_b}}{2^k}$. This goes to zero exponentially fast as k goes to infinity. \square

The next example uses a two-dimensional constellation.

EXAMPLE 42. (Phase-Shift-Keying (PSK): Single Shot) Let $\tau = [0, T]$ and define

$$s_i(t) = \sqrt{\frac{2\mathcal{E}}{T}} \cos(2\pi f_0 t + \frac{2\pi}{m}i) 1_\tau(t), \quad i = 0, 1, \dots, m-1. \quad (4.2)$$

We assume $f_0 T = \frac{k}{2}$ for some integer k , so that $\|s_i\|^2 = \mathcal{E}$ for all i . The signal space representation may be obtained by using the trigonometric equivalence $\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$ to rewrite (4.2) as

$$s_i(t) = s_{i,1}\psi_1(t) + s_{i,2}\psi_2(t),$$

where

$$\begin{aligned} s_{i1} &= \sqrt{\mathcal{E}} \cos\left(\frac{2\pi i}{m}\right), & \psi_1(t) &= \sqrt{\frac{2}{T}} \cos(2\pi f_0 t) 1_\tau(t), \\ s_{i2} &= \sqrt{\mathcal{E}} \sin\left(\frac{2\pi i}{m}\right), & \psi_2(t) &= -\sqrt{\frac{2}{T}} \sin(2\pi f_0 t) 1_\tau(t). \end{aligned}$$

Hence, the n -tuple representation of the signals is

$$\mathbf{s}_i = \sqrt{\mathcal{E}} \begin{pmatrix} \cos 2\pi i/m \\ \sin 2\pi i/m \end{pmatrix}.$$

In Example 7 we have already studied this constellation and derived the following upper bound to the error probability

$$P_e \leq 2Q \left(\sqrt{\frac{\mathcal{E}}{\sigma^2}} \sin \frac{\pi}{m} \right),$$

where $\sigma^2 = \frac{N_0}{2}$ is the variance of the noise in each coordinate.

From the decoding regions plotted in Example 7 we also immediately see that for each $0 \leq i, j \leq m-1$, there is an isometry $a : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that $a(\mathbf{s}_i) = \mathbf{s}_j$ and $a(\mathcal{R}_i) = \mathcal{R}_j$. Thus, as an application of what we have learned in the previous section, we can tell the rather obvious fact that $P_e(i)$ is the same for all $i \in \mathcal{H}$.

As in the previous example we are interested in understanding what happens as k goes to infinity while \mathcal{E}_b remains constant. Since $\mathcal{E} = k\mathcal{E}_b$ grows linearly with k , the circle that contains the signal points has radius $\sqrt{\mathcal{E}} = \sqrt{k}\mathcal{E}_b$ that grows with \sqrt{k} , while the number $m = 2^k$ of points on this circle grows exponentially with k . Hence the minimum distance between points goes to zero (indeed exponentially fast). As a consequence, the argument of the Q function that upperbounds the probability of error for PSK goes to 0 and the probability of error goes to 1. Recall from Example 7 that the upperbound becomes tight as m grows. \square

As they are, the signal constellations used in the above two examples are not suitable to transmit a large amount of data. Indeed, to do so, we would have to let m be large enough so that $\log_2 m$ is the number of bits we want to transmit. As m grows, the probability of error goes to 1. The problem with these two examples is that, as m grows, we are trying to pack more and more signal points into a space that also grows in size but does not grow fast enough. The space becomes “crowded” as m grows, meaning that the minimum distance becomes smaller, and the probability of error increases.

In the next example we try to do better. So far we have not made use of the fact that we should expect to use more time to transmit more bits. In both of the above examples, the length T of the time interval used to communicate was constant. In the next example we let T grow linearly with the number of bits. This will free up a number of dimensions that grows linearly with k . (Recall that $n = 2BT$ is possible.) Each dimension may be used with the signal constellation of Example 41. Alternatively, every two dimensions may be used with the constellation of Example 42. Other possibilities exist.

EXAMPLE 43. (*Bit by Bit on a Pulse Train*) The idea is to transmit a signal of the form

$$S_i(t) = \sum_{j=1}^k S_j \psi_j(t), \quad (4.3)$$

and let $\psi_j(t) = \psi(t - jT)$ for some appropriate basic pulse $\psi(t)$ while fulfilling the requirement $\langle \psi_i, \psi_j \rangle = \delta_{ij}$. Assuming that it is indeed possible to find such a pulse, we obtain

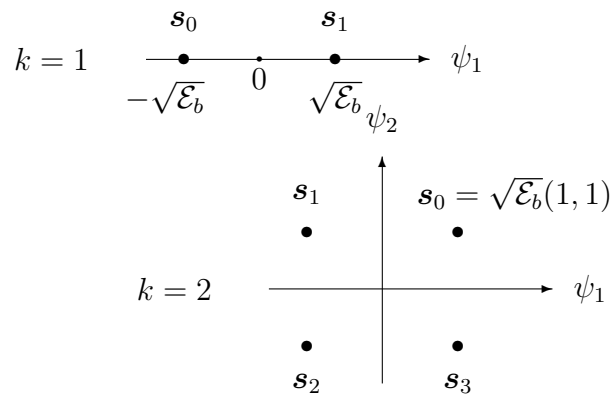
$$S_i(t) = \sum_{j=i}^k S_j \psi(t - jT_s). \quad (4.4)$$

In this example we use binary symbols, i.e., $S_j \in \{\pm\sqrt{\mathcal{E}_b}\}$, $j = 1, 2, \dots, k$. To be specific, we let $S_j = \sqrt{\mathcal{E}_b} D_j$, where $D_j \in \{\pm 1\}$ is the random variable that represents the j th

source bit. (For this example, it is convenient to consider bits as taking value in $\{\pm 1\}$ rather than in $\{0, 1\}$.) The subscript b indicates that \mathcal{E}_b is the energy per bit. For reasons that should be obvious, the above signaling method will be called bit-by-bit on a pulse train.

There are various possible choices for $\psi(t)$. Common choices for $\psi(t)$ are sinc pulses, rectangular pulses, and raised-cosine pulses (to be defined later). We will see how to choose $\varphi(t)$ in Chapter 5.

To gain insight in the operation of the receiver and to determine the error probability, it is always a good idea to try to picture the signal space constellation. In this case $\mathbf{s}_0, \dots, \mathbf{s}_{m-1}$, $m = 2^k$, are the vertices of a k -dimensional hypercube as shown in the figures below for $k = 1, 2$.



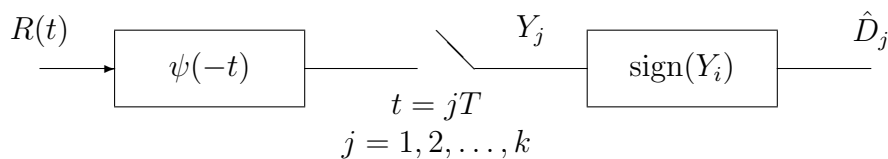
From the picture, you should be able to see what the decoding regions are, but let us proceed analytically. The ML receiver decide that the constellation point used by the sender is one of the $\mathbf{s} = (s_1, s_2, \dots, s_k) \in \{\pm\mathcal{E}_b\}^k$ that maximizes $\langle y, \mathbf{s} \rangle - \frac{\|\mathbf{s}\|^2}{2}$. Since $\|\mathbf{s}\|^2$ is the same for all constellation points, the previous expression is maximized iff $\langle \mathbf{y}, \mathbf{s}_i \rangle = \sum y_j s_j$ is maximized. We do so by making sure that the right hand side has non-negative terms, which means that we pick the sign of s_j to match the sign of y_j . When $y_j = 0$, it does not matter which sign we choose for s_j . Hence we may define

$$\text{sign}(y) = \begin{cases} 1 & y \geq 0 \\ 0 & y < 0 \end{cases},$$

and let our ML receiver be the one that decides the j th bit according to

$$\hat{d}_j = \text{sign}(y_j).$$

The next figure shows the block diagram of our ML receiver. Notice that we need only one matched filter to do the k projections. This is one of the reasons why we choose $\psi_i(t) = \psi(t - iT_s)$. Other reasons will be discussed in the next chapter.



We now compute the error probability. As usual, we first compute the error probability conditioned on the event $\mathbf{S} = \mathbf{s} = (s_1, \dots, s_k)$ for some arbitrary constellation point \mathbf{s} . From the geometry of the signal constellation, we expect that the error probability will not depend on \mathbf{s} . If s_j is positive, $Y_j = \sqrt{\mathcal{E}_b} + Z_j$ and \hat{D}_j will be correct iff $Z_j \geq -\sqrt{\mathcal{E}_b}$. This happens with probability $1 - Q\left(\frac{\sqrt{\mathcal{E}_b}}{\sigma}\right)$. Reasoning similarly, you should verify that the probability of error is the same if s_j is negative. Now let C_j be the event that the decoder makes the correct decision of bit j . The probability of C_j depends only on Z_j . The independence of the noise components implies the independence of the events C_1, C_2, \dots, C_k . Thus, the probability that all k bits are decoded correctly when $\mathbf{S} = \mathbf{s}$ is

$$P_c(\mathbf{s}) = \left[1 - Q\left(\frac{\sqrt{\mathcal{E}_b}}{\sigma}\right)\right]^k.$$

Since this probability does not depend on \mathbf{s} , we obtain that the unconditional error probability is also the same, namely

$$P_c = \left[1 - Q\left(\frac{\sqrt{\mathcal{E}_b}}{\sigma}\right)\right]^k.$$

Notice that $P_c \rightarrow 0$ as $k \rightarrow \infty$. However, the probability that a specific symbol (bit) be decoded incorrectly is $Q\left(\frac{\sqrt{\mathcal{E}_b}}{\sigma}\right)$. This is constant with respect to k .

The following properties (due to our choice $\psi_j(t) = \psi(t - jT)$) are worth noticing: (i) k may be arbitrary and may vary from one message file to the other without changing the structure of the transmitter and the receiver. (This would not be true with a general choice of ψ_1, \dots, ψ_k .); (ii) The transmitter does not have to wait until it has received all k information bits to start transmission. This is important in real time applications, e.g. speech, video, etc.; (iii) A ML receiver decides for each bit independently. Moreover, it can decide bit i as soon as the signal transmitted in the i th time interval has been received. All of the above properties are desirable for practical systems. \square

The drawback of bit-by-bit signaling is that $P_c \rightarrow 0$ as $k \rightarrow \infty$. Hence, as it is, it is not appropriate to communicate long files either. We are, however, in a better situation than with the first two examples of this section. In those examples the probability of error was going to one since signal points were getting closer as k increased. To the contrary, in bit-by-bit on a pulse train the probability that we make an error in decoding one or more of the k bits goes to one because the number of neighbors increases. Coding will fix this problem by ensuring that the distance between neighboring signal points grows enough to compensate for the growing number of neighbors.

While in the last example we have chosen to transmit a single bit per dimension, we could have transmitted instead some small number of bits per dimension by means of one of the methods discussed in the previous two examples. In that case we call the signaling scheme *symbol by symbol on a pulse train*. Symbol by symbol on a pulse train will come up often in the remainder of this course. In fact it is the basis for most digital communication systems.

The following question seems natural at this point: Is it possible to map k bits into a signal s_i and avoid that $P_c \rightarrow 0$ as $k \rightarrow \infty$? The next example shows that it is indeed possible.

EXAMPLE 44. (*Frequency Shift Keying (FSK): An Example Of Orthogonal Signaling*) Let $n = m = 2^k$. We do this by using m equi-norm orthogonal functions $s_1(t), \dots, s_m(t)$:

$$s_i = \sqrt{\mathcal{E}}\psi_i, \quad \langle \psi_i, \psi_j \rangle = \delta_{ij}.$$

This is called *block orthogonal signaling*. The name stems from the fact that one collects a block of k bits and maps them into one of 2^k orthogonal waveforms. Notice that $\|s_i\| = \sqrt{\mathcal{E}}$ for all i .

There are many ways to choose the 2^k waveforms ψ_i . One way is to choose $\psi_i(t) = \psi(t - iT)$ for some basic pulse $\psi(t)$ such that $\langle \psi(t - iT), \psi(t - jT) \rangle = \delta_{ij}$ as in bit-by-bit signaling. For reasons that should be obvious, this is sometimes called *pulse position modulation*. Notice, however, that now we need 2^k such shifts of ψ as opposed to only k such shifts. Another way is what is called *m-FSK* (*m-ary frequency shift keying*). Specifically,

$$s_i(t) = \sqrt{\frac{2\mathcal{E}}{T}} \cos(2\pi f_i t) 1_\tau(t) \quad (4.5)$$

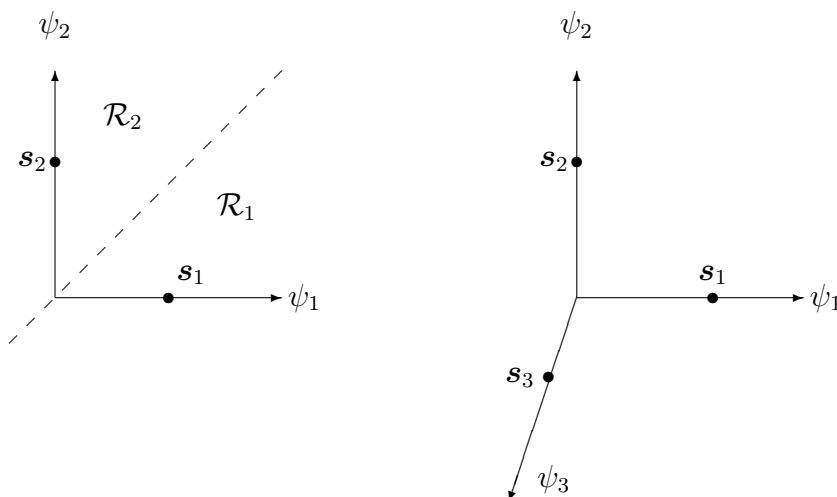
for some $\tau = [0, T]$ and $i = 1, 2, \dots, m$. (For FSK it is convenient to index this way rather than letting $i = 0, 1, \dots, m - 1$ as usual.) For convenience we assume $f_i T = k_i/2$ for some integer k_i such that $k_i \neq k_j$ if $i \neq j$. Then

$$\begin{aligned} \langle s_i, s_j \rangle &= \frac{2\mathcal{E}}{T} \int_0^T \cos(2\pi f_i t) \cos(2\pi f_j t) dt \\ &= \frac{2\mathcal{E}}{T} \int_0^T \left[\frac{1}{2} \cos[2\pi(f_i + f_j)t] + \frac{1}{2} \cos[2\pi(f_i - f_j)t] \right] dt \\ &= \frac{\mathcal{E}}{T} \int_0^T \cos[2\pi(f_i - f_j)t] dt \\ &= \mathcal{E} \delta_{ij}. \end{aligned}$$

Letting $\psi_i(t) = \sqrt{\frac{2}{T}} \cos 2\pi f_i t 1_\tau(t)$ we obtain

$$s_i = \sqrt{\mathcal{E}}\psi_i, \quad i = 1, \dots, m. \quad (4.6)$$

Hence we have an orthogonal signal set as desired.



When $m \geq 3$, it is not easy to visualize the decoding regions. However we can proceed analytically using the fact that \mathbf{s}_i is 0 everywhere except at position i where it is $\sqrt{\mathcal{E}}$. Hence,

$$\begin{aligned}\hat{H}_{ML}(\mathbf{y}) &= \arg \max_i \langle \mathbf{y}, \mathbf{s}_i \rangle - \frac{\mathcal{E}}{2} \\ &= \arg \max_i \langle \mathbf{y}, \mathbf{s}_i \rangle \\ &= \arg \max_i y_i.\end{aligned}$$

To compute (or bound) the error probability, we start as usual with a fixed \mathbf{s}_i . We pick $i = 1$. When $H = 1$,

$$Y_j = \begin{cases} Z_j & \text{if } j \neq 1, \\ \sqrt{\mathcal{E}} + Z_j & \text{if } j = 1. \end{cases}$$

Then

$$P_c(1) = \Pr\{Y_1 > Z_2, Y_1 > Z_3, \dots, Y_1 > Z_m | H = 1\}.$$

To evaluate the right side, we start by conditioning on $Y_1 = \alpha$, where $\alpha \in \mathbb{R}$ is an arbitrary number

$$\begin{aligned}\Pr\{c | H = 1, Y_1 = \alpha\} &= \Pr\{\alpha > Z_2, \dots, \alpha > Z_m\} \\ &= \left[1 - Q\left(\frac{\alpha}{\sqrt{N_0/2}}\right)\right]^{m-1},\end{aligned}$$

and then remove the conditioning on Y_1 ,

$$\begin{aligned}P_c(1) &= \int_{-\infty}^{\infty} f_{Y_1|H}(\alpha|1) \left[1 - Q\left(\frac{\alpha}{\sqrt{N_0/2}}\right)\right]^{m-1} d\alpha \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(\alpha - \sqrt{\mathcal{E}})^2}{N_0}} \left[1 - Q\left(\frac{\alpha}{\sqrt{N_0/2}}\right)\right]^{m-1} d\alpha,\end{aligned}$$

where we used the fact that when $H = 1$, $Y_1 \sim \mathcal{N}(\sqrt{\mathcal{E}}, \frac{N_0}{2})$. The above expression for $P_c(1)$ cannot be simplified further but one can evaluate it numerically. By symmetry,

$$P_c = P_c(1) = P_c(i)$$

for all i .

The union of events bound is especially useful when the signal set $\{\mathbf{s}_1, \dots, \mathbf{s}_m\}$ is completely symmetric, like for orthogonal signals. In this case:

$$\begin{aligned} P_e = P_e(i) &\leq (m-1)Q\left(\frac{d}{2\sigma}\right) \\ &= (m-1)Q\left(\sqrt{\frac{\mathcal{E}}{N_0}}\right) \\ &< 2^k \exp\left[-\frac{\mathcal{E}}{2N_0}\right] \\ &= \exp\left[-k\left(\frac{\mathcal{E}/k}{2N_0} - \ln 2\right)\right], \end{aligned}$$

where we used $\sigma^2 = \frac{N_0}{2}$ and $d = \sqrt{2\mathcal{E}}$. The latter follows from $d = \|\mathbf{s}_i - \mathbf{s}_j\|$ and

$$\|\mathbf{s}_i - \mathbf{s}_j\|^2 = \|\mathbf{s}_i\|^2 + \|\mathbf{s}_j\|^2 - 2\langle \mathbf{s}_i, \mathbf{s}_j \rangle = \|\mathbf{s}_i\|^2 + \|\mathbf{s}_j\|^2 = 2\mathcal{E},$$

which is Pythagorean's Theorem.

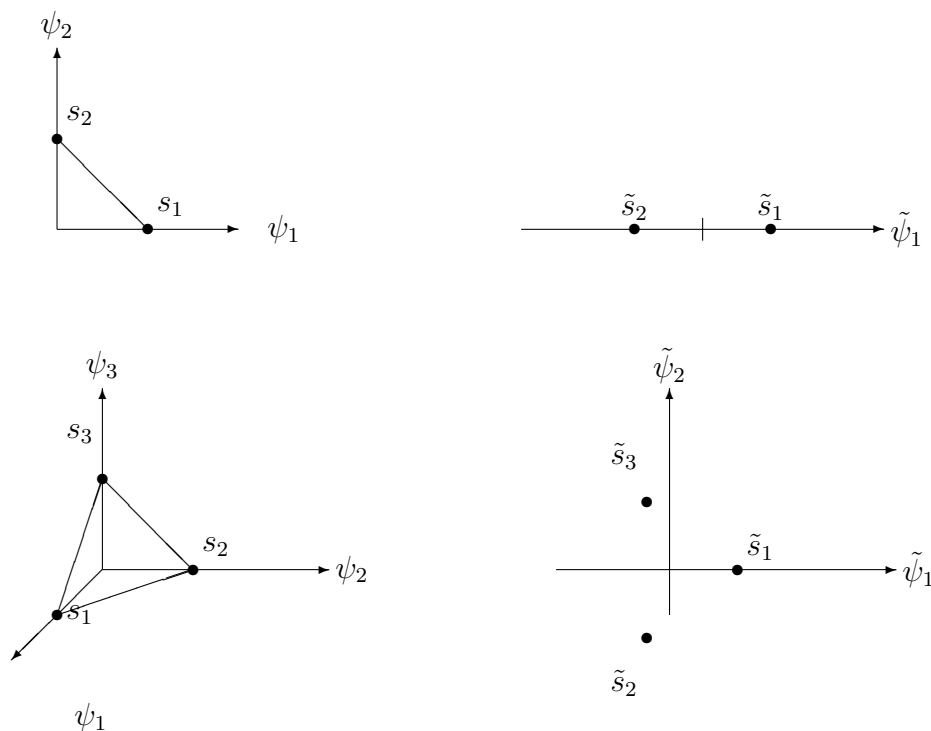
Here \mathcal{E} is the signal's energy. If we let $\mathcal{E} = \mathcal{E}_b k$, meaning that we let the signal's energy grow linearly with the number of bits as in bit-by-bit signaling, then we obtain

$$P_e < e^{-k(\frac{\mathcal{E}_b}{2N_0} - \ln 2)}.$$

Here $P_e \rightarrow 0$ as $k \rightarrow \infty$, provided that $\frac{\mathcal{E}_b}{N_0} > 2 \ln 2$. ($2 \ln 2$ is approximately 1.39.)

□

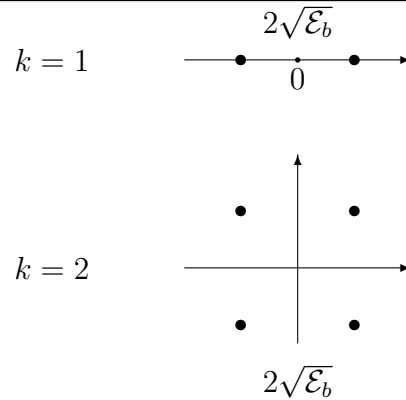
A useful application of the energy minimization idea (See Problem 1) applied to an orthogonal signal constellation leads to the *simplex* signal set.



4.5 Bit By Bit Versus Block Orthogonal

In one of the last two examples we have let the number of dimensions n increased linearly with the number k of bits and in the other example we have let n increased exponentially with k . In both cases we kept the energy per bit \mathcal{E}_b fixed, which means that the signal energy $\mathcal{E} = k\mathcal{E}_b$ grew linearly with the number k of bits. Let us compare the two cases.

In bit-by-bit on a pulse train the bandwidth is constant (we have not proved this yet, but this is consistent with the asymptotic limit $2B = n/T$ seen in Section 4.3 applied with $T = nT_s$), and the time and the energy increased linearly with k . These are all desirable properties. (We have also seen that the delay at the sender and at the receiver are small and that we need only one matched filter to do the projections but we will not take complexity and delay into this discussion). The drawback of bit-by-bit on a pulse train was found to be the fact that the probability of error goes to 1 as k goes to infinity. The union of events bound is a useful tool to understand what is going on. Let us use it to bound the probability of error when $H = i$. The union of events bound has one term for each alternative j . The dominating terms in the bound are those that correspond to signals \mathbf{s}_j that are the closest neighbors to \mathbf{s}_i . There are k closest neighbors, obtained by changing \mathbf{s}_i in exactly one component, and each of them is at distance $2\sqrt{\mathcal{E}_b}$ from \mathbf{s}_i (see the figure below). As k increases, the number of dominant terms goes up and so does the probability of error.



Let us now consider block orthogonal signaling. Since the dimensionality of the space it occupies grows exponentially with k , the expression $n = 2BT$ tells us that either the time or the bandwidth has to grow exponentially also. This is a significant drawback. Now let us consider the error probability. Using the bound

$$Q\left(\frac{d}{2\sigma}\right) \leq \frac{1}{2} \exp\left[-\frac{d^2}{8\sigma^2}\right] = \frac{1}{2} \exp\left[-\frac{k\mathcal{E}_b}{2N_0}\right]$$

we see that the probability that the noise carries a signal closer to a specific neighbor goes down as $\exp\left(-\frac{k\mathcal{E}_b}{2N_0}\right)$. There are $2^k - 1 = e^{k \ln 2} - 1$ nearest neighbors (all alternative signals are nearest neighbors). For $\frac{\mathcal{E}_b}{2N_0} > k \ln 2$, the growth in distance dominates the probability of error behavior. For $\frac{\mathcal{E}_b}{2N_0} < k \ln 2$ the number of neighbors dominates. Finally notice that the bit error probability P_b can not be larger than the block error probability P_e . Indeed they are the same iff every time that the decoder selects a wrong message the bit sequence that corresponds to this message has all bits flipped with respect to the bit sequence that corresponds to the correct message.

4.6 Conclusion

We have discussed some of the trade-offs between the number of transmitted bits, the duration, the bandwidth, and the energy of the signal we use to transmit those bits, and the resulting error probability. We have seen that, rather surprisingly, it is possible to transmit an increasing number k of bits at a fixed energy per bit \mathcal{E}_b and make the probability that even a single bit is decoded incorrectly go to zero as k increases. However, the scheme we used to prove this has the undesirable property of requiring an exponential growth of the time bandwidth product. Ideally we would like to make the probability of error go to zero with a scheme similar to bit by bit on a pulse train. Is it possible? The answer is yes and the technique to do so is coding. We will give an example of coding in Chapter 6.

The study of the fundamental relationships between the rate at which we want to communicate (e.g. in bits per second per Hz), the power of the signal (measured at the receiver),

and the probability of error that can be achieved is a typical subject of information theory. For instance, the *capacity* of the additive white Gaussian noise channel of power spectral density N_0 is

$$C = B \log_2 \left(1 + \frac{P}{N_0 B} \right) \quad [\text{bits/sec}],$$

where B is the bandwidth [Hz] and P is the power (energy per second) [Watts] that we are allowed to use. One can show that at rates smaller than C one can make the communication arbitrarily reliable. This is not possible at rates above C .

4.7 Problems

PROBLEM 1. (Minimum-energy Signals.)

Consider a given signal constellation consisting of vectors $\{s_1, s_2, \dots, s_m\}$. Let signal s_i occur with probability p_i . In this problem, we study the influence of moving the origin of the coordinate system of the signal constellation. That is, we study the properties of the signal constellation $\{s_1 - a, s_2 - a, \dots, s_m - a\}$ as a function of a .

- (i) Draw a sample signal constellation, and draw its shift by a sample vector a .
- (ii) Does the average error probability, $Pr\{e\}$, depend on the value of a ? Explain.
- (iii) The average energy per symbol depends on the value of a . For a given signal constellation $\{s_1, s_2, \dots, s_m\}$ and given signal probabilities p_i , prove that the value of a that minimizes the average energy per symbol is the centroid (the center of gravity) of the signal constellation, i.e.,

$$a = \sum_{i=1}^m p_i s_i. \quad (4.7)$$

Hint: First prove that if X is a real-valued zero-mean random variable and $b \in \mathbb{R}$, then $E[X^2] \leq E[(X - b)^2]$ with equality iff $b = 0$. Then extend your proof to vectors and consider $\mathbf{X} = \mathbf{S} - E[\mathbf{S}]$ where $\mathbf{S} = s_i$ with probability p_i .

PROBLEM 2. (Orthogonal Signal Sets.)

Consider the following situation: A signal set $\{s_j(t)\}_{j=0}^{m-1}$ has the property that all signals have the same energy \mathcal{E}_s and that they are mutually orthogonal:

$$\langle s_i, s_j \rangle = \mathcal{E}_s \delta_{ij}. \quad (4.8)$$

Assume also that all signals are equally likely. The goal is to transform this signal set into a minimum-energy signal set $\{s_j^*(t)\}_{j=0}^{m-1}$. It will prove useful to also introduce the unit-energy signals $\phi_j(t)$ such that $s_j(t) = \sqrt{\mathcal{E}_s} \phi_j(t)$.

- (i) Find the minimum-energy signal set $\{s_j^*(t)\}_{j=0}^{m-1}$.
- (ii) What is the dimension of $\text{span}\{s_0^*(t), \dots, s_{m-1}^*(t)\}$? For $m = 3$, sketch $\{s_j(t)\}_{j=0}^{m-1}$ and the corresponding minimum-energy signal set.
- (iii) What is the average energy per symbol if $\{s_j^*(t)\}_{j=0}^{m-1}$ is used? What are the savings in energy (compared to when $\{s_j(t)\}_{j=0}^{m-1}$ is used) as a function of m ?

PROBLEM 3. (Antipodal Signaling with Rayleigh Fading.)

Suppose that we use antipodal signaling (i.e. $s_0(t) = -s_1(t)$). When the energy per symbol is \mathcal{E}_b and the power spectral density of the additive white Gaussian noise in the channel is $N_0/2$, then we know that the average probability of error is

$$\Pr\{e\} = Q\left(\sqrt{\frac{\mathcal{E}_b}{N_0/2}}\right). \quad (4.9)$$

In mobile communications, one of the dominating effects is fading. A simple model of fading is as follows: Let the channel attenuate the signal by a random variable A . Specifically, if \mathbf{s}_i is transmitted, the received signal is $Y = A\mathbf{s}_i + N$. The probability density function of A depends on the particular channel that is to be modeled.³ Suppose A assumes the value a . From the receiver point of view this is as if there is no fading and the transmitter uses the signals $as_0(t)$ and $-as_0(t)$. Hence,

$$\Pr\{e|A = a\} = Q\left(\sqrt{\frac{a^2\mathcal{E}_b}{N_0/2}}\right). \quad (4.10)$$

The average probability of error can thus be computed by taking the expectation over the random variable A , i.e.

$$\Pr\{e\} = E_A[\Pr\{e|A\}] \quad (4.11)$$

An interesting, yet simple model is to take A to be a Rayleigh random variable, i.e.

$$f_A(a) = \begin{cases} 2ae^{-a^2}, & \text{if } a \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (4.12)$$

This type of fading, which can be justified especially for wireless communications is called Rayleigh fading.

- (i) Compute the average probability of error for antipodal signaling subject to Rayleigh fading.
- (ii) Comment on the difference between Eqn. (4.9) (the average error probability without fading) and your result in (i) (the average error probability with Rayleigh fading). Is it significant? For an average error probability $\Pr\{e\} = 10^{-5}$, find the necessary \mathcal{E}_b/N_0 for both cases.

³In a more realistic model, not only the amplitude, but also the phase of the channel transfer function is a random variable.

PROBLEM 4. (i) The root-mean square (rms) bandwidth of a low-pass signal $g(t)$ of finite energy is defined by

$$W_{rms} = \left[\frac{\int_{-\infty}^{\infty} f^2 |G(f)|^2 df}{\int_{-\infty}^{\infty} |G(f)|^2 df} \right]^{1/2}$$

where $|G(f)|^2$ is the energy spectral density of the signal. Correspondingly, the root-mean-square (rms) duration of the signal is defined by

$$T_{rms} = \left[\frac{\int_{-\infty}^{\infty} t^2 |g(t)|^2 dt}{\int_{-\infty}^{\infty} |g(t)|^2 dt} \right]^{1/2}.$$

Using these definitions and assuming that $|g(t)| \rightarrow 0$ faster than $1/\sqrt{|t|}$ as $|t| \rightarrow \infty$, show that

$$T_{rms} W_{rms} \geq \frac{1}{4\pi}.$$

Hint: Use Schwarz's inequality

$$\left\{ \int_{-\infty}^{\infty} [g_1^*(t)g_2(t) + g_1(t)g_2^*(t)] dt \right\}^2 \leq 4 \int_{-\infty}^{\infty} |g_1(t)|^2 dt \int_{-\infty}^{\infty} |g_2(t)|^2 dt$$

in which we set

$$g_1(t) = tg(t)$$

and

$$g_2(t) = \frac{dg(t)}{dt}.$$

(ii) Consider a Gaussian pulse defined by

$$g(t) = \exp(-\pi t^2).$$

Show that for this signal, the equality

$$T_{rms} W_{rms} = \frac{1}{4\pi}$$

can be reached.

Hint:

$$\exp(-\pi t^2) \xrightarrow{\mathcal{F}} \exp(-\pi f^2).$$

PROBLEM 5. (Minimum Energy for Orthogonal Signaling)

Let $H \in \{1, \dots, m\}$ be uniformly distributed and consider the communication problem described by:

$$H = i : \quad \mathbf{Y} = s_i + Z, \quad Z \sim \mathcal{N}(0, \sigma^2 I_m),$$

where s_1, \dots, s_m , $s_i \in \mathbb{R}^m$, is a set of constant-energy orthogonal signals. Without loss of generality we assume

$$s_i = \sqrt{\mathcal{E}} e_i,$$

where e_i is the i th unit vector in \mathbb{R}^m , i.e., the vector that contains 1 at position i and 0 elsewhere, and \mathcal{E} is some positive constant.

- (a) Describe the statistic of Y_j (the j th component of \mathbf{Y}) for $j = 1, \dots, m$ given that $H = 1$.
- (b) Consider a suboptimal receiver that uses a threshold $t = \alpha\sqrt{\mathcal{E}}$ where $0 < \alpha < 1$. The receiver declares $\hat{H} = i$ if i is the only integer such that $Y_i \geq t$. If there is no such i or there is more than one index i for which $Y_i \geq t$, the receiver declares that it can't decide. This will be viewed as an error.

Let $E_i = \{Y_i \geq t\}$, $E_i^c = \{Y_i < t\}$, and describe, in words, the meaning of the event

$$E_1 \cap E_2^c \cap E_3^c \cap \dots \cap E_m^c.$$

- (c) Find an upper bound to the probability that the above event does not occur when $H = 1$. Express your result using the Q function.
- (d) Now we let \mathcal{E} and $\ln m$ go to ∞ while keeping their ratio constant, namely $\mathcal{E} = \mathcal{E}_b \ln m \log_2 e$. (Here \mathcal{E}_b is the energy per transmitted bit.) Find the smallest value of \mathcal{E}_b/σ^2 (according to your bound) for which the error probability goes to zero as \mathcal{E} goes to ∞ . Hint: Use $m - 1 < m = \exp(\ln m)$ and $Q(x) < \frac{1}{2} \exp(-\frac{x^2}{2})$.

PROBLEM 6. (Pulse Amplitude Modulated Signals)

Consider using the signal set

$$s_i(t) = s_i \phi(t), \quad i = 0, 1, \dots, m-1,$$

where $\phi(t)$ is a unit-energy waveform, $s_i \in \{\pm \frac{d}{2}, \pm \frac{3}{2}d, \dots, \pm \frac{m-1}{2}d\}$, and $m \geq 2$ is an even integer.

- (a) Assuming that all signals are equally likely, determine the average energy \mathcal{E}_s as a function of m . Hint: $\sum_{i=0}^n i^2 = \frac{n}{6} + \frac{n^2}{2} + \frac{n^3}{3}$. Note: If you prefer you may determine an approximation of the average energy by assuming that $S(t) = S\phi(t)$ and S is a continuous random variable which is uniformly distributed in the interval $[-\frac{m}{2}d, \frac{m}{2}d]$.
- (b) Draw a block diagram for the ML receiver, assuming that the channel is AWGN with power spectral density $\frac{N_0}{2}$.
- (c) Give an expression for the error probability.
- (d) For large values of m , the probability of error is essentially independent of m but the energy is not. Let k be the number of bits you send every time you transmit $s_i(t)$ for some i , and rewrite \mathcal{E}_s as a function of k . For large values of k , how does the energy behaves when k increases by 1?

Chapter 5

Controlling the Spectrum

5.1 Introduction

In many applications, notably cellular communications, the power spectral density of the transmitted signal has to fit a certain frequency-domain mask. This restriction is enforced to limit the amount of interference that a user can cause to users of adjacent bands. In this chapter we learn some basics on how to fulfill certain frequency domain constraints. We will always assume the additive white Gaussian noise (AWGN) channel with noise power spectral density $\frac{N_0}{2}$.

Before reading the next section you may want to review the sampling theorem in the appendix of this chapter.

5.2 The Ideal Lowpass Case

An instructive case to start with is when the spectrum of the transmitted signal has to vanish outside a frequency interval $[-B, B]$ for some bandwidth B . When this is the case we talk about *lowpass communication*. Sometimes the limitation is due to regulations and sometimes it is due to the channel characteristic. For instance, analog telephone lines have a lowpass filter with cut-off frequency roughly at 4 KHz. In either case it is convenient to imagine that the channel model is that of Figure 5.1 where the channel impulse response h is that of an ideal lowpass filter, i.e.,

$$h_{\mathcal{F}}(f) = \begin{cases} 1, & |f| \leq B \\ 0, & \text{otherwise.} \end{cases}$$

The sampling theorem is the right tool to deal with this situation:

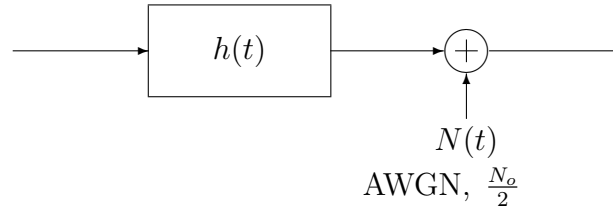


Figure 5.1: Lowpass channel model.

THEOREM 45. (Sampling Theorem)¹ Let $s(t)$ be a function in \mathcal{L}_2 that is lowpass limited to B . Then $s(t)$ is specified by its values at a sequence of points spaced $T = \frac{1}{2B}$ apart. In particular

$$s(t) = \sum_{n=-\infty}^{\infty} s(nT) \operatorname{sinc}\left(\frac{t}{T} - n\right) \quad (5.1)$$

where $\operatorname{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$. □

The sinc pulse does not have unit energy. Hence we define (its normalized version) $\psi(t) = \frac{1}{\sqrt{T}} \operatorname{sinc}\left(\frac{t}{T}\right)$. The set $\{\psi(t - jT)\}_{j=-\infty}^{\infty}$ forms an orthonormal set. Hence (5.1) can be rewritten as

$$s(t) = \sum_{j=-\infty}^{\infty} s_j \psi(t - jT), \quad \psi(t) = \frac{1}{\sqrt{T}} \operatorname{sinc}\left(\frac{t}{T}\right) \quad (5.2)$$

where $s_j = s(jT)\sqrt{T}$. This highlights the way the sampling theorem should be seen, namely as an orthogonal (but not orthonormal) expansion. In this expansion the basis is formed by time translates of sinc pulses. Implicit in the sampling theorem is that the set $\{\psi(t - iT)\}_{i=-\infty}^{\infty}$ is a complete orthonormal basis for the inner product space of finite energy waveforms that are lowpass limited to $B = \frac{1}{2T}$.

Now let us go back to our communication problem. The filter in the channel model is lowpass. Hence, any component of the transmitted signal $s(t)$ that lies outside the frequency range $[-B, B]$ will not be visible to the receiver; thus we may as well consider only signals $s(t)$ that have no frequency components outside $[-B, B]$. All such signals may be written in the form (5.2). Hence, without loss of generality, we may decide to transmit only signals of the form (5.2). Notice that (5.2) is the familiar form $\sum_j s_j \psi_j(t)$ with $\psi_j(t) = \psi(t - jT)$. For these signals the filter is transparent, which implies that the optimal receiver derived in Chapter 3 for the AWGN channel (without lowpass filter) is also optimal for the channel model at hand. Figure 5.2 shows the system block diagram. For now s_j is a symbol that may take values in any discrete subset of \mathbb{R} such as PAM. (Later on we will see that there is a benefit in letting it be from a discrete subsets of \mathbb{C} . An example of this is PSK where symbols are seen as elements of \mathbb{C} rather than of \mathbb{R}^2)

¹See the Appendix for a proof of the sampling Theorem.

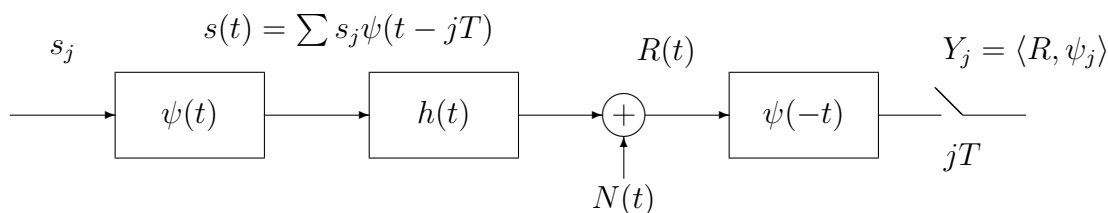


Figure 5.2: Lowpass system.

It is interesting to observe that the sampling theorem is somewhat used backwards in the diagram of Figure 5.2 . Until now you have probably seen sampling being done at the source and reconstruction at the destination. In Figure 5.2 the reconstruction part of the sampling theorem is used at the sender, and the sampling part at the receiver. The receiver also contains a filter with impulse response $\psi(-t)$ but this filter only scales $s(t)$. In fact, due to the symmetry of the sinc pulse, $\psi(-t) = \psi(t)$. Hence the filter at the receiver is an ideal lowpass (up to some scaling factor) that lets $s(\tau)$ through and cuts off the components of the noise $N(t)$ that fall outside $[-B, B]$. It is also interesting to observe that the sender and the receiver have the same structure as that of bit-by-bit on a pulse train described in Example 43.

From the input to the output of the block diagram of Figure 5.2 we see the discrete-time Gaussian channel depicted in Figure 5.3 and studied in Chapter 2. The channel takes and delivers a new symbol every T seconds. In the sampling theorem s_j is an arbitrary real number, but we would only use values that are element of the signal constellation. As we will see, the spectrum does not change if subsequent symbols are selected independently. If a symbol carries k_s bits, i.e. it is element of a constellation that contains 2^{k_s} signal points, then the system carries $\frac{k_s}{T}$ bits per second.

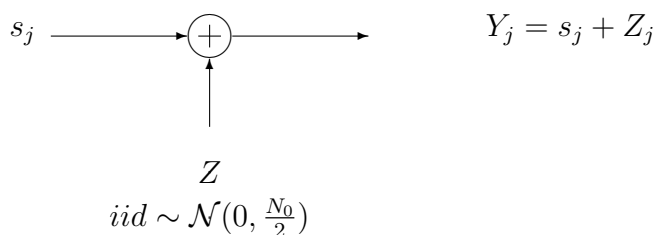


Figure 5.3: Equivalent discrete time channel.

5.3 Generalization Using Nyquist Pulses

The system described in the previous section, if implemented, produces a strictly rectangular spectrum. Nobody requires the spectrum to end with a strictly vertical edge and for a good reason: it would require using sinc pulses and sinc pulses are both non-causal and of infinite duration. What if we do not use sinc pulses?

The front end of the receiver of Figure 5.2 is optimal provided that $\{\psi(t - iT)\}, i \in \mathbb{Z}$, is an orthonormal sequence. As long as this is the case, the probability of error will not depend on ψ . However, ψ does affect the spectrum of the transmitted signal. We will see that the power spectral density of the transmitted signal when symbols are iid and have second moment \mathcal{E} is $\frac{\mathcal{E}|\psi_{\mathcal{F}}(f)|^2}{T}$. This suggests that we may want to choose ψ starting from its Fourier transform $\psi_{\mathcal{F}}$. Nyquist theorem, that we now derive, gives the necessary and sufficient condition that $\psi_{\mathcal{F}}$ has to satisfy to ensure that the time domain pulse ψ and its shifts form an orthonormal sequence, i.e.,

$$\int_{-\infty}^{\infty} \psi(t - nT)\psi^*(t)dt = \delta_n. \quad (5.3)$$

First define the following periodic function of period $\frac{1}{T}$

$$g(f) = \sum_{k \in \mathbb{Z}} \psi_{\mathcal{F}}(f + \frac{k}{T})\psi_{\mathcal{F}}^*(f + \frac{k}{T}).$$

We can now transform (5.3) as follows:

$$\delta_n = \int_{-\infty}^{\infty} \psi(t - nT)\psi^*(t)dt \stackrel{(a)}{=} \int_{-\infty}^{\infty} \psi_{\mathcal{F}}(f)\psi_{\mathcal{F}}^*(f)e^{-j2\pi nTf}df \stackrel{(b)}{=} \int_{-\frac{1}{2T}}^{\frac{1}{2T}} g(f)e^{-j2\pi nTf}df, \quad (5.4)$$

where in (a) we used Parseval's relationship and the shift property of the Fourier transform, and in (b) we made repeated use of the fact that for an arbitrary function $u(x)$ and an arbitrary interval $[-\frac{a}{2} + ia, \frac{a}{2} + ia]$ in the domain of u ,

$$\int_{-\frac{a}{2} + ia}^{\frac{a}{2} + ia} u(x)dx = \int_{-\frac{a}{2}}^{\frac{a}{2}} u(x + ia)dx.$$

But the last expression in (5.4) is $\frac{A_n}{T}$, where A_n is the n -th Fourier series coefficient of $g(f)$ (see 5.9). Comparing the first and the last term in (5.4), we also see that $A_0 = T$ and $A_n = 0$, $n \neq 0$. This means that $g(f) = T$. Since $g(f)$ has period $\frac{1}{T}$, $g(t)$ is constant over all $f \in \mathbb{R}$ iff it is constant over an arbitrary interval of length $\frac{1}{T}$.

We have proved the following

THEOREM 46. (Nyquist). A waveform $\psi(t)$ is orthonormal to each shift $\psi(t - nT)$ if and only if

$$\sum_{k=-\infty}^{\infty} |\psi_{\mathcal{F}}(f + \frac{k}{T})|^2 = T \quad \text{for all } f \text{ in some interval of length } \frac{1}{T}. \quad (5.5)$$

□

Waveforms that fulfill Nyquist theorem are called Nyquist pulses. A few comments are in order:

- (a) Often we are interested in Nyquist pulses that have small bandwidth, between $1/2T$ and $1/T$. For pulses that are strictly bandlimited to $1/T$ or less, the Nyquist criterion is satisfied if and only if $|\psi_{\mathcal{F}}(\frac{1}{2T} - \epsilon)|^2 + |\psi_{\mathcal{F}}(-\frac{1}{2T} - \epsilon)|^2 = T$ for $\epsilon \in [-\frac{1}{2T}, \frac{1}{2T}]$ (See the picture below). If we assume (as we do) that $\psi(t)$ is real-valued, then $|\psi_{\mathcal{F}}(-f)|^2 = |\psi_{\mathcal{F}}(f)|^2$. In this case the above relationship is equivalent to

$$|\psi_{\mathcal{F}}(\frac{1}{2T} - \epsilon)|^2 + |\psi_{\mathcal{F}}(\frac{1}{2T} + \epsilon)|^2 = T, \quad \epsilon \in [0, \frac{1}{2T}].$$

This means that $|\psi_{\mathcal{F}}(\frac{1}{2T})|^2 = \frac{T}{2}$ and the amount by which $|\psi_{\mathcal{F}}(f)|^2$ increases when we go from $f = \frac{1}{2T}$ to $f = \frac{1}{2T} - \epsilon$ equals the decrease when we go from $f = \frac{1}{2T}$ to $f = \frac{1}{2T} + \epsilon$.

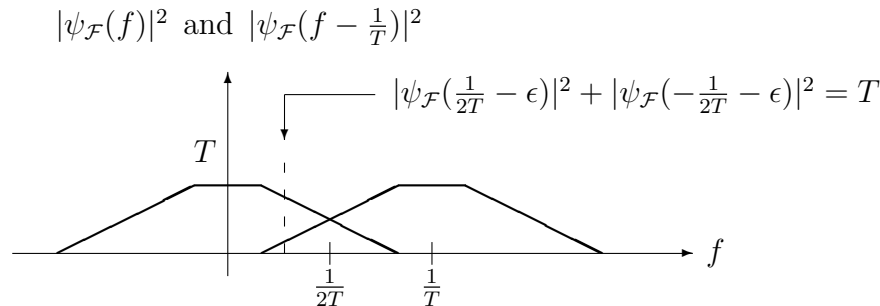


Figure 5.4: Nyquist condition for pulses $\psi_{\mathcal{F}}(f)$ that have support within $[-\frac{1}{T}, \frac{1}{T}]$.

- (b) The sinc pulse is just a special case of a Nyquist pulse. It has the smallest possible bandwidth, namely $1/2T$ [Hz], among all pulses that satisfy Nyquist criterion for a given T . (Draw a picture if this is not clear to you).
- (c) Nyquist criterion is a condition expressed in the frequency domain. It is equivalent to the time domain condition (5.3). Hence if one asks you to “verify that $\psi(t)$ fulfills Nyquist criterion” it does not mean that you have to take the Fourier transform of ψ and then check that $\psi_{\mathcal{F}}$ fulfills (5.5). It may be easier to check if ψ fulfills the time-domain condition (5.3).

(d) Any pulse $\psi(t)$ that satisfies

$$|\psi_{\mathcal{F}}(f)|^2 = \begin{cases} T, & |f| \leq \frac{1-\alpha}{2T} \\ \frac{T}{2} \left(1 + \cos \left[\frac{\pi T}{\alpha} \left(|f| - \frac{1-\alpha}{2T}\right)\right]\right), & \frac{1-\alpha}{2T} < |f| < \frac{1+\alpha}{2T} \\ 0, & \text{otherwise} \end{cases}$$

for some $\alpha \in (0, 1)$ fulfills Nyquist criterion. Such a pulse is called a root-raised-cosine pulse. (Draw a picture).

(e) We have derived Nyquist criterion inspired by what we have done in Section 5.2. However, Nyquist criterion is not limited to lowpass signals. If ψ fulfills Nyquist criterion and has bandpass characteristic then it will give rise to a bandpass signal $s(t)$.

5.4 Summary and Conclusion

The communication problem, as we see it in this course, consists of signaling to a recipient a message chosen by a sender. The message is one out of m possible messages. Each message has a unique signal used as a proxy to communicate the message across the channel.

Regardless of how we pick the m signals, which are assumed to be finite-energy and known to the sender and the receiver, there exists an orthonormal basis $\psi_1 \dots \psi_n$ and a constellation of points $\mathbf{s}_0, \dots, \mathbf{s}_{m-1}$ in \mathbb{C}^n (the signal space) such that

$$s_i(t) = \sum_{j=1}^n s_{ij} \psi_j(t), \quad i = 0, \dots, m-1. \quad (5.6)$$

A minimum-probability-of-error receiver that observes the received signal R may decide which message was signaled by means of a vector $\mathbf{Y} = (Y_1, \dots, Y_n)^T \in \mathbb{C}^n$ where $Y_j = \langle R, \psi_j \rangle$.

It is up to us to decide if we want to start by choosing the m waveforms s_i , $i = 0, \dots, m-1$ and then use the Gram Schmidt procedure to find an orthonormal basis $\psi_1 \dots \psi_n$ and the associated signal-space points $\mathbf{s}_0 \dots, \mathbf{s}_{m-1}$, or if we want to start with an arbitrary orthonormal basis $\psi_1 \dots \psi_n$ and a selection of m signal-space points $\mathbf{s}_0 \dots, \mathbf{s}_{m-1}$ and let the signaling waveforms be obtained through (5.6). The latter approach has the advantage of decoupling design choices that can be made independently: The orthonormal basis affects the duration of the signals and the bandwidth, whereas the signal space points affects the transmit power and the probability of error.

In Chapter 4 we have already come across the idea of letting ψ_1, \dots, ψ_n be obtained from a single pulse ψ by the assignment $\psi_i(t) = \psi(t - iT)$. In that occasion we motivated our choice by saying that a single matched filter is sufficient to generate the n components

of \mathbf{Y} . In the present chapter we have seen that choosing the orthonormal basis by time-translating a single pulse also lets us control the power spectral density of the transmitted signal which is $\mathcal{E} \frac{|\psi_{\mathcal{F}}(f)|^2}{T}$. There is a restriction though: we are allowed to use $\psi(t - iT)$ as $\psi_i(t)$ iff

$$\langle \psi(t), \psi(t - iT) \rangle = \delta_i. \quad (5.7)$$

The main result of this chapter was the derivation of Nyquist condition: it tells us which functions $\frac{|\psi_{\mathcal{F}}(f)|^2}{T}$ (thus which power spectral densities) are compatible with (5.7).

5.5 Problems

PROBLEM 1. Consider the transmitted signal, $S(t) = \sum_i X_i \psi(t - iT)$, where $X_i \in \{\pm 1\}$ are i.i.d random variables and $\{\psi(t - iT)\}_{i=-\infty}^{\infty}$ forms an orthonormal set. Let $Y(t)$ be the matched filter output at the receiver. Then in the absence of noise, X_i 's are the samples of $Y(t)$, sampled at integer multiples of T i.e $Y(iT) = X_i$. In this MATLAB exercise we will try to see how crucial it is to sample at $t = iT$ as opposed to $t = iT + \epsilon$. Towards that goal we plot the so-called eye diagram.

An eye diagram is the plot of $Y(t + iT)$ versus $t \in [-\frac{T}{2}, \frac{T}{2}]$, plotted on top of each other for each $i = 0 \cdots K - 1$, where K is the number of transmitted symbols. Thus at $t = 0$ on the eye diagram lies our sampling points mentioned earlier.

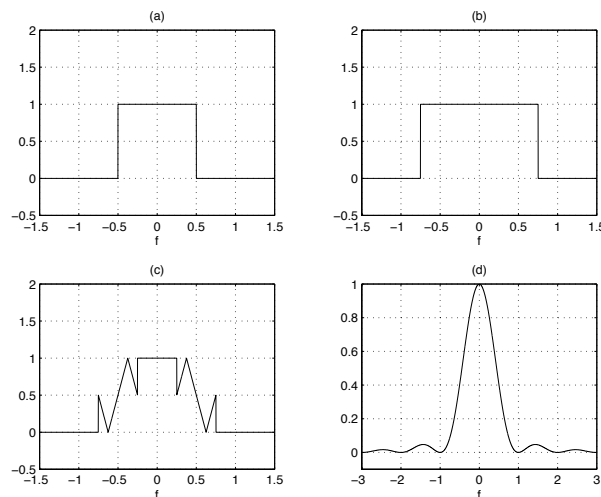
(a) Assuming $K = 100$, $T = 1$ and 10 samples per time period T , plot the eye diagrams when,

- (i) $\psi(t)$ is a raised cosine with $\alpha = 1$.
- (ii) $\psi(t)$ is a raised cosine with $\alpha = \frac{1}{2}$.
- (iii) $\psi(t)$ is a raised cosine with $\alpha = 0$ (or sinc).

(b) From the plotted eye diagrams what can you say about the cruciality of the sampling points with respect to α .

PROBLEM 2. (Nyquist Pulses.)

(i) Consider the following $|\theta_{\mathcal{F}}(f)|^2$. The unit on the frequency axis is $1/T$ and the unit on the vertical axis is T . Which ones correspond to Nyquist pulses $\theta(t)$ for symbol rate $1/T$? Note: Figure (d) shows a sinc² function.

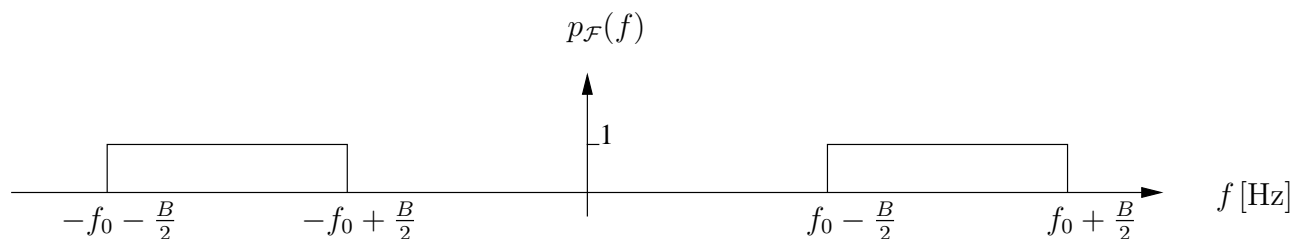


(ii) Design a (non-trivial) Nyquist pulse yourself.

(iii) Sketch the block diagram of a binary communication system that employs Nyquist pulses. Write out the formula for the signal after the matched filter. Explain the advantages of using Nyquist pulses.

PROBLEM 3. (Nyquist Pulses.)

Consider a pulse $p(t)$ defined via its Fourier transform $p_{\mathcal{F}}(f)$ as follows:



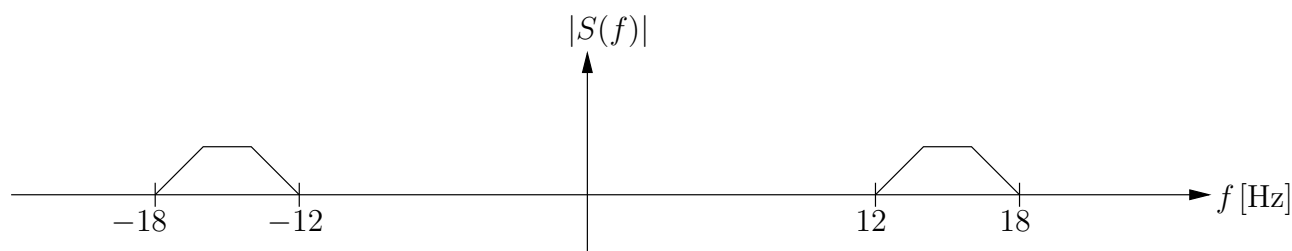
- (a) What is the expression for $p(t)$? (If you can't determine a mathematical expression, you may draw $p(t)$ qualitatively).
- (b) Determine the constant c so that $\psi(t) = cp(t)$ has unit energy.
- (c) Assume that $f_0 - \frac{B}{2} = B$ and consider the infinite set of functions $\dots, \psi(t+T), \psi(t), \psi(t-T), \psi(t-2T), \dots$. Do they form an orthonormal set for $T = \frac{1}{2B}$? (Explain).
- (d) Determine all possible values of $f_0 - \frac{B}{2}$ so that $\dots, \psi(t+T), \psi(t), \psi(t-T), \psi(t-2T), \dots$ forms an orthonormal set.

PROBLEM 4. (Bandpass Sampling)

- (a) Let f_s be fixed. Consider a bandpass signal $s(t)$ with support $[-(l+1)\frac{f_s}{2}, -l\frac{f_s}{2}] \cup [l\frac{f_s}{2}, (l+1)\frac{f_s}{2}]$. Prove that $s(t)$ can be reconstructed from samples of the form $s(kT_s)$, where $k \in \mathbb{Z}$ and $T_s = \frac{1}{f_s}$.

Hint: If we define $h_{\mathcal{F}}(f) = 1_{[-(l+1)\frac{f_s}{2}, -l\frac{f_s}{2}] \cup [l\frac{f_s}{2}, (l+1)\frac{f_s}{2]}(f)$, then $s_{\mathcal{F}}(f) = \tilde{s}_{\mathcal{F}}(f)h_{\mathcal{F}}(f)$, where $\tilde{s}_{\mathcal{F}}(f) = \sum_{k \in \mathbb{Z}} s_{\mathcal{F}}(f - f_s k)$ is the periodic extension of $s_{\mathcal{F}}(f)$.

- (b) Now, let the signal $s(t)$ be as given in the following Figure:



For the following sampling frequencies f_s , indicate (with yes or no) whether or not the signal $s(t)$ can be reconstructed from its samples taken every $T_s = \frac{1}{f_s}$ seconds apart.

- (i) $f_s = 10 \text{ Hz}$
- (ii) $f_s = 12 \text{ Hz}$
- (iii) $f_s = 14 \text{ Hz}$
- (iv) $f_s = 16 \text{ Hz}$
- (v) $f_s = 18 \text{ Hz}$

PROBLEM 5. (Nyquist Criterion) Consider transmitting

$$S(t) = \sum_{i=-\infty}^{\infty} X_i \psi(t - iT)$$

across an AWGN channel, where $\psi(t)$ is a Nyquist pulse. We know that an optimal thing to do at the receiver front end is to send the received signal $R(t)$ through the filter with impulse response $\psi^*(-t)$ and sample the filter output $Y(t)$ at time $t = iT$.

- (a) Show that, in absence of noise, the filter output $Y(iT)$ equals X_i .
- (b) Now assume that you transmit $S(t) = \sum_{i=-\infty}^{\infty} X_i p(t - iT)$ and let the received signal through a filter of real-valued impulse response $q(t)$. You would like to retain the property that, in absence of noise, the filter output at time $t = iT$ be X_i . Show that this is equivalent to

$$\int_{-\infty}^{\infty} p(kT + t)q(-t)dt = \delta(k).$$

- (c) Show that the equivalent condition in the frequency domain is

$$\sum_{l=-\infty}^{\infty} p_{\mathcal{F}}\left(f - \frac{l}{T}\right)q_{\mathcal{F}}^*\left(f - \frac{l}{T}\right) = T \quad \text{for } -\frac{1}{2T} \leq f \leq \frac{1}{2T}.$$

PROBLEM 6. (Minimum Energy Bandpass Signals)

Let the transmitted bandpass signal be of the form

$$x(t) = a \cos\left(2\pi\left(f_c + \frac{1}{T}\right)t\right) + b \cos\left(2\pi\left(f_c + \frac{2}{T}\right)t\right)$$

where a and b are parameters, $a \in \{0, 1\}$ and $b \in \{0, 1\}$.

- (a) Find the baseband equivalent signal $x_E(t)$ for the transmitted signal.
- (b) Draw the constellation for the set of signals corresponding to all choices of a and b in either baseband or bandpass.
- (c) If $a = \{0, 1\}$ equally likely and $b = \{0, 1\}$ equally likely find the average energy of the baseband signal. Is this a minimum energy configuration? If not how can you modify the constellation so that it is minimum energy?

PROBLEM 7. (Signals and Sampling: Multiple Choice Test)

Mark the correct choice.

- (a) Consider the signal $x(t) = \cos(2\pi t) \left(\frac{\sin(\pi t)}{\pi t} \right)^2$. Assume that we sample $x(t)$ with sampling period T . What is the maximum T that guarantees signal recovery?
 a) $T = 1/8$ b) $T = 1/4$ c) $T = 1/2$
- (b) Consider the three signals $s_1(t) = 1$, $s_2(t) = \cos(2\pi t)$, $s_3(t) = \sin^2(\pi t)$, for $0 \leq t \leq 1$. What is the dimension of the signal space spanned by $\{s_1(t), s_2(t), s_3(t)\}$?
 a) 1 b) 2 c) 3
- (c) You are given a pulse $p(t)$ with spectrum $p_{\mathcal{F}}(f) = T(1 - |f|T)$, $0 \leq |f| \leq \frac{1}{T}$. What is the value of $\int p(t)p(t - 3T)dt$? (Hint: First think, then calculate!)
 a) 0 b) $3T$ c) $\frac{1}{3T}$

Appendix 5.A Fourier Series and Sampling Theorem

We briefly review the Fourier series focusing on the big picture and on how to remember things.

Let $f(x)$ be a periodic function, $x \in \mathbb{R}$. It has period p if $f(x) = f(x+p)$ for all $x \in \mathbb{R}$. Its fundamental period is the smallest such p . We are using the “physically unbiased” variable x instead of t (which usually represents time) since we want to emphasize that we are dealing with a general periodic function, not necessarily a function of time.

The periodic function $f(x)$ can be represented as a linear combination of complex exponentials of the form $e^{j2\pi\frac{x}{p}i}$. These are all the complex exponentials that have period p . Hence

$$f(x) = \sum_{i \in \mathbb{Z}} A_i e^{j2\pi\frac{x}{p}i} \quad (5.8)$$

for some sequence of coefficients $\dots A_{-1}, A_0, A_1, \dots$ with value in \mathbb{C} . This says that a function of fundamental period p may be written as a linear combination of all the complex exponentials of period p . You should remember this.

The expression for A_i can also be easily remembered (derived). Two functions of fundamental period p are identical iff they coincide over a period. Hence to check if a given series of coefficients $\dots A_{-1}, A_0, A_1, \dots$ is the correct series, it is sufficient to verify that

$$f(x)1_{[-\frac{p}{2}, \frac{p}{2}]}(x) = \sum_{i \in \mathbb{Z}} \sqrt{p}A_i \frac{e^{j\frac{2\pi}{p}xi}}{\sqrt{p}} 1_{[-\frac{p}{2}, \frac{p}{2}]}(x).$$

Since $\phi_i(x) = \frac{e^{j\frac{2\pi}{p}xi}}{\sqrt{p}} 1_{[-\frac{p}{2}, \frac{p}{2}]}(x)$, $i \in \mathbb{Z}$, is an orthonormal basis, the right side of the above expression is an orthonormal expansion of the left. The coefficients of an orthonormal expansion are always found in the same way, namely

$$\sqrt{p}A_i = \langle f, \phi_i \rangle.$$

Hence

$$A_i = \frac{1}{p} \int_{-\frac{p}{2}}^{\frac{p}{2}} f(x) e^{-j\frac{2\pi}{p}xi} dx. \quad (5.9)$$

We hope that this will make it easier for you to remember (or re-derive) (5.8) and (5.9).

As an example of the utility of this relationship we derive the sampling theorem. Recall that the sampling theorem states that any \mathcal{L}_2 function $s(t)$ which is bandlimited to B may be written as

$$s(t) = \sum_k s(kT) \operatorname{sinc} \left(\frac{t - nT}{T} \right)$$

where $T = \frac{1}{2B}$.

Proof of the sampling theorem: By assumption, $s_{\mathcal{F}}(f) = 0$, $f \notin [-B, B]$. Hence, we may think of $s_{\mathcal{F}}(f)$ as being obtained by multiplying its periodic extension, say $\tilde{s}_{\mathcal{F}}(f)$, with $1_{[-B, B]}(f)$. The periodic extension may be written as a Fourier series, yielding

$$s_{\mathcal{F}}(f) = \tilde{s}_{\mathcal{F}}(f)1_{[-B, B]}(f) = \sum_k A_k e^{+j\frac{2\pi}{2B}fk} 1_{[-B, B]}(f).$$

Taking the Fourier transform on both sides, using

$$1_{[-B, B]}(f) \Leftrightarrow \frac{1}{T} \text{sinc}\left(\frac{t}{T}\right), \quad T = \frac{1}{2B},$$

and the time shift property of the Fourier transform

$$h(t - \tau) \Leftrightarrow h_{\mathcal{F}}(f)e^{-j2\pi f\tau},$$

we obtain

$$s(t) = \sum_k \frac{A_k}{T} \text{sinc}\left(\frac{t + kT}{T}\right).$$

We still need to determine $\frac{A_k}{T}$. It is straightforward to determine A_k from its definition, but it is easier to observe that if we plug in $t = nT$ on both sides of the expression above we obtain $s(nT) = \frac{A_{-n}}{T}$. This completes the proof. To show that it is straightforward, we also determine A_k from the definition:

$$A_k = \frac{1}{2B} \int_{-B}^B s_{\mathcal{F}}(f) e^{-j\frac{2\pi}{2B}kf} df = \frac{1}{2B} \int_{-\infty}^{\infty} s_{\mathcal{F}}(f) e^{-j\frac{2\pi}{2B}kf} df = Ts(-kT),$$

where the first equality is the definition of the Fourier coefficient A_k , the second uses the fact that $s_{\mathcal{F}}(f) = 0$ for $f \notin [-B, B]$, and the third is the inverse Fourier transform evaluated at $t = -kT$. \square

Appendix 5.B Power Spectral Density

In this appendix we derive an expression for the power spectral density of a general signal in the form

$$X(t) = \sum_{i=-\infty}^{\infty} X_i \psi(t - iT - \Theta), \quad (5.10)$$

where $\psi(t)$ is some unit-energy pulse, $\{X_i\}_{i=-\infty}^{\infty}$ is a discrete-time wide-sense stationary sequence, and Θ is a random variable that is uniformly distributed over the interval $[0, T)$. Except for the random delay Θ and the fact that the symbol sequence is infinite, this signal has the general form of symbol-by-symbol on a pulse train. Accounting for a random delay Θ is both realistic and necessary. It is realistic since without it it would imply that an observer of the signal $X(t)$ uses a time reference known to the

sender and that the sender emits the signal in such a way that by the time it reaches the observer it has a predetermined position on the time line. In reality, an observer expects a random shift in time due to clock misalignments but also due to the time it takes for the signal to reach the observer. The reason why we have not mentioned Θ so far is that its presence up until now would have made no difference other than making the notation more cumbersome. Now we need Θ since without it the stochastic process $X(t)$ is not wide-sense stationary and its power spectral density is not defined.

The first step towards the power spectral density is to compute the correlation $R_X(t + \tau, t) := E[X(t + \tau)X^*(t)]$. It will be expressed as a function of

$$R_X[i] = E[X_{j+i}X_j] \quad (5.11)$$

and

$$R_\psi(v) = \int_{-\infty}^{\infty} \psi(\alpha + v)\psi(\alpha)d\alpha : \quad (5.12)$$

$$\begin{aligned} R_X(t + \tau, t) &= E[X(t + \tau)X^*(t)] \\ &= E\left[\sum_{i=-\infty}^{\infty} X_i\psi(t + \tau - iT - \Theta) \sum_{j=-\infty}^{\infty} X_j^*\psi^*(t - jT - \Theta)\right] \\ &= E\left[\sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} X_iX_j^*\psi(t + \tau - iT - \Theta)\psi^*(t - jT - \Theta)\right] \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} E[X_iX_j^*]E[\psi(t + \tau - iT - \Theta)\psi^*(t - jT - \Theta)] \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} R_X[i - j]E[\psi(t + \tau - iT - \Theta)\psi^*(t - jT - \Theta)] \\ &= \sum_{k=-\infty}^{\infty} R_X[k] \sum_{i=-\infty}^{\infty} \frac{1}{T} \int_0^T \psi(t + \tau - iT - \theta)\psi^*(t - iT + kT - \theta)d\theta \\ &= \sum_{k=-\infty}^{\infty} R_X[k] \frac{1}{T} \int_{-\infty}^{\infty} \psi(t + \tau - \theta)\psi^*(t + kT - \theta)d\theta. \end{aligned}$$

Hence

$$R_X(\tau) = \sum_{k=-\infty}^{\infty} R_X[k] \frac{1}{T} R_\psi(\tau - kT), \quad (5.13)$$

where, with a slight abuse of notation, we have written $R_X(\tau)$ instead of $R_X(t + \tau, t)$ to emphasize that $R_X(t + \tau, t)$ depends only on the difference τ between the first and the second variable. It is straightforward to verify that $E[X(t)]$ does not depend on t either. Hence $X(t)$ is a wide-sense stationary process.

The power spectral density S_X is the Fourier transform of R_X . Hence,

$$S_X(f) = \frac{|\psi_{\mathcal{F}}(f)|^2}{T} \sum_k R_X[k] \exp(-j2\pi k f T).$$

Notice that the summation is the discrete-time Fourier transform of $\{R_X[k]\}_{k=-\infty}^{\infty}$ evaluated at fT .

In many cases of interest $\{X_i\}_{i=-\infty}^{\infty}$ is a zero-mean iid sequence. Then $R_X[k] = \mathcal{E}\delta_k$ where $\mathcal{E} = E[|X_j|^2]$ and the formulas simplify to

$$R_X(\tau) = \mathcal{E}R_{\psi}(\tau) \quad (5.14)$$

$$S_X(f) = \mathcal{E} \frac{|\psi_{\mathcal{F}}(f)|^2}{T}. \quad (5.15)$$

EXAMPLE 47. When $\psi(t) = \sqrt{\frac{1}{T}} \text{sinc}(\frac{t}{T})$ and $R_X[k] = \mathcal{E}\delta_k$, the spectrum is $S_X(f) = \mathcal{E}1_{[-B,B]}(f)$, where $B = \frac{1}{2T}$. By integrating the power spectral density we obtain the power $2B\mathcal{E} = \frac{\mathcal{E}}{T}$. This is consistent with our expectation: When we use the pulse $\text{sinc}(\frac{t}{T})$ we expect to obtain a spectrum which is flat for all frequencies in $[-B, B]$ and vanishes outside this interval. The energy per symbol is \mathcal{E} . Hence the power is $\frac{\mathcal{E}}{T}$. \square

Chapter 6

Convolutional Coding and Viterbi Decoding

In Chapter 3 we have seen that each element of the signal constellation can be written as $s_i(t) = \sum s_{ij}\varphi_j(t)$ and we used this fact to derive the MAP (and the ML) receiver for the AWGN channel. At that stage it was convenient to assume that the signal constellation $s_i(t)$, $i = 1, 2, \dots, m$ had been given to us. In so doing we did not have to worry about choosing the signal constellation and it made it clear that the approach we followed is quite general.

In practice, one picks the signal constellation by choosing directly what goes into the right hand side of $s_i(t) = \sum s_{ij}\varphi_j(t)$. In Chapter 4 we have seen that, whenever possible, choosing $\varphi_j(t) = \varphi(t - jT)$ has a number of desirable properties. This is possible if we can find a convenient pulse $\varphi(t)$ such that the resulting $\varphi_j(t) = \varphi(t - jT)$ form an orthonormal set. In Chapter 5 we have seen how to choose $\varphi(t)$ as a function of the spectral slope we want to achieve. In the present chapter we show how to choose the s_{ij} of $s_i(t) = \sum s_{ij}\varphi_j(t)$. Up to a scaling factor, we will let them be the outputs of a convolutional encoder.

The receiver will implement the Viterbi algorithms – a neat and clever way to decode efficiently in many circumstances. To analyze the bit-error probability we will introduce a few new tools, notably detour flow graphs and generating functions.

The signals that we will construct will have the following properties: The transmitter and the receiver adapt naturally to the number k of bits that need to be communicated and the duration of the transmission grows linearly with the number of bits; the bandwidth is constant (independently of the number of bits transmitted) and the encoding and decoding operations are done at low complexity and essentially “on the fly.” As usual, we assume that signals are passed through the AWGN channel.

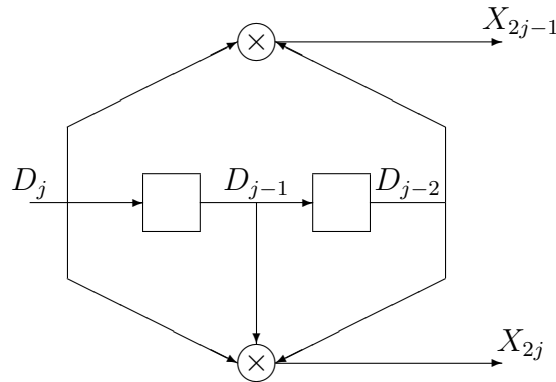


Figure 6.1: Rate $\frac{1}{2}$ convolutional encoder.

6.1 The Transmitter

Like in bit-by-bit on a pulse train, we assume that the entire transmitted signal has the form¹

$$s_i(t) = \sum_{j=1}^n s_{ij} \psi(t - jT),$$

where $i \in \{0, 1, \dots, 2^k - 1\}$ for some integer k ,

$$\begin{aligned} s_{ij} &= x_{ij} \sqrt{E_s} \\ x_{ij} &\in \{\pm 1\}. \end{aligned}$$

In both cases we will go straight from the bit sequence to the symbol sequence without explicitly passing through the message index i that labels $s_i(t)$. Unlike for bit-by-bit on a pulse train, $k < n$. This implies that when the index i runs over all possible 2^k values, the *codeword* $\mathbf{x} = (x_{i1}, \dots, x_{in})$ does *not* range over all possible 2^n n -length sequences with value in $\{\pm 1\}^n$. Only a subset of such sequences are used. This is what coding is all about. It helps since the fact that not all sequences with components taking value in ± 1 produce a valid signal decreases the chances that a transmitted signal plus noise looks more like an alternative signal rather than the original. In bit-by-bit on a pulse train we would not speak of coding since any n tuple in $\{\pm \mathcal{E}_b\}^n$ is a valid signal space point. For block-orthogonal signaling all components of \mathbf{s}_i except one are zero. This is a form of coding.

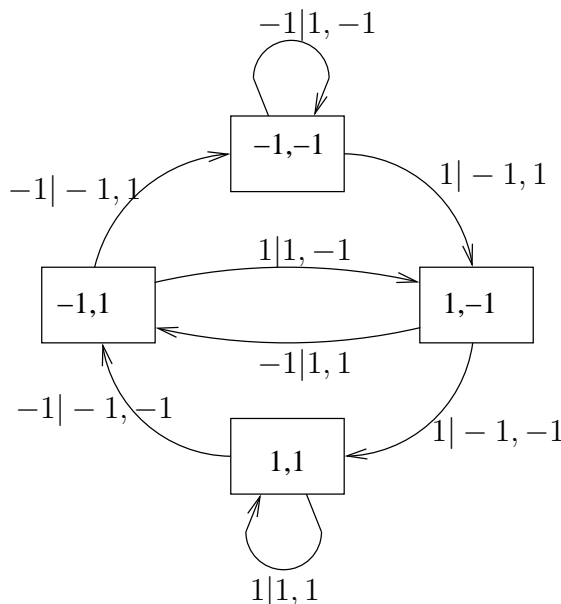
The next step is to specify the signal space points that we use. We specify our 2^k sequences x_{i1}, \dots, x_{in} by means of an *encoder*. For our encoder, depicted in Figure ??, $n = 2k$.

The encoder input is a uniformly distributed k -length random vector $\mathbf{D} = (D_1, \dots, D_k)$ that models the information bits (the source output) to be transmitted. The k encoder

¹For most of what we do in this chapter we could assume that the signal has the general form $s_i(t) = \sum_{j=1}^n s_{ij} \psi_j(t)$. We prefer to be more specific and focus to situations of real practical interest.

output pairs are serialized to form an n -length random sequence $\mathbf{X} = (X_1, X_2, \dots, X_n)$ used to modulate the pulse train. Hence $\mathbf{X}\sqrt{E_b}$ is the corresponding signal space point. The multiplication symbol in Figure ?? represents the usual multiplication over \mathbb{R} . However, this operation is actually the addition over the field that one can define on the set $\{\pm 1\}$. Hence the encoder is linear. More on this in problem 6.

There are alternative ways to describe the encoder. One such alternative way that will turn out to be useful in determining the error probability is by means of the state diagram shown below.



This is a four state machine. Each box represents a state. States are labeled by the content of the two shift registers of the encoder (previous picture). Arrows represent transitions. There are two possible transitions from each state. The input symbol D_j decides which of the two possible transitions is taken at time j . Transitions are labeled by $D_j|X_{2j-1}, X_{2j}$.

We agree that initially the encoder is in some arbitrary but fixed state. We assume that this state is $(1, 1)$.

The following example shows a source output sequence of length $k = 5$ and the corresponding encoder output sequence of length $n = 10$.

D_j	1	-1	-1	1	1
X_{2j-1}, X_{2j}	1, 1	-1, -1	-1, 1	-1, 1	-1, -1
j	1	2	3	4	5

6.2 The Receiver

Let $\|s_i\|^2 = \sum_{j=1}^n E_s x_{ij}^2 = nE_s$ be the signal's energy (the same for all signals).

A maximum likelihood (ML) decoder decides for one of the i that maximizes

$$\langle r, s_i \rangle - \frac{nE_s}{2},$$

where the second term is irrelevant since it does not depend on i and r is the received signal.

Hence the ML decoder picks (one of) the sequence(s) s_{i1}, \dots, s_{in} that maximizes

$$\begin{aligned} & \int r(t) \sum_{j=1}^n s_{ij} \psi(t - jT) dt \\ &= \sum_{j=1}^n s_{ij} \int r(t) \psi(t - jT) dt \\ &= \sum_{j=1}^n s_{ij} y_j \\ &= \sqrt{E_s} \sum_{j=1}^n x_{ij} y_j \end{aligned}$$

where we have defined

$$y_j = \int r(t) \psi(t - jT) dt.$$

Recall that y_j is the output at time jT of the filter with impulse response $\psi(-t)$ and input $r(t)$.

The difficulty in finding one of the i that maximizes $\langle \mathbf{x}_i, \mathbf{y} \rangle$ is that it appears at first that we have to test all 2^k such inner products. Typically k is much larger than 100. If $k = 100$, with a computer that does 10^9 inner products $\langle \mathbf{x}_i, \mathbf{y} \rangle$ in a second it takes roughly (using 2^{30} to approximate 10^9) $2^{100}/2^{30} = 2^{70}$ seconds to compute all inner products. This makes almost 4×10^{13} years. The universe is only 20×10^9 years old!

What we need is a method that finds the maximum $\langle \mathbf{x}, \mathbf{y} \rangle$ by making a number of operations that grows linearly (as opposed to exponentially) in k . By cleverly making use of the encoder structure we will see that this can be done. The result is the Viterbi algorithm.

To describe the Viterbi algorithm (VA) we need the notion of a trellis. The trellis is an unfolded transition diagram that keeps track of the passage of time. If we assume that we start at state $(1, 1)$, that we transmit $k = 5$ source digits, and that we append 2 *dummy* source digits to ensure that at the end of the transmission the encoder is again in state

(1, 1), the trellis diagram looks as in the top drawing of figure 6.2. For each index i , there is exactly one path in the trellis between the *root* (the leftmost node) and the *toor*² (the rightmost node) which is labeled by the sequence \mathbf{x}_i .

Unlike in the state diagram, in the mentioned trellis we have labeled transitions with the encoder output symbols only. We have dropped the encoder input from the label since we have ordered states in such a way that the lower branch leaving a state at depth j corresponds to $D_j = 1$. Hence it is straightforward to associate a path in the trellis with the corresponding source sequence.

To decode we do the following. Let $\mathbf{y} = (y_1, y_2 \dots y_n)^T$ be the n tuple of matched filter output symbols. Each \mathbf{x}_i corresponds to a path in the trellis and the sequence of labels that we read out along that path is exactly \mathbf{x}_i . We use \mathbf{y} to relabel the path in the trellis corresponding to \mathbf{x}_i . Specifically instead of $x_{i,2j-1}, x_{i,2j}$ we now write the *branch metric* $\langle (x_{i,2j-1}x_{i,2j})^T, (y_{2j-1}, y_{2j})^T \rangle$. Then, by adding the path metric along the path corresponding to \mathbf{x}_i we obtain $\sum_j \langle (x_{i,2j-1}x_{i,2j})^T, (y_{2j-1}, y_{2j})^T \rangle = \langle \mathbf{x}_i, \mathbf{y} \rangle$. We call this the *path metric*. The path metric is the sum of the branch metric along a specific path between root and toor. The second trellis in Fig. 6.2 has been labeled with branch metrics.

Maximum likelihood decoding now amounts to finding (one of) the path(s) with the largest path metrics. The VA is an efficient way to do this. It is convenient to think of the trellis as a road map with branch metrics instead of distances. The job of the VA is to find the longest (not the shortest) path between root and toor. How to do this is best explained by means of an example. (See example in class and bottom two trellises in Fig. 6.2).

²Toor is root read backwards.

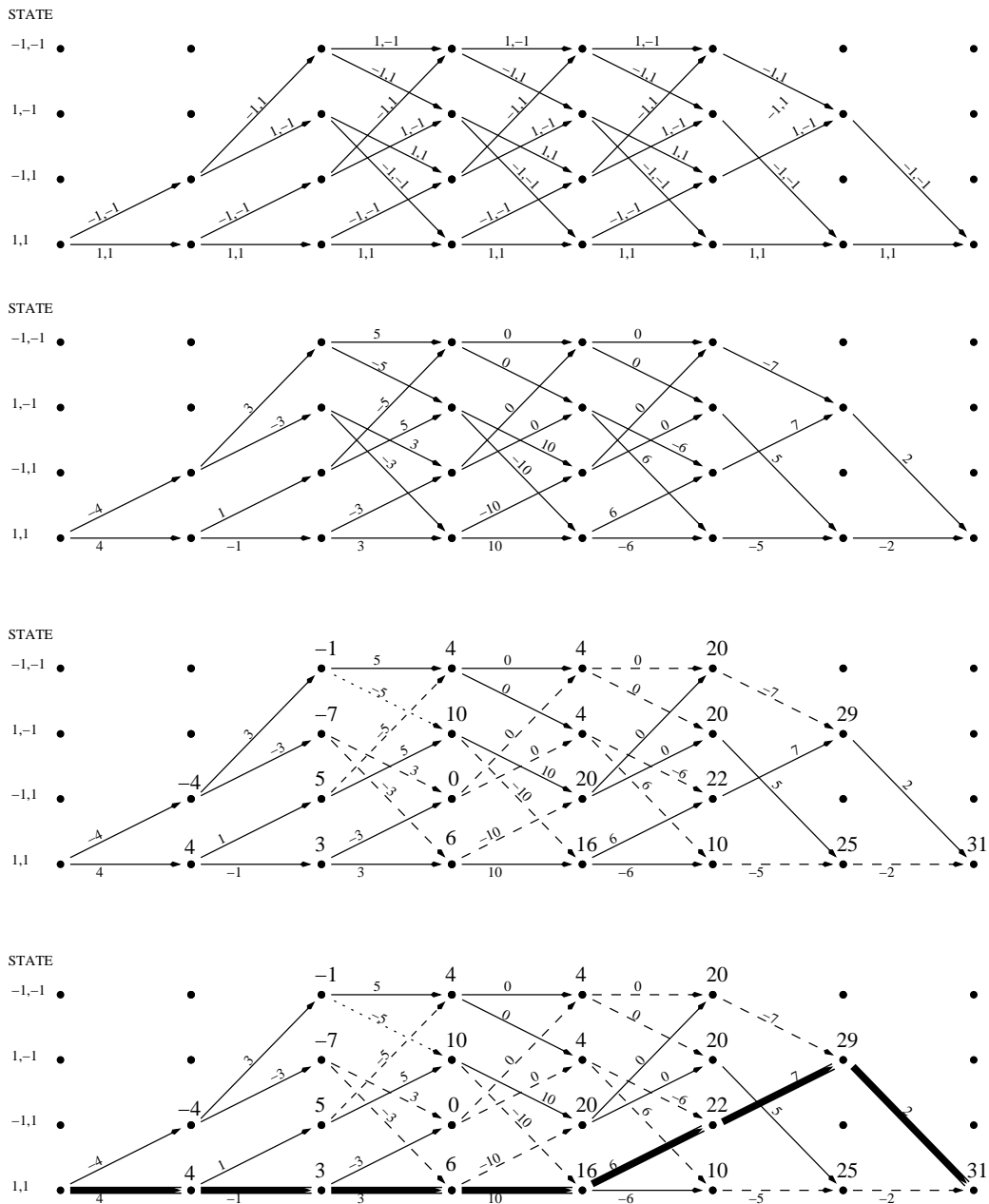


Figure 6.2: The Viterbi algorithm. Top figure: Trellis representing the encoder. The upper transition leaving a state corresponds to source symbol -1 , the lower transition to source symbol 1 . Transitions are labeled with the corresponding output symbols; Second figure: Transitions have been labeled with the branch metric corresponding to the received sequence $(1, 3), (-2, 1), (4, -1), (5, 5), (-3, -3), (1, -6), (2, -4)$ (parentheses have been inserted for convenience only); Third figure: Each state has been labeled with the metric of the surviving path and non-surviving transitions have been dashed; Fourth figure: Tracing back from the end we find the decoded path (bold). It corresponds to source sequence $1, 1, 1, 1, -1, 1, 1$.

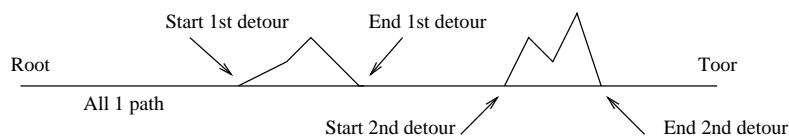
6.3 Bit-Error Probability

We assume that the initial state is $(1, 1)$ and we transmit a number k of bits.

As we have done so far, we determine (an upper bound to) the probability of error by first conditioning on a fixed transmitted signal. It will turn out that our expression does not depend on the selected transmitted signal.

Each signal that can be produced by the transmitter has a corresponding trellis path. The actual path that we assume as being the correct one when we compute (or when we bound) the bit error probability will be called the *reference* path. For now, the reference path is the *all-one* path. This is the path generated by the all-one source symbols. The corresponding encoder output sequence also consists of all ones. In this section we assume that trellis paths are labeled by the corresponding *input* sequence.

The first new concept needed is that of a *detour*. Detours are those segments of the path selected by the Viterbi decoder that do not correspond to the reference path.³ A detour starts at a node where the decoded path diverges from the reference path and ends at some later node where the decoded path merges again with the reference path. (See the figure below.)



6.3.1 Counting Detours

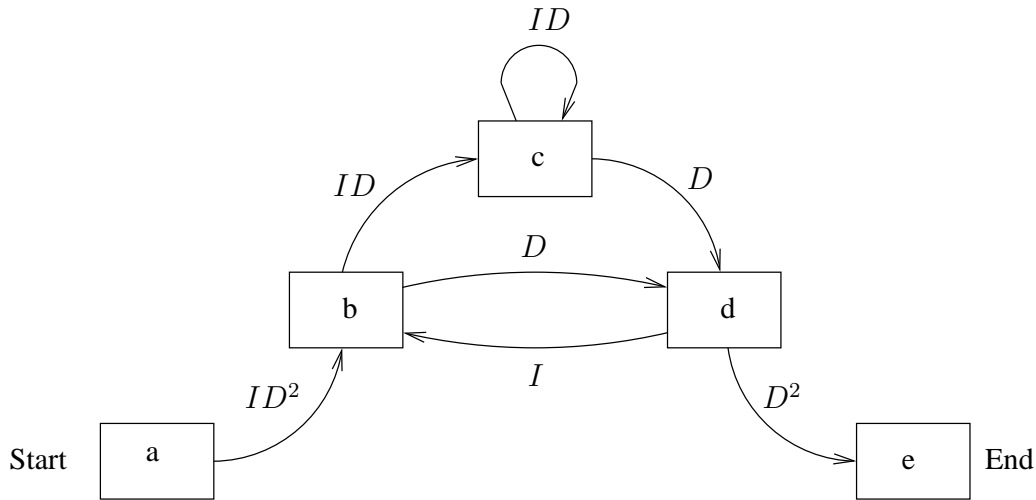
The basic idea to determine the bit error probability is simple. For each detour we determine the probability that the Viterbi decoder takes this detour and the number of information bit errors i we make when this happens. These two quantities allow us, in principle, to determine the average bit-error probability. We will actually work with an upper bound to the probability that the Viterbi decoder takes a given detour. For a given detour, our upper bound will depend on the number d of discrepancies between the encoder output sequence corresponding to the detour and that of the reference path.

EXAMPLE 48. *Regardless of the reference path, for our example there is a shortest detour starting at any state in the trellis (provided we are sufficiently away from the final state to avoid "edge effect"). This shortest detour spans 3 trellis sections. (A trellis section is the portion of the trellis between all states at one depth and all states at the next depth). The corresponding parameters are $i = 1$ and $d = 5$.* \square

³For an analogy, think of the trellis as a road map, of the reference path as of an intended road for your journey, and the path selected by the Viterbi decoder as of the actual road that you take. Once on a while you are forced to take a detour with respect to the intended road.

To avoid end effects, in this section we will assume that the trellis is semi infinite, i.e., it extends to infinity to the right. For any given point (depth) on the reference path, how many detours are there with given parameters i and d ? We now proceed to find this number, denoted by $a(i, d)$. It will be one of the main ingredient in determining our upper bound to the bit error probability.

There is a one-to-one correspondence between a detour with respect to the all-one path and a path between state a and e in the following *detour flow graph*.



The label $I^i D^d$, (i and d nonnegative integers), on a transition denotes that this transition increases the discrepancies between the two input sequences (reference path and detour) by i and between the two output sequences by d .

Now we show how to determine $a(i, d)$. We will actually determine the *generating function* $T(I, D)$ of $a(i, d)$ defined as

$$T(I, D) = \sum_{i,d} I^i D^d a(i, d)$$

You should think of I and D as “place holders” without any physical meaning. It is like describing the coefficients a_0, a_1, \dots, a_{n-1} by means of the polynomial $p(x) = a_0 + a_1 x + \dots + a_{n-1} x^{n-1}$. In our case, as we will see, having the generating function $T(I, D)$ is more convenient than having $a(i, d)$ for all i and all d . We determine $T(I, D)$ recursively as follows. We introduce auxiliary generating functions, one for each intermediate state of the detour flow graph, namely:

$$T_b(I, D) = \sum_{i,d} I^i D^d a_b(i, d) \quad (6.1)$$

$$T_c(I, D) = \sum_{i,d} I^i D^d a_c(i, d) \quad (6.2)$$

$$T_d(I, D) = \sum_{i,d} I^i D^d a_d(i, d) \quad (6.3)$$

As $a(i, d)$ is the number of paths in the detour flow graph that start at Start, end at End, and have parameters i and d , so $a_b(i, d)$ is the number of paths in the detour flow graph that start at Start, end at b , and have parameters i and d . Similar definitions apply for $a_c(i, d)$ and $a_d(i, d)$. From the detour flow graph, we immediately see that the following relationships hold:

Writing T_c instead of $T_c(I, D)$ we have the following relationships:

$$\begin{aligned} T_b &= T_a ID^2 + T_d I \\ T_c &= T_b ID + T_c ID \\ T_d &= T_b D + T_c D \\ T &= T_d D^2 \end{aligned}$$

The above system may be solved for T by pure formal manipulations. (Like solving a system of equations). The result is

$$T(I, D) = \frac{ID^5}{1 - 2ID}.$$

The above expression $T(I, D)$ is what we need. However, to show that one can indeed obtain $a(i, d)$ from $T(I, D)$, using the expansion $\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots$ we write

$$T(I, D) = \frac{ID^5}{1 - 2ID} = ID^5(1 + 2ID + (2ID)^2 + (2ID)^3 + \dots) \quad (6.4)$$

$$= ID^5 + 2I^2D^6 + 2^2I^3D^7 + 2^3I^4D^8 + \dots \quad (6.5)$$

This means that there is one path with parameters $d = 5$, $i = 1$, that there are two paths with $d = 6$, $i = 2$, etc. In general, for $i = 1, 2, \dots$ we have

$$a(i, d) = \begin{cases} 2^{i-1}, & d = i + 4 \\ 0, & \text{otherwise.} \end{cases}$$

You can verify on the detour flow graph that there is a detour with $i = 1$ and $d = 5$, two detours with $d = 2$ $d = 6$, four with $i = 3$, $d = 7$, etc. Next we show that $a(i, d)$ does not depend on the reference path, provided that the encoder is linear.

Let \mathcal{D}^* be the set of finite length strings of symbols in $\{\pm 1\}$. Let $\mathbf{a} \in \mathcal{D}^*$, be the reference encoder-input sequence. Let $f : \mathcal{D}^* \rightarrow \mathcal{D}^*$ be the encoder map. Hence $f(\mathbf{a})$ is the encoder output that corresponds to input sequence \mathbf{a} .

For a sequence $\mathbf{a} \in \mathcal{D}^*$ and positive integers k, l , we define $\mathbf{a}_k^l = (a_k, a_{k+1}, \dots, a_l)$ to be the subsequence that starts with index k and ends with index l , $k \leq l$. The set of all such sequences is denoted by $\{\pm 1\}_k^l$.

We will be interested in knowing how many “-1” are contained in a given sequence $\mathbf{a} \in \{\pm 1\}_k^l$. Let $w(\mathbf{a})$ denote this number.

$$w(\mathbf{a}) = | \{j : a_j = -1\} |$$

Let $\mathcal{A}_{j,i,d}$ be the set of all treillis paths that contain a detour that leave the all-one sequence at depth j ($j = 0, 1, \dots$) and has parameters i and d , i.e.

$$\mathcal{A}_{j,i,d} = \{\mathbf{c} \in \{\pm 1\}_1^\infty : w(\mathbf{c}_1^j) = 0, c_{j+1} = -1, w(\mathbf{c}) = i, w(f(\mathbf{c})) = d\}$$

Then $a(i, d) = |\mathcal{A}_{j,i,d}|$, where the right hand side does not depend on the depth j since the encoder is time-invariant in the following sense. If input $\mathbf{a} = a_1, a_2, \dots$ produces a detour to the all-one path that starts at depth j and has parameters i and d , then the input $1, \mathbf{a} = 1, a_1, a_2, \dots$ produces a detour that starts at depth $j+1$ and has parameters i and d .

Similarly, define $\mathcal{A}_{j,i,d}(\mathbf{a})$ to be the set of all treillis paths that contain a detour that starts at time i and has parameters i and d when we consider the path label $f(\mathbf{a})$ as the reference path:

$$\mathcal{A}_{j,i,d}(\mathbf{a}) = \{\mathbf{e} \in \mathcal{D}^* : w((\mathbf{e}\mathbf{a})_1^j) = 0, (\mathbf{e}\mathbf{a})_{j+1} = -1, w(\mathbf{e}\mathbf{a}) = i, w(f(\mathbf{e})f(\mathbf{a})) = d\},$$

where $\mathbf{e}\mathbf{a}$ denotes the componentwise product of \mathbf{e} and \mathbf{a} .

Let $a_{j,i,d}(\mathbf{a}) \triangleq |\mathcal{A}_{j,i,d}(\mathbf{a})|$ be the number of such detours. We want to show that for all $\mathbf{a} \in \mathcal{D}^*$, $a_{j,i,d}(\mathbf{a}) = a(i, d)$. It suffices to show that there exists a one-to-one correspondence between the elements of $\mathcal{A}_{j,i,d}(\mathbf{a})$ and those of $\mathcal{A}_{j,i,d}$. We claim that the mapping

$$g : \mathcal{A}_{j,i,d}(\mathbf{a}) \rightarrow \mathcal{A}_{j,i,d}$$

that sends $\mathbf{e} \in \mathcal{A}_{j,i,d}(\mathbf{a})$ to $\mathbf{e}\mathbf{a}$ is such a correspondence.

If we let $\mathbf{c} = \mathbf{e}\mathbf{a}$ and use the definition of $\mathcal{A}_{j,i,d}(\mathbf{a})$ and the linearity of f which implies $f(\mathbf{e})f(\mathbf{a}) = f(\mathbf{e}\mathbf{a}) = f(\mathbf{c})$, we see immediately that $\mathbf{c} \in \mathcal{A}_{j,i,d}$. Now let $\mathbf{c} \in \mathcal{A}_{j,i,d}$. The inverse mapping g^{-1} maps \mathbf{c} to $\mathbf{e} = \mathbf{c}\mathbf{a}$. Using the definition of $\mathcal{A}_{j,i,d}$, the fact that $\mathbf{e}\mathbf{a} = \mathbf{c}\mathbf{a}\mathbf{a} = \mathbf{c}$ and the linearity of f which implies $f(\mathbf{e})f(\mathbf{a}) = f(\mathbf{e}\mathbf{a}) = f(\mathbf{c})$ we immediately verify that $\mathbf{e} \in \mathcal{A}_{j,i,d}(\mathbf{a})$. This completes the proof.

6.4 Upper Bound to P_b

We are now ready for the final step, namely the derivation of a tight upper bound to the bit-error probability.

Fix an arbitrary encoder input sequence, let $\mathbf{x} = x_1, x_2, \dots, x_n$ be the corresponding encoder output sequence and $\mathbf{s} = \sqrt{E_s}\mathbf{x}$ be the corresponding point in signal space. The transmitted signal is

$$s(t) = \sum s_i \psi(t - iT).$$

We transmit this signal over an AWGN channel with power spectral density $N_0/2$. Let $r(t) = s(t) + z(t)$ be the received signal (where $z(t)$ is a sample path of the noise process $Z(t)$) and let

$$\mathbf{y} = (y_1, \dots, y_n)^T, \quad y_i = \langle r, \psi_i \rangle$$

be a sufficient statistic.

The Viterbi algorithm labels each branch in the trellis with the corresponding branch metric and finds the path through the trellis with the largest path metric. A branch from depth j to $j + 1$ with output symbols x_{2j-1}, x_{2j} is assigned the branch metric $y_{2j-1}x_{2j-1} + y_{2j}x_{2j}$.

The ML path selected by the Viterbi decoder may contain several detours. Let w_j , $j = 0, 1, \dots, k-1$, be the number of errors made on a detour that *begins* at depth j . If at depth j the VD is on the correct path or if it follows a detour started earlier then we let $w_j = 0$. Let W_j be the corresponding random variable (over all possible noise realizations).

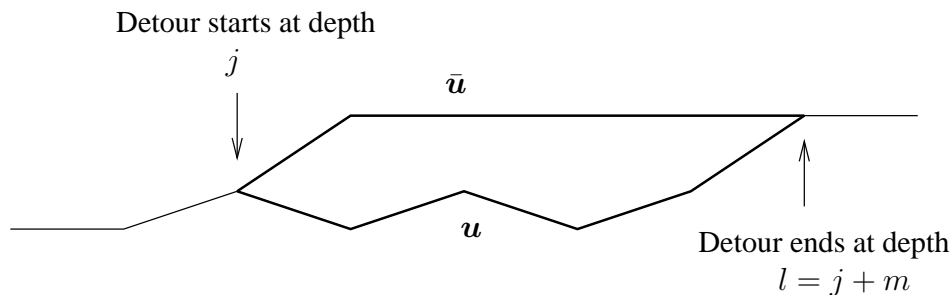
The total of source symbol errors for the path selected by the VD is

$$\sum_{j=0}^{k-1} w_j$$

and $\frac{1}{k} \sum_{j=0}^{k-1} w_j$ is the fraction errors with respect to the k source symbols. Hence we define the bit-error probability

$$P_b \triangleq E \frac{1}{k} \left[\sum_{j=0}^{k-1} W_j \right] = \frac{1}{k} \sum_{j=0}^{k-1} E W_j$$

Let us now focus on a detour. If it starts at depth j and ends at depth $l = j + m$, then the corresponding encoder-output symbols are some $2m$ tuple $\bar{\mathbf{u}} \in \{\pm 1\}^{2m}$. Let $\mathbf{u} = (x_{2j}, \dots, x_{2l-1})^T \in \{\pm 1\}^{2m}$ be the corresponding sub-sequence of the correct path and $\boldsymbol{\rho} = (y_{2j}, \dots, y_{2l-1})^T$ the corresponding channel output subsequence.



Let d be the number of positions in which \mathbf{u} and $\bar{\mathbf{u}}$ differ (also called Hamming distance $d(\mathbf{u}, \bar{\mathbf{u}})$ between \mathbf{u} and $\bar{\mathbf{u}}$).

Notice that the Euclidean distance between the corresponding waveforms is the distance between $\sqrt{E_s}\mathbf{u}$ and $\sqrt{E_s}\bar{\mathbf{u}}$ which is $d_E = 2\sqrt{E_s d}$.

A necessary (but not sufficient) condition for the Viterbi decoder to take the detour under consideration is that

$$\langle \boldsymbol{\rho}, \sqrt{E_s}\mathbf{u} \rangle \leq \langle \boldsymbol{\rho}, \sqrt{E_s}\bar{\mathbf{u}} \rangle.$$

This condition is satisfied iff

$$\| \boldsymbol{\rho} - \sqrt{E_s} \mathbf{u} \|^2 \geq \| \boldsymbol{\rho} - \sqrt{E_s} \bar{\mathbf{u}} \|^2.$$

If \mathbf{u} is the correct sequence and $\bar{\mathbf{u}}$ is the only alternative, then the above event happens with probability

$$Q\left(\frac{d_E}{2} \frac{1}{\sigma}\right)$$

where $\sigma^2 = \frac{N_0}{2}$ and $d_E = 2\sqrt{E_s d}$.

Recall (the Bhattacharyya bound for the Gaussian channel) that

$$Q\left(\frac{d_E}{2\sigma}\right) = Q\left(\sqrt{\frac{2E_s d}{N_0}}\right) \leq e^{-\frac{E_s d}{N_0}} = z^d,$$

where we have defined $z = e^{-\frac{E_s}{N_0}}$.

For the purpose of the next step in the calculation, let us imagine that we enumerate all detours that leave at depth j the correct path. Explicitly, if there are L_{max} such detours, let us assign to each detour a unique index from the set $\{1, 2, \dots, L_{max}\}$ and let $i(l)$ be the Hamming distance between the information subsequence that corresponds to detour l and that of the correct path. Let $d(l)$ be the Hamming distance between the encoder output subsequence that corresponds to the l th detour and that of the correct patch. With this notation,

$$EW_j = \sum_{l=1}^{L_{max}} i(l) \pi_l$$

where π_l stands for the probability that detour l is taken. Using $\pi_l \leq z^{d(l)}$ we obtain

$$\begin{aligned} EW_j &\leq \sum_{l=1}^{L_{max}} i(l) z^{d(l)} \\ &= \sum_{i=1}^{\infty} \sum_{d=0}^{\infty} i z^d a(i, d), \end{aligned}$$

where $a(i, d) = |\{l : i(l) = i, d(l) = d\}|$. Using the relationship

$$\sum_{i=1}^{\infty} i f(i) = \frac{\partial}{\partial I} \sum_{i=0}^{\infty} I^i f(i) \Big|_{I=1},$$

which holds for any function f , we may write

$$\begin{aligned} EW_j &\leq \frac{\partial}{\partial I} \sum_{i=1}^{\infty} \sum_{d=0}^{\infty} I^i z^d a(i, d) \Big|_{I=1} \\ &= \frac{\partial}{\partial I} T(I, D) \Big|_{I=1, D=z}. \end{aligned}$$

Since the upperbound to EW_j does not depend on j , in the following expression it can be moved in front of the sum to obtain

$$P_b = \frac{1}{k} \sum_{j=0}^{k-1} EW_j \leq \frac{\partial}{\partial I} T(I, D) \Big|_{I=1, D=z}.$$

This result can be generalized to an encoder that, in each trellis section, takes k_0 information symbols and produces n_0 channel symbols, $n_0 \geq k_0$ (in our example $k_0 = 1$ and $n_0 = 2$), and for any memoryless channel (not just the AWGN).

In our particular example

$$T(D, I) = \frac{ID^5}{1 - 2ID}$$

$$\frac{\partial T}{\partial I} = \frac{D^5}{(1 - 2ID)^2}.$$

Thus

$$P_b \leq \frac{z^5}{(1 - 2z)^2},$$

where $z = e^{-\frac{E_b}{2N_0}}$ and $E_s = \frac{E_b}{2}$ (we are transmitting two channel symbols per information bit).

6.5 Concluding Remarks

What have we done and how does it compare to what we have done before?

It is convenient to think of the bit-by-bit as our starting point.

$$s(t) = \sum_{i=0}^{n-1} s_i \psi(t - iT)$$

$$s_i \in \{\pm \sqrt{E_s}\}.$$

The relevant design choices for this system are:

$$R_b = R_s = \frac{1}{T} \quad \text{bit rate}$$

$$E_b = E_s \quad \text{energy per bit}$$

$$P_b = Q\left(\frac{\sqrt{E_s}}{\sigma}\right), \quad \text{bit-error probability.}$$

Using $\sigma = \sqrt{\frac{N_0}{2}}$ and the upper bound $Q(x) \leq e^{-\frac{x^2}{2}}$ we obtain

$$P_b \leq e^{-\frac{E_b}{N_0}}.$$

The novelty of this chapter was to have an encoder in front of the bit-by-bit on a pulse train. The encoder trades P_b for R_b . The new parameters are:

$$\begin{aligned} R_b &= \frac{R_s}{2} = \frac{1}{2T_s} \\ E_b &= \frac{E_s}{r} = 2E_s \\ P_b &\leq \frac{z^5}{(1-2z)^2} \text{ where } z = e^{-\frac{E_b}{2N_0}}. \end{aligned}$$

where r is the dimensionless rate of the encoder in $\frac{\text{bits}}{\text{symbol}}$ ($r = \frac{1}{2}$ in our example). It should be emphasized that source symbols are called *bits* since they carry one bit of information each. The notion of bits as a unit of information will be introduced to you in the Information Theory class (7th semester). For now it suffices to say that a binary random variable carries one bit of information iff it is uniformly distributed, and each symbol of a sequence of binary random variables carries one bit of information if, in addition, the symbols are i.i.d.

For a fixed encoder output rate, the power spectral density of the transmitted signal is the same as that of an uncoded system with the same symbol rate (see Homework). However, the encoder output has a symbol rate which is twice that of the encoder input rate. Hence, in effect, we are occupying twice as much bandwidth. We have reduced the bit-error probability but we have increased the bandwidth by a factor two. With more powerful codes we can further decrease the bit error probability without further expanding the bandwidth. As already mentioned in the previous chapter, there is a fundamental limit, studied in information theory (seen next semester) that says that we can make the error probability arbitrarily small as long as the bit rate is smaller than a computable number called *channel capacity*. For the AWGN channel with bandwidth B , the channel capacity is $B \log_2 \left(1 + \frac{P}{BN_0} \right)$ [bits/sec], where P is the transmitted power. The probability of error may be made arbitrarily small by a clever choice of signaling method. In fact, across a bandlimited channel, one may always transmit signals of the form

$$s_i(t) = \sum_j s_{ij} \psi(t - jT)$$

where $\psi(t) = \frac{1}{\sqrt{T}} \text{sinc}(\frac{t}{T})$. This follows from the sampling theorem. In general, however, the s_{ij} is not constrained to be of the form $\{\pm\sqrt{E_s}\}$. Like in the example of this chapter, the signal space points used to make the probability of error arbitrary small are obtained from an encoder.

As already mentioned, the encoder is just meant to form appropriate signal points in n dimensions, where n is typically large. Intuitively speaking, here is what a large n buys us. Let us start first from the opposite situation and let us assume that we are operating in one dimension like in pulse amplitude modulation (PAM). Independently of the transmitted point s , the received point $y = s + Z$, can be anywhere in \mathbb{R} (certain regions are more likely than other and if we choose the decoding region according to the

MAP rule then we minimize the error probability, but there is a limit to how small we can make the probability of error). If we move to n dimensions, then we send some $\mathbf{s} \in \mathbb{R}^n$ and receive $\mathbf{y} = \mathbf{s} + \mathbf{Z}$, where $\mathbf{Z} \sim \mathcal{N}(0, I_n \sigma^2)$. By the law of large numbers, $\sqrt{\frac{1}{n} \sum z_i^2}$ goes to σ as n goes to infinity. This means that with probability going to 1, \mathbf{Y} will be in a thin shell of radius $\sqrt{n}\sigma$ around \mathbf{s} . Intuitively, the key here is that in n dimensional space there are points in \mathbb{R}^n that are “off limit” for $\mathbf{Y} = \mathbf{s} + \mathbf{Z}$. This was not the case for the corresponding one-dimensional problem. By choosing the signal points cleverly, there is a hope that we can find a large number of such points that can be distinguished from one another with high probability even after they have been sent through the channel.

6.6 Problems

PROBLEM 1. (Power Spectral Density of the convolutional code that was considered in class.)

Block-orthogonal signaling may be the simplest coding method that achieves $\Pr\{e\} \rightarrow 0$ as $N \rightarrow \infty$ for a non-zero data rate. However, we have seen in class that the price to pay is that block-orthogonal signaling requires infinite bandwidth to make $\Pr\{e\} \rightarrow 0$. This may be a small problem for one space explorer communicating to another; however, for terrestrial applications, there are always constraints on the bandwidth consumption. Therefore, in the examination of any coding method, an important issue is to compute its bandwidth consumption. Compute the bandwidth occupied by the rate-1/2 convolutional code studied in this chapter. The signal that is put onto the channel is given by

$$X(t) = \sum_{i=-\infty}^{\infty} X_i \sqrt{E_s} \psi(t - iT_s), \quad (6.6)$$

where $\psi(t)$ is some unit-energy function of duration T_s and we assume that the trellis extends to infinity on both ends, but as usual we actually assume that the signal is the wide-sense stationary signal

$$\tilde{X}(t) = \sum_{i=-\infty}^{\infty} X_i \sqrt{E_s} \psi(t - iT_s - T_0), \quad (6.7)$$

where T_0 is a random delay which is uniformly distributed over the interval $[0, T_s)$.

(i) Find the expectation $E[X_i X_j]$ for $i = j$, for $(i, j) = (2n, 2n + 1)$ and for $(i, j) = (2n, 2n + 2)$ for the convolutional code that was studied in class. Then give the autocorrelation function $R_X[i - j] = E[X_i X_j]$ for all i and j . Hint: Consider the infinite trellis of the code. Recall that the convolution code studied in the class can be defined as

$$\begin{aligned} X_{2n} &= D_n D_{n-2} \\ X_{2n+1} &= D_n D_{n-1} D_{n-2} \end{aligned}$$

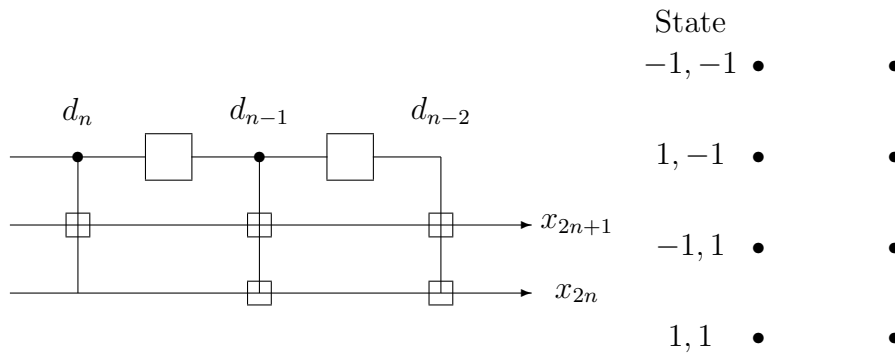


Figure 6.3: (2, 3) convolutional encoder: $x_{3n} = d_n d_{n-2}$; $x_{3n+1} = d_{n-1} d_{n-2}$; $x_{3n+2} = d_n d_{n-1} d_{n-2}$.

(ii) Find the autocorrelation function of the signal $\tilde{X}(t)$, that is

$$R_{\tilde{X}}(\tau) = E[\tilde{X}(t)\tilde{X}(t + \tau)] \tag{6.8}$$

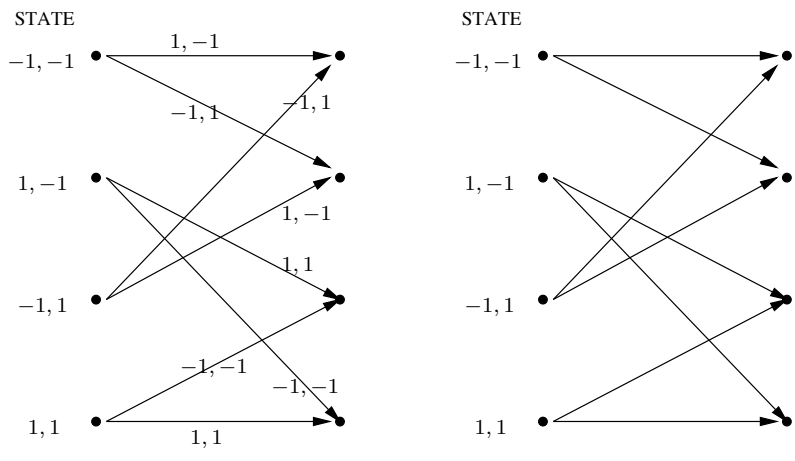
in terms of $R_X[k]$ and $R_\psi(\tau) = \frac{1}{T_s} \int_0^{T_s} \psi(t + \tau)\psi(t)dt$.

(iii) Give the expression of power spectral density of the signal $\tilde{X}(t)$.

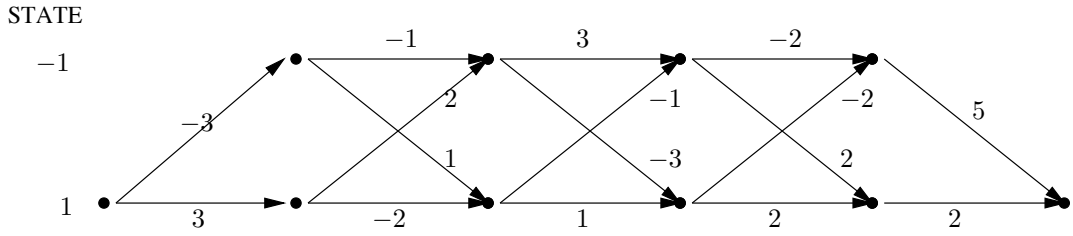
(iv) Find and plot the power spectral density that results when $\psi(t)$ is a rectangular pulse of width T_s centered at 0.

PROBLEM 2. For the convolutional encoder shown below on the left, fill in the section of the trellis shown below on the right, that is, find the correct arrows and label them with the corresponding output value pairs (x_{2n}, x_{2n+1}) . The input sequence d_n takes values in $\{\pm 1\}$.

PROBLEM 3. Consider the convolutional code described by the trellis section below on the left. You may assume that each of the encoder output symbols (x_{2n}, x_{2n+1}) , are mapped into orthogonal waveforms, $\phi_1(t)$ if $x_i = +1$ and $\phi_2(t)$ if $x_i = -1$. The waveforms are of equal energy E_s . At the receiver we perform matched filtering with the filters matched to $\phi_1(t)$ and $\phi_2(t)$. Suppose the output of the matched filter at time n are $(y_{2n}, y_{2n+1}) = (1, -2)$. Find the branch metric values to be used by the Viterbi algorithm and enter them into the trellis section on the right.



PROBLEM 4. In the trellis below, the received sequence has already been preprocessed. The labels on the edges of the trellis are the branch metric values. Find the maximum likelihood path.



PROBLEM 5. (Intersymbol Interference)

An information sequence $\underline{U} = (U_1, U_2, \dots, U_5)$, $U_i \in \{0, 1\}$ is transmitted over a noisy intersymbol interference channel. The i th sample of the receiver-front-end filter (e.g. a filter matched to the pulse used by the sender)

$$Y_i = S_i + Z_i,$$

where the noise Z_i forms an independent and identically distributed (i.i.d.) sequence of Gaussian random variables,

$$S_i = \sum_{j=0}^{\infty} U_{i-j} h_j, \quad i = 1, 2, \dots$$

and

$$h_i = \begin{cases} 1, & i = 0 \\ -2, & i = 1 \\ 0, & \text{otherwise.} \end{cases}$$

You may assume that $U_i = 0$ for $i \geq 6$ and $i \leq 0$.

(a) Rewrite S_i in a form that explicitly shows by which symbols of the information sequence it is affected.

(b) Sketch a trellis representation of a finite state machine that produces the output sequence $\underline{S} = (S_1, S_2, \dots, S_6)$ from the input sequence $\underline{U} = (U_1, U_2, \dots, U_5)$. Label each trellis transition with the specific value of $U_i|S_i$.

(c) Specify a metric $f(\underline{s}, \underline{y}) = \sum_{i=1}^6 f(s_i, y_i)$ whose minimization or maximization with respect to \underline{s} leads to a maximum likelihood decision on \underline{S} . Specify if your metric needs to be minimized or maximized. Hint: Think of a vector channel $\underline{Y} = \underline{S} + \underline{Z}$, where $\underline{Z} = (Z_1, \dots, Z_6)$ is a sequence of i.i.d. components with $Z_i \sim \mathcal{N}(0, \sigma^2)$.

(d) Assume $\underline{Y} = (Y_1, Y_2, \dots, Y_5, Y_6) = (2, 0, -1, 1, 0, -1)$. Find the maximum likelihood estimate of the information sequence \underline{U} . Please: Do not write into the trellis that you have drawn in Part (b); work on a copy of that trellis.

PROBLEM 6. (Linear Transformations.)

(i)(a) First review the notion of a field. (See e.g. K. Hoffman and R. Kunze, Linear Algebra, Prentice Hall or your favorite linear algebra book.)

Now consider the set $\mathcal{F} = \{0, 1\}$ with the following addition and multiplication tables:

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \quad \begin{array}{c|cc} \times & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

Does \mathcal{F} , “+”, and “ \times ” form a field?

(i)(b) Repeat using $\mathcal{F} = \{\pm 1\}$ and the following addition and multiplication tables:

$$\begin{array}{c|cc} + & 1 & -1 \\ \hline 1 & 1 & -1 \\ -1 & -1 & 1 \end{array} \quad \begin{array}{c|cc} \times & 1 & -1 \\ \hline 1 & 1 & 1 \\ -1 & 1 & -1 \end{array}$$

(ii)(a) Now first review the notion of a vector space.

Let \mathcal{F} , + and \times be as defined in (i)(a). Let $\mathcal{V} = \mathcal{F}^\infty$. (The latter is the set of infinite sequences with components in \mathcal{F} . Does \mathcal{V} , \mathcal{F} , + and \times form a vector space?

(ii)(b) Repeat using \mathcal{F} , + and \times as in (i)(b).

(iii)(a) Review the concept of linear transformation from a vector space \mathcal{I} to a vector space \mathcal{O} . Now let $f : \mathcal{I} \rightarrow \mathcal{O}$ be the mapping implemented by the encoder described in this chapter. Specifically, let $\mathbf{x} = f(\mathbf{d})$ be specified by

$$\begin{aligned} x_{2j-1} &= d_{j-1} \oplus d_{j-2} \oplus d_{j-3} \\ x_{2j} &= d_j \oplus d_{j-2}. \end{aligned}$$

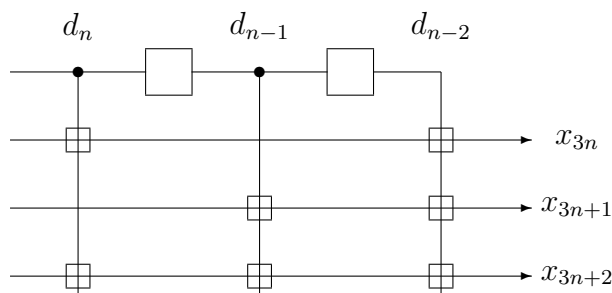


Figure 6.4: Convolutional encoder. $x_{3n} = d_n d_{n-2}$; $x_{3n+1} = d_{n-1} d_{n-2}$;
 $x_{3n+2} = d_n d_{n-1} d_{n-2}$.

Is this encoder linear?

PROBLEM 7. (Rate 1/3 Convolutional Code.)

Consider the following convolutional code, to be used for the transmission of some information sequence $d_i \in \{-1, 1\}$:

- (i) Draw the state diagram for this encoder.
- (ii) Suppose that this code is decoded using the Viterbi algorithm. Draw the detour flowgraph.
- (iii) This encoder/decoder is used on an AWGN channel. The energy available per source digit is E_b and the power spectral density of the noise is $N_0/2$. Give an upper bound on the bit error probability P_b as a function of E_b/N_0 .

PROBLEM 8. (Convolutional Code.)

The following equations define a convolutional code for a data sequence $d_i \in \{-1, 1\}$:

$$x_{3n} = d_{2n} \cdot d_{2n-1} \cdot d_{2n-2} \quad (6.9)$$

$$x_{3n+1} = d_{2n+1} \cdot d_{2n-2} \quad (6.10)$$

$$x_{3n+2} = d_{2n+1} \cdot d_{2n} \cdot d_{2n-2} \quad (6.11)$$

- (i) Draw an implementation of the encoder of this convolutional code, using only delay elements D and multipliers. Hint: Split the data sequence d into two sequences, one containing only the even-indexed samples, the other containing only the odd-indexed samples.

- (ii) What is the rate of this convolutional code?
- (iii) Draw the state diagram for this convolutional encoder.
- (iv) Does the formula for the upper bound on P_b that was derived in class still hold? If not, make the appropriate changes.
- (v) (optional) Now suppose that the code is used on an AWGN channel. The energy available per source digit is E_b and the power spectral density of the noise is $N_0/2$. Give the detour flowgraph, and derive an upper bound on the bit error probability P_b as a function of E_b/N_0 .

PROBLEM 9. (PSD of a Basic Encoder)

Consider the transmitter shown in Figure 6.5, when $\dots D_{-i}, D_i, D_{i+1}, \dots$ is a sequence of independent and uniformly distributed random variables taking value in $\{\pm 1\}$.

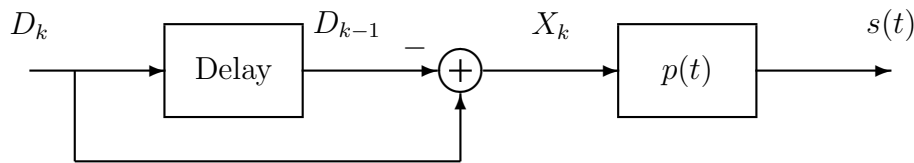


Figure 6.5: Encoder

The transmitted signal is

$$s(t) = \sum_{i=-\infty}^{\infty} X_i p(t - iT - \Theta),$$

where Θ is a random variable, uniformly distributed in $[0, T]$.

$$X_i = D_i - D_{i-1}$$

$$p(t) = 1_{[-\frac{T}{2}, \frac{T}{2}]}(t).$$

- (a) Determine $R_X[k] = E[X_{i+k}X_i]$.
- (b) Determine $R_p(\tau) = \int_{-\infty}^{\infty} p(t + \tau)p(t)dt$.
- (c) Determine the autocorrelation function $R_s(\tau)$ of the signal $s(t)$.
- (d) Determine the power spectral density $S_s(f)$.

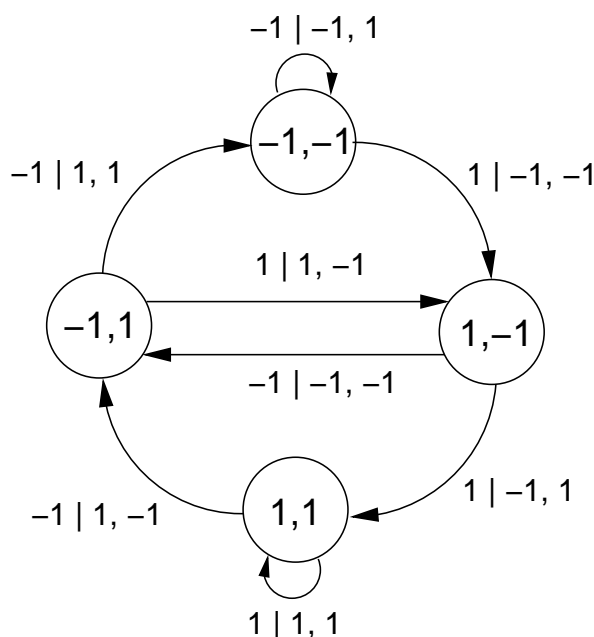
PROBLEM 10. (Convolutional Encoder, Decoder and Error Probability Analysis)

Consider a channel, where a transmitter wants to send a sequence $\{D_j\}$ taking values in $\{-1, +1\}$, for $j = 0, 1, 2, \dots, k-1$. This sequence is encoded using a convolutional

encoder. The channel adds white Gaussian noise to the transmitted signal. If we let X_j denote the transmitted value, then, the received value is: $Y_j = X_j + Z_j$, where $\{Z_j\}$ is a sequence of i.i.d. zero-mean Gaussian random variables with variance $\frac{N_0}{2}$. The receiver has to decide which sequence was transmitted using the optimal decoding rule.

(a) Convolutional Encoder

Consider the convolutional encoder corresponding to the finite state machine drawn below. The transitions are labeled by $D_j|X_{2j}, X_{2j+1}$, and the states by D_{j-1}, D_{j-2} . We assume that the initial content of the memory is $(1, 1)$.



- (i) What is the rate of this encoder?
- (ii) Sketch the filter (composed of shift registers and multipliers) corresponding to this finite state machine. How many shift registers do you need?
- (iii) Draw a section of the trellis representing this encoder.

(b) Viterbi Decoder

Let X_j^i denote the output of the convolutional encoder at time j when we transmit hypothesis i , $i = 0, \dots, m - 1$, where m is the number of different hypotheses.

Assume that the received vector is $\bar{Y} = (Y_1, Y_2, Y_3, Y_4, Y_5, Y_6) = (-1, -3, -2, 0, 2, 3)$. It is the task of the receiver to decide which hypothesis i was chosen or, equivalently, which vector $\bar{X}^i = (X_1^i, X_2^i, X_3^i, X_4^i, X_5^i, X_6^i)$ was transmitted.

- (i) Use the Viterbi algorithm to find the most probable transmitted vector \bar{X}^i .

(c) Performance Analysis

- (i) Suppose that this code is decoded using the Viterbi algorithm. Draw the detour flow graph, and label the edges by the input weight using the symbol I , and the output weight using the symbol D .
- (ii) Considering the following generating function

$$T(I, D) = \frac{ID^4}{1 - 3ID},$$

What is the value of

$$\sum_{i,d} ia(i, d)e^{-\frac{d}{2N_0}},$$

where $a(i, d)$ is the number of detours with i bit errors and d channel errors? First compute this expression, then give an interpretation in terms of probability of error of this quantity.

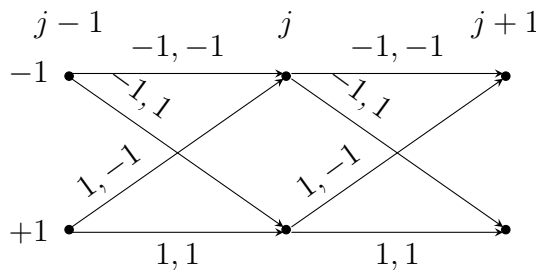
Hints: Recall that the generating function is defined as $T(I, D) = \sum_{i,d} a(i, d)D^d I^i$. You may also use the formula $\sum_{k=1}^{\infty} kq^{k-1} = \frac{1}{(1-q)^2}$ if $|q| < 1$.

PROBLEM 11. (Trellis with Antipodal Signals)

Assume that the sequence X_1, X_2, \dots is sent over an additive white Gaussian noise channel, i.e.,

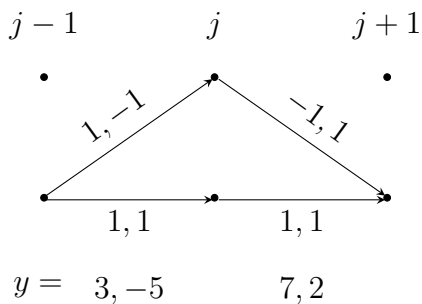
$$Y_i = X_i + Z_i,$$

where the Z_i are i.i.d. zero-mean Gaussian random variables with variance σ^2 . The sequence X_i is the output of a convolutional encoder described by the following trellis.

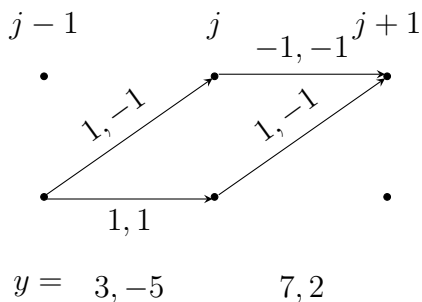


As the figure shows, the trellis has two states labeled with $+1$ and -1 , respectively. The probability assigned to each of the two branches leaving any given state is $1/2$. The trellis is also labeled with the output produced when a branch is traversed and with the trellis depths $j - 1, j, j + 1$.

- (a) Consider the two paths in the following picture. Which of the two paths is more likely if the corresponding channel output subsequence $y_{2j-1}, y_{2j}, y_{2j+1}, y_{2(j+1)}$ is $3, -5, 7, 2$?

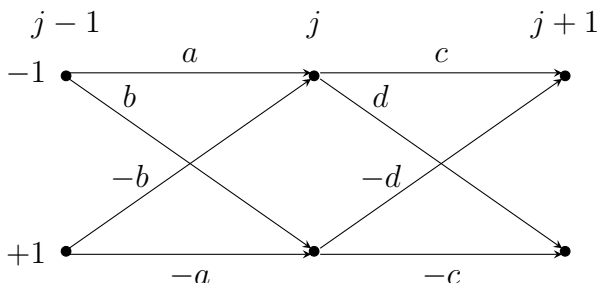


(b) Now, consider the following two paths with the same channel output as in the previous question. Find again the most likely of the two paths.



(c) If you have made no mistake in the previous two questions, the state at depth j of the most likely paths is the same in both cases. This is no coincidence as we will now prove.

The first step is to remark that the metric has to be as in the following picture for some value of $a, b, c,$ and d .



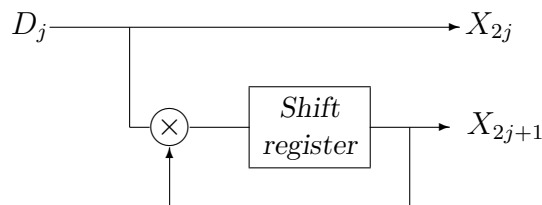
- (i) Now let us denote by $\sigma_k \in \{\pm 1\}$ the state at depth $k, k = 0, 1, \dots,$ of the maximum likelihood path. Assume that a genie tells you that $\sigma_{j-1} = 1$ and $\sigma_{j+1} = 1$. In terms of $a, b, c, d,$ write down a necessary condition for $\sigma_j = 1$. (The condition is also sufficient up to ties.)
- (ii) Now assume that $\sigma_{j-1} = 1$ and $\sigma_{j+1} = -1$. What is the condition for choosing $\sigma_j = 1$?
- (iii) Now assume that $\sigma_{j-1} = -1$ and $\sigma_{j+1} = 1$. What is the condition for $\sigma_j = 1$?
- (iv) Now assume that $\sigma_{j-1} = -1$ and $\sigma_{j+1} = -1$. What is the condition for $\sigma_j = 1$?

- (v) Are the four conditions equivalent? Justify your answer.
- (d) Comment on the advantage, if any, implied by your answer to part (v) of question (c).

PROBLEM 12. (Convolutional Code: Complete Analysis)

(a) Convolutional Encoder

Consider the following convolutional encoder. The input sequence D_j takes values in $\{-1, +1\}$ for $j = 0, 1, 2, \dots, k-1$. The output sequence, call it X_j , $j = 0, \dots, 2k-1$, is the result of passing D_j through the filter shown below, where we assume that the initial content of the memory is 1.



- (i) In the case $k = 3$, how many different hypotheses can the transmitter send using the input sequence (D_0, D_1, D_2) , call this number m .
- (ii) Draw the finite state machine corresponding to this encoder. Label the transitions with the corresponding input and output bits. How many states does this finite state machine have?
- (iii) Draw a section of the trellis representing this encoder.
- (iv) What is the rate of this encoder?
(number of information bits / number of transmitted bits).
- (b) Viterbi Decoder

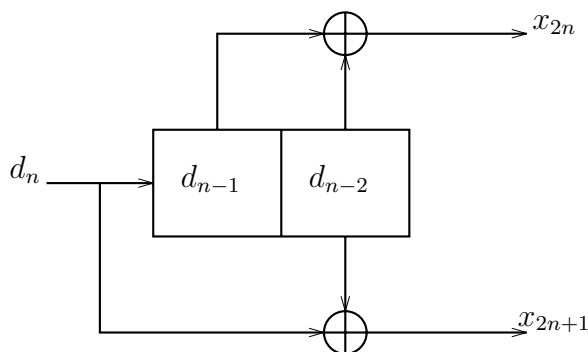
Consider the channel defined by $Y_j = X_j^i + Z_j$. Let X_j^i denote the output of the convolutional encoder at time j when we transmit hypothesis i , $i = 0, \dots, m-1$. Further, assume that Z_j is a zero-mean Gaussian random variable with variance $\sigma^2 = 4$ and let Y_j be the output of the channel.

Assume that the received vector is $\bar{Y} = (Y_1, Y_2, Y_3, Y_4, Y_5, Y_6) = (1, 2, -2, -1, 0, 3)$. It is the task of the receiver to decide which hypothesis i was chosen or, equivalently, which vector $\bar{X}^i = (X_1^i, X_2^i, X_3^i, X_4^i, X_5^i, X_6^i)$ was transmitted.

- (i) Without using the Viterbi algorithm, write formally (in terms of \bar{Y} and \bar{X}^i) the optimal decision rule. Can you simplify this rule to express it as a function of inner products of vectors? In that case, how many inner products do you have to compute to find the optimal decision?
- (ii) Use the Viterbi algorithm to find the most probable transmitted vector \bar{X}^i .
- (c) Performance Analysis.
- (i) Draw the detour flow graph corresponding to this decoder and label the edges by the input weight using the symbol I , the output weight (of both branches) using the symbol D .

PROBLEM 13. (Viterbi for the Binary Erasure Channel)

Consider the following convolutional encoder. The input sequence belongs to the binary alphabet $\{0, 1\}$. (This means we are using XOR over $\{0, 1\}$ instead of multiplication over $\{\pm 1\}$.)



- What is the rate of the encoder?
- Draw one trellis section for the above encoder.
- Consider communication of this sequence through the channel known as Binary Erasure Channel (BEC). The input of the channel belongs to $\{0, 1\}$ and the output belongs to $\{0, 1, ?\}$. The “?” denotes an erasure which means that the output is equally likely to be either 0 or 1. The transition probabilities of the channel are given by

$$P_{Y|X}(0 | 0) = P_{Y|X}(1 | 1) = 1 - \epsilon,$$

$$P_{Y|X}(? | 0) = P_{Y|X}(? | 1) = \epsilon.$$

Starting from first principles derive the branch metric of the optimal (MAP) decoder. (Hint: Start with $p(x | y)$. Hopefully you are not scared of ∞ ?)

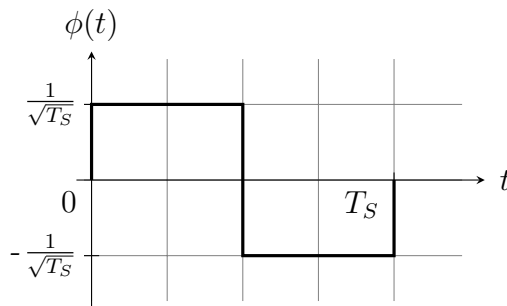
- Assuming that the initial state is $(0, 0)$, what is the most likely input corresponding to $\{0, ?, ?, 1, 0, 1\}$?
- What is the maximum number of erasures the code can correct? (Hint: What is the minimum distance of the code? Just guess from the trellis, don't use the detour graph. :-)

PROBLEM 14. (Power Spectrum: Manchester Pulse)

In this problem you will derive the power spectrum of a signal

$$X(t) = \sum_{i=-\infty}^{\infty} X_i \phi(t - iT_s - \Theta)$$

where $\{X_i\}_{i=-\infty}^{\infty}$ is an iid sequence of uniformly distributed random variables taking values in $\{\pm\sqrt{E_s}\}$, Θ is uniformly distributed in the interval $[0, T_s]$, and $\phi(t)$ is the so-called Manchester pulse shown in the following figure



- (a) Let $r(t) = \sqrt{\frac{1}{T_s}} 1_{[-\frac{T_s}{4}, \frac{T_s}{4}]}(t)$ be a rectangular pulse. Plot $r(t)$ and $r_{\mathcal{F}}(f)$, both appropriately labeled, and write down a mathematical expression for $r_{\mathcal{F}}(f)$.
- (b) Derive an expression for $|\phi_{\mathcal{F}}(f)|^2$. Your expression should be of the form $A \frac{\sin^m(\cdot)}{(\cdot)^n}$ for some A , m , and n . Hint: Write $\phi(t)$ in terms of $r(t)$ and recall that $\sin x = \frac{e^{jx} - e^{-jx}}{2j}$ where $j = \sqrt{-1}$.
- (c) Determine $R_X[k] \triangleq E[X_{i+k}X_i]$ and the power spectrum

$$S_X(f) = \frac{|\phi_{\mathcal{F}}(f)|^2}{T_s} \sum_{k=-\infty}^{\infty} R_X[k] e^{-j2\pi k f T_s}.$$

Chapter 7

Communication Across Bandpass AWGN Channels

7.1 Introduction

In the last part of this course we consider communication across bandpass AWGN channels. The block diagram of a general channel model is shown in Fig. 7.1. It is similar to the ideal lowpass channel model considered in Section 5.2, but the filter's frequency response is now that of an *ideal bandpass filter*, i.e.,

$$h_{\mathcal{F}}(f) = \begin{cases} 1, & ||f| - f_0| \leq \frac{B}{2} \\ 0, & \text{otherwise.} \end{cases}$$

$N(t)$ is white Gaussian noise of power spectral density $\frac{N_0}{2}$.

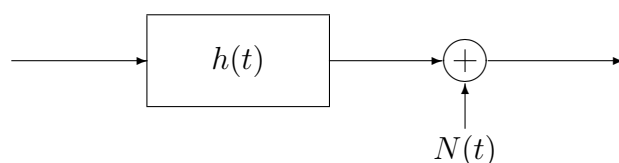


Figure 7.1: Bandpass AWGN Channel

There are various reasons for being interested in knowing how to communicate across a bandpass AWGN channel. Some of them are rooted into physics and some are dictated by practical choices. Among the former we mention that in wireless communications the channel seen between the transmit and the receiver antenna is always a bandpass channel (low frequency components do not generate electromagnetic waves capable of traveling long distances with small attenuation). The usable bandwidth of this channel

is typically quite large however (it depends on the antennas among other things). More severe restrictions are dictated by international agreements that specify which portion of the electromagnetic spectrum can be used for what.

Regardless whether we have decided to use the better portion of the bandpass determined by physical constraints, or because we are complying with international regulations, what we “see” is often a bandpass channel as the one in Fig. 7.1.

If f_0 and B are in a certain relationship (See Problem 3 of Chapter 5), then $\psi(t) = h(t)/\|h\|$ fulfills the Nyquist criterion. In this case we can, in principle, proceed exactly as in the Chapter 5 using this $\psi(t)$ as the basic pulse.

Now we proceed to derive an alternative (more general and widely used) approach that works regardless of the center frequency f_0 and bandwidth B . The idea is to produce a baseband signal via a baseband Nyquist pulse $\psi(t)$ and then shift the spectrum of the obtained signal (that we will denote by $s_E(t)$) to the desired location for $s(t)$. The receiver will do a somewhat inverse operation.

7.2 Baseband Equivalent of a Passband Signal

In this section we learn that for each passband signal $s(t)$ and for each “center frequency” f_0 (typically a frequency in the middle of the positive support set of $s_{\mathcal{F}}(f)$), there is a well-defined baseband signal $s_E(t)$ called the baseband equivalent of $s(t)$. One can go from $s(t)$ to $s_E(t)$ and back by means of a relatively simple operation. You may want to review the Appendix before reading on.

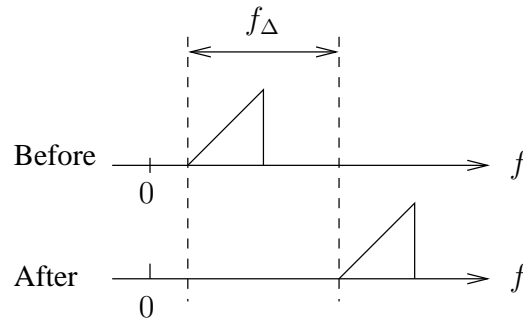
We care about $s_E(t)$ since at the transmitter it is often much easier to “assemble” $s_E(t)$ and then convert it to $s(t)$ rather than aiming directly at $s(t)$. Similarly, at the receiver it is easier to deal with the baseband equivalent of the received signal rather than with the received signal itself.

What makes $s_E(t)$ easier to work with is the fact that it contains relatively low frequencies. For a low frequency signals a wire is like a pipe. The same wire can act as an antenna to high-frequency signals. This means interference among signals that are not supposed to interfere. It also means that the signal at the output of an amplifier may feedback to the input and turn the amplifier into an oscillator (much as when you put a microphone near the speaker).

The other reason why we prefer to work with baseband signals as much as possible is that an amplifier is good only in a given frequency range. Now $s_E(t)$ has a fixed frequency range, whereas we want the center frequency of $s(t)$ to be selectable within some wider range. This means that an amplifier for $s(t)$ is a more sophisticated device than one for $s_E(t)$. To be more concrete, $s_E(t)$ could have a bandwidth of a few MHz, whereas we may want to be able to vary the center frequency of $s(t)$ over hundreds of MHz.

Translating the center frequency of an analytic signal $\hat{s}(t)$ can be easily done using the

frequency shift property of the Fourier transform: we just multiply it by $e^{j2\pi f_{\Delta}t}$ where f_{Δ} is the desired frequency translate. (See figure below)



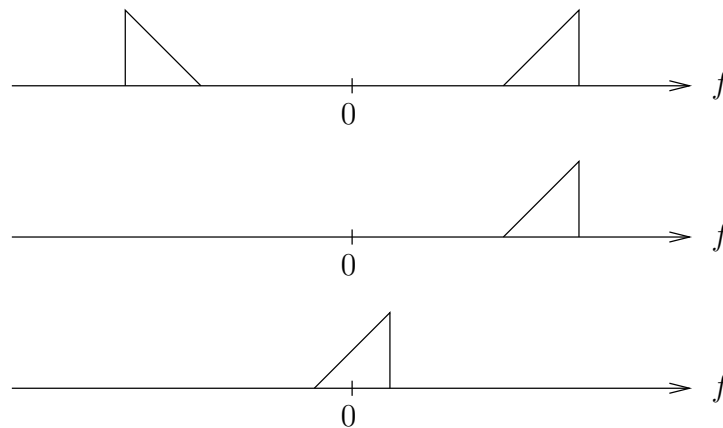
Let $s(t)$ be a real-valued bandpass signal. By this we mean that $s_{\mathcal{F}}(|f| - f_0) = 0$ for $|f| - f_0 \geq B$ for some B . The Fourier transform (amplitude only) of such a signal is shown on the top figure below. The figure also depicts the analytic equivalent signal $\hat{s}(t)$ (middle figure) and the signal

$$s_E(t) = \hat{s}(t)e^{-j2\pi f_0 t} \quad (\text{Complex-Valued Baseband-Equivalent of } s(t))$$

whose Fourier transform is

$$s_{E,\mathcal{F}}(f) = \hat{s}_{\mathcal{F}}(f + f_0)$$

(bottom figure). (The figure does not show the scaling by $\sqrt{2}$ in going from $s(t)$ to $\hat{s}(t)$.)



Going the other way is straightforward. As shown in the Appendix, from $s_E(t)$ one immediately obtains $s(t)$ according to

$$s(t) = \sqrt{2}\Re\{s_E(t)e^{j2\pi f_0 t}\}.$$

7.3 Up/Down Conversion

As the name indicates, $s_E(t)$ is a baseband signal. We know how to generate any such signal via the sampling theorem or via Nyquist pulses as explained in Chapter 5.¹

Hence, to generate a passband signal $s(t)$ we may start with its baseband-equivalent signal $s_E(t)$ and up-convert it as shown in Fig. 7.2.

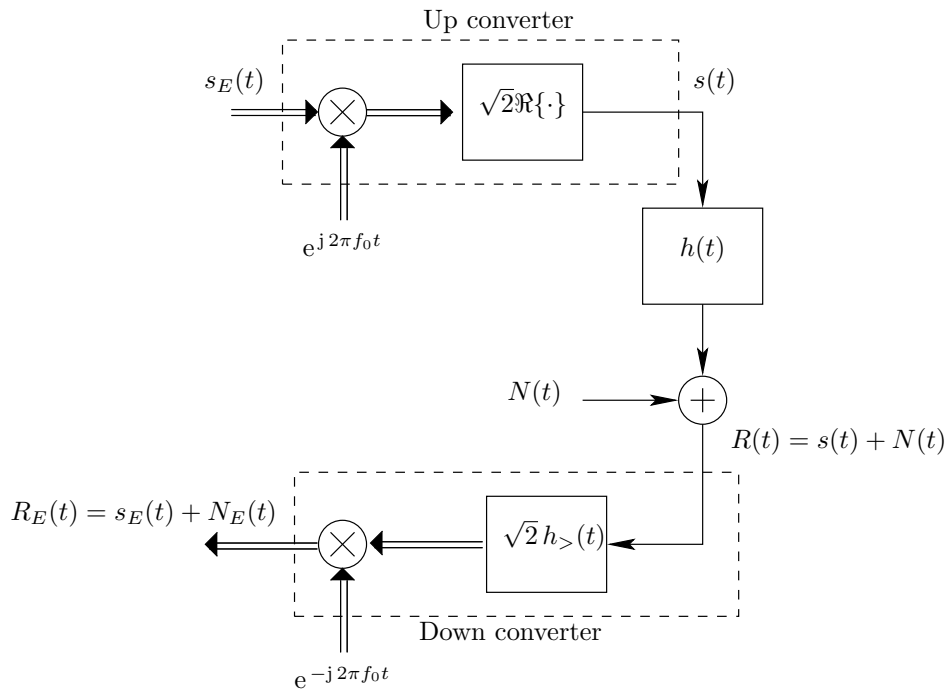


Figure 7.2: Bandpass communication system. Double lines denote complex signals.

Let the channel output signal be $R(t) = s(t) + N(t)$. Its baseband equivalent signal is $R_E(t) = s_E(t) + N_E(t)$ and it is obtained as shown in Figure 7.2. $R_E(t)$ is a sufficient statistic (since we can go back to $R(t)$ if we wish). We will see in Chapter 8 that over the frequency range of interest $N_E(t)$ is a complex-valued white Gaussian noise process. Hence the signal $R_E(t)$ is the output of a baseband AWGN channel.

To summarize, by means of the up/down-converter we have reduced the bandpass AWGN channel to a baseband AWGN channel. Input and output signals to the baseband AWGN channel are complex-valued but this is not a problem since all the tools we have used so far (e.g. the \mathcal{L}_2 space) are valid for complex-valued signals.

Besides being conceptually nice that we can reduce the bandpass channel to a lowpass channel as shown in Fig. 7.2, the approach described in Fig. 7.2 is not just academic. It is indeed common practice to carry out as much as possible of the needed signal processing

¹The sampling theorem holds unchanged for complex-valued signals.

at low frequencies (in baseband) where one does not have to worry about a piece of wire acting variously as an antenna, as a capacitor, or as an inductor.

On the other hand, a common practice of communication system engineers/designers is to look at the up-converter and at the downconverter as part of the channel and pretend that he/she has to design a system for baseband communication. We now develop this view.

7.4 Baseband-Equivalent Channel Impulse Response

Assume that the bandpass channel has an arbitrary real-valued impulse response $h(t)$, i.e., $h_{\mathcal{F}}(f)$ is arbitrary in the range $||f| - f_0| \leq B$ and vanishes outside this range. The fact that $h_{\mathcal{F}}(f)$ vanishes outside the passband range is only important to the extent that it reminds us that we can't use the channel outside this range. We also assume that the noise power spectrum is arbitrary but in this section we neglect the noise and focus on the channel impulse response.

Neglecting the noise for the moment, we now show that with the up-converter at the channel input and the down-converter at the channel output we convert the original bandpass channel into an equivalent baseband channel. We will deal with the noise in Chapter 8. Notice that we are allowed to study the signal and the noise separately since the system is linear.

Without noise the input/output relationship is

$$r(t) = (h \star s)(t).$$

Taking the Fourier transform on both sides we get the first of the equations below. The other follow from straightforward manipulations of the first one.

$$\begin{aligned} r_{\mathcal{F}}(f) &= h_{\mathcal{F}}(f)s_{\mathcal{F}}(f) \\ r_{\mathcal{F}}(f)h_{>,\mathcal{F}}(f)\sqrt{2} &= h_{\mathcal{F}}(f)h_{>,\mathcal{F}}(f)s_{\mathcal{F}}(f)h_{>,\mathcal{F}}(f)\sqrt{2} \\ \hat{r}_{\mathcal{F}}(f) &= \frac{\hat{h}_{\mathcal{F}}(f)}{\sqrt{2}}\hat{s}_{\mathcal{F}}(f) \\ r_{E,\mathcal{F}}(f) &= \frac{h_{E,\mathcal{F}}(f)}{\sqrt{2}}s_{E,\mathcal{F}}(f). \end{aligned} \tag{7.1}$$

Hence, when we send a signal $s(t)$ through a channel of impulse response $h(t)$ it is as sending the baseband equivalent signal $s_E(t)$ through a channel of baseband equivalent impulse response $\frac{h_E(t)}{\sqrt{2}}$.

The baseband equivalent channel is shown in Fig. 7.3.

The properties of the noise $N_Z(t)$ will be derived in Chapter 9. We may use the baseband equivalent channel for instance in a simulation in which alternatives for $s_E(t)$ are being tested.

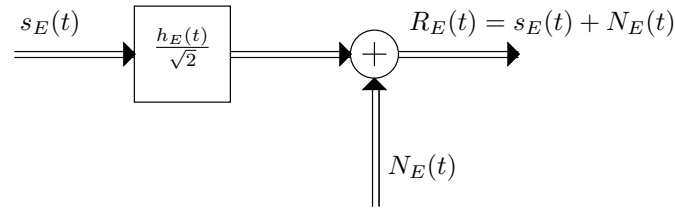


Figure 7.3: Baseband-equivalent communication system.

7.5 Problems

PROBLEM 1. (Baseband Equivalent Relationship)

In this problem we neglect noise and consider the situation in which we transmit a signal $X(t)$ and receive

$$R(t) = \sum_i \alpha_i X(t - \tau_i).$$

Show that the baseband equivalent relationship is

$$R_E(t) = \sum_i \beta_i X_E(t - \tau_i).$$

Express β_i explicitly.

PROBLEM 2. (Fun with Sine and Cosine)

A bandpass signal $x(t)$ may be written as $x(t) = \sqrt{2}\Re\{x_E(t)e^{j2\pi f_0 t}\}$, where $x_E(t)$ is the baseband equivalent of $x(t)$.

- Show that a signal $x(t)$ can also be written as $x(t) = a(t) \cos[2\pi f_0 t + \theta(t)]$ and describe $a(t)$ and $\theta(t)$ in terms of $x_E(t)$. Interpret this result.
- Show that the signal $x(t)$ can also be written as $x(t) = x_{EI}(t) \cos 2\pi f_0 t - x_{EQ}(t) \sin(2\pi f_0 t)$, and describe $x_{EI}(t)$ and $x_{EQ}(t)$ in terms of $x_E(t)$. (This shows how you can obtain $x(t)$ without doing complex-valued operations.)
- Find the baseband equivalent of the signal $x(t) = A(t) \cos(2\pi f_0 t + \varphi)$, where $A(t)$ is a real-valued lowpass signal. Hint: You may find it easier to guess an answer and verify that it is correct.

PROBLEM 3. (Equivalent Baseband Signal)

- You are given a “passband” signal $\psi(t)$ whose spectrum is centered around f_0 . Write down in a generic form the different steps needed to find the baseband equivalent signal.

(b) Consider the waveform

$$\psi(t) = \text{sinc}\left(\frac{t}{T}\right) \cos(2\pi f_0 t).$$

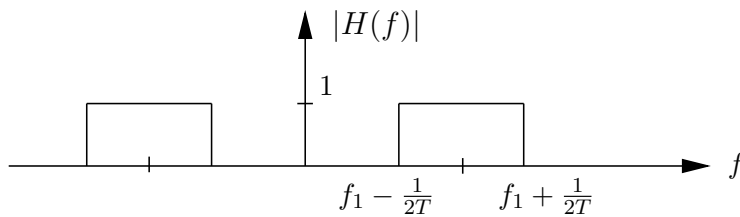
What is the equivalent baseband signal of this waveform.

(c) Assume that the signal $\psi(t)$ is passed through the filter with impulse response $h(t)$ where $h(t)$ is specified by its baseband equivalent impulse response $h_E(t) = \frac{1}{T\sqrt{2}} \text{sinc}^2\left(\frac{t}{2T}\right)$. What is the output signal, both in passband as well as in baseband?

Hint: The Fourier transform of $\cos(2\pi f_0 t)$ is $\frac{1}{2}\delta(f - f_0) + \frac{1}{2}\delta(f + f_0)$. The Fourier transform of $\frac{1}{T}\text{sinc}\left(\frac{t}{T}\right)$ is equal to $\mathbf{1}_{[-\frac{1}{2T}, \frac{1}{2T}]}(f)$ with $\mathbf{1}_{[-\frac{1}{2T}, \frac{1}{2T}]}(f) = 1$ if $f \in [-\frac{1}{2T}, \frac{1}{2T}]$ and 0 otherwise.

PROBLEM 4. (Up-Down Conversion)

We want to send a “passband” signal $\psi(t)$ whose spectrum is centered around f_0 , through a waveform channel defined by its impulse response $h(t)$. The Fourier transform $H(f)$ of the impulse response is given by



where $f_1 \neq f_0$.

(a) Write down in a generic form the different steps needed to send $\psi(t)$ at the correct frequency f_1 .

(b) Consider the waveform

$$\psi(t) = \text{sinc}\left(\frac{t}{T}\right) \cos(2\pi f_0 t).$$

What is the output signal, in passband (at center frequency f_1) as well as in baseband?

(c) Assume that $f_0 = f_1 + \epsilon$, with $\epsilon \ll \frac{1}{2T}$, and that the signal $\psi(t)$ is directly transmitted without any frequency shift. What will be the central frequency of the output signal?

Hint: The Fourier transform of $\cos(2\pi f_0 t)$ is $\frac{1}{2}\delta(f - f_0) + \frac{1}{2}\delta(f + f_0)$. The Fourier transform of $\frac{1}{T}\text{sinc}\left(\frac{t}{T}\right)$ is equal to $\mathbf{1}_{[-\frac{1}{2T}, \frac{1}{2T}]}(f)$ with $\mathbf{1}_{[-\frac{1}{2T}, \frac{1}{2T}]}(f) = 1$ if $f \in [-\frac{1}{2T}, \frac{1}{2T}]$ and 0 otherwise.

PROBLEM 5. (Smoothness of Bandlimited Signals)

In communications one often finds the statement that if $s(t)$ is a signal of bandwidth W , then it can't vary too much in a small interval $\tau \ll 1/W$. Based on this, people sometimes substitute $s(t)$ for $s(t + \tau)$. In this problem we will derive an upper bound for $|s(t + \tau) - s(t)|$. It is assumed that $s(t)$ is a finite energy signal with Fourier transform satisfying $S(f) = 0, |f| > W$.

- (a) Let $H(f)$ be the frequency response of the ideal lowpass-filter defined as 1 for $|f| \leq W$ and 0 otherwise. Show that

$$s(t + \tau) - s(t) = \int s(\xi)[h(t + \tau - \xi) - h(t - \xi)]d\xi. \quad (7.2)$$

- (b) Use Schwarz inequality to prove that

$$|s(t + \tau) - s(t)|^2 \leq 2E_s[E_h - R_h(\tau)], \quad (7.3)$$

where E_s is the energy of $s(t)$,

$$R_h(\tau) = \int h(\xi + \tau)h(\xi)d\xi$$

is the (time) autocorrelation function of $h(t)$, and $E_h = R_h(0)$.

- (c) Show that $R_h(\tau) = h * h(\tau)$, i.e., for h the convolution with itself equals its autocorrelation function. What makes h have this property?
- (d) Show that $R_h(\tau) = h(\tau)$.
- (e) Put things together to derive the upperbound

$$|s(t + \tau) - s(t)| \leq \sqrt{2E_s[E_h - h(\tau)]} = \sqrt{4WE_s \left(1 - \frac{\sin(2\pi W\tau)}{2\pi W\tau}\right)}. \quad (7.4)$$

[Can you determine the impulse response $h(t)$ without looking it up and without solving integrals? Remember the "mnemonics" given in class?] Verify that for $\tau = 0$ the bound is tight.

- (f) Let E_D be the energy in the difference signal $s(t + \tau) - s(t)$. Assume that the duration of $s(t)$ is T and determine an upperbound on E_D .
- (g) Consider a signal $s(t)$ with parameters $2W = 5$ Mhz and $T = 5/2W$. Find a numerical value T_m for the time difference τ so that $E_D(\tau) \leq 10^{-2}E_s$ for $|\tau| \leq T_m$.

Appendix 7.A Some Review from Fourier Analysis

Recall that if $s(t)$ is a real-valued signal, then its Fourier transform $s_{\mathcal{F}}(f)$ satisfies the symmetry property

$$s_{\mathcal{F}}(f) = s_{\mathcal{F}}^*(-f) \quad (\text{Symmetry Property})$$

where $s_{\mathcal{F}}^*$ denotes the complex conjugate of $s_{\mathcal{F}}$. If $s(t)$ is a *purely imaginary* signal, then its Fourier transform satisfies the anti-symmetry property

$$s_{\mathcal{F}}(f) = -s_{\mathcal{F}}^*(-f) \quad (\text{Anti-Symmetry Property}).$$

The symmetry and the anti-symmetry properties can be easily verified directly from the definition of the Fourier transform. (Would you be able to know how?).

We will also often use the frequency shift property of the Fourier transform, namely

$$s(t)e^{j2\pi f_0 t} \longleftrightarrow s_{\mathcal{F}}(f - f_0) \quad (\text{Frequency Shift Property}).$$

The frequency shift property is also an immediate consequence of the definition of the Fourier transform.

The symmetry property implies that the Fourier transform $s_{\mathcal{F}}(f)$ of a real-valued signal $s(t)$ has redundant information. If we know $s_{\mathcal{F}}(f)$ for $f \geq 0$ then we can infer $s_{\mathcal{F}}(f)$ for $f \leq 0$ and thus we know $s_{\mathcal{F}}(f)$ and $s(t)$ for all frequencies and all times, respectively. This also implies that a real-valued signal $s(t)$ is in one-to-one correspondence with a complex-valued signal $\hat{s}(t)$ that occupies 1/2 the bandwidth of $s(t)$, namely the signal obtained by removing the negative frequency component from $s(t)$. If we shift $\hat{s}_{\mathcal{F}}(f)$ by some proper amount f_0 we obtain a signal $s_{E\mathcal{F}}(f)$ with support centered at $f = 0$. Its inverse Fourier transform $s_E(t)$ is a baseband complex-valued signal. (It is complex-valued in general since it does not necessarily fulfill the symmetry property.) This shows that there is a one-to-one correspondence between the set of complex-valued (baseband) signals of frequency-domain support $[-\frac{B}{2}, \frac{B}{2}]$ and the set of real-valued (passband) signals of frequency-domain support $[-\frac{B}{2} - f_0, -f_0 + \frac{B}{2}] \cup [-\frac{B}{2} + f_0, f_0 + \frac{B}{2}]$. We now develop this relationship.

Define the filter with impulse response $h_{>}(t)$ via its Fourier transform $h_{>,\mathcal{F}}(f)$, namely

$$h_{>,\mathcal{F}}(f) = \begin{cases} 1 & \text{for } f > 0 \\ 1/2 & \text{for } f = 0 \\ 0 & \text{for } f < 0. \end{cases} \quad (7.5)$$

This is a filter that removes the negative portion of the spectrum.

If $s(t)$ is an arbitrary real-valued signal, define $\hat{s}(t)$ to be the signal with Fourier transform

$$\hat{s}_{\mathcal{F}}(f) = \sqrt{2}s_{\mathcal{F}}(f)h_{>,\mathcal{F}}(f).$$

In going from $s(t)$ to $\hat{s}(t)$ we have removed the negative part of the spectrum. The factor $\sqrt{2}$ is introduced so that $s(t)$ and $\hat{s}(t)$ have the same \mathcal{L}_2 norm. A signal $\hat{s}(t)$ such that its Fourier transform vanishes at negative frequencies is said to be *analytic* and $\hat{s}(t)$ is called the *analytic equivalent* of $s(t)$.

Since $s(t)$ is real-valued, the positive portion of the spectrum contains all the information about the signal. Hence we should be able to go back from $\hat{s}(t)$ to $s(t)$. Indeed

$$s(t) = \sqrt{2}\Re\{\hat{s}(t)\}. \quad (7.6)$$

One way to see this is to use the relationship

$$h_{>,\mathcal{F}}(f) = \frac{1}{2} + \frac{1}{2}\text{sign}(f)$$

to obtain

$$\begin{aligned} \hat{s}_{\mathcal{F}}(f) &= \sqrt{2}s_{\mathcal{F}}(f)h_{>,\mathcal{F}}(f) \\ &= \sqrt{2}s_{\mathcal{F}}(f)\left[\frac{1}{2} + \frac{1}{2}\text{sign}(f)\right] \\ &= \frac{s_{\mathcal{F}}(f)}{\sqrt{2}} + \frac{s_{\mathcal{F}}(f)}{\sqrt{2}}\text{sign}(f). \end{aligned}$$

The first term of last line satisfies the symmetry property (by assumption) and therefore the second term satisfies the anti-symmetry property. Hence, taking the inverse Fourier transform, $\hat{s}(t)$ equals $\frac{s(t)}{\sqrt{2}}$ plus an imaginary term, implying (7.6).

Another way to prove the same thing is to write

$$\sqrt{2}\Re\{\hat{s}(t)\} = \frac{1}{\sqrt{2}}(\hat{s}(t) + \hat{s}^*(t))$$

and take the Fourier transform on the right side. The result is

$$s_{\mathcal{F}}(f)h_{>,\mathcal{F}}(f) + s_{\mathcal{F}}^*(-f)h_{>,\mathcal{F}}^*(-f).$$

For positive frequencies the first term equals $s_{\mathcal{F}}(f)$ and the second term vanishes. Hence $\sqrt{2}\Re\{\hat{s}(t)\}$ and $s(t)$ agree for positive frequencies. But two real-valued signals whose Fourier transforms agree for positive frequencies must agree everywhere.

We summarize. Let $s(t)$ be real-valued. If we remove its negative frequencies and multiply by $\sqrt{2}$ to compensate for the energy loss, we obtain its analytic signal $\hat{s}(t)$. The real part of this signal is $\frac{s(t)}{\sqrt{2}}$. From $\hat{s}(t)$ to $s_E(t)$ is just a matter of using the frequency shift property.

Chapter 8

Complex-Valued Random Variables and Processes

8.1 Introduction

In this chapter we define and study complex-valued random variables and complex-valued stochastic processes. We need them to model the noise of the baseband equivalent channel. Besides being practical in many situations, working with complex-valued random variables and processes turns out to be more elegant than working with the real-valued counterparts. We will focus on complex-valued random variables and processes called *proper* (to be defined) since they are what we need and since they are easier to deal with than with non-proper counterparts.

In Sections 8.2–8.6 we define proper random variables and processes, study some of their properties, and derive the probability density function of proper Gaussian random vectors. In subsequent chapters we will learn how to use eigenvectors to simplify our channel model.

To make the chapter self-contained, Appendix 8.A contains a review of relevant facts that will be useful throughout. You should review them now.

8.2 Complex-Valued Random Variables

A complex-valued random variable U (hereafter simply called complex random variable) is defined as a random variable of the form

$$U = U_R + jU_I, \quad j = \sqrt{-1},$$

where U_R and U_I are real-valued random variables.

The statistical properties of $U = U_R + jU_I$ are determined by the joint distribution $P_{U_R U_I}(u_R, u_I)$ of U_R and U_I .

A real random variable X is specified by cumulative distribution function $F_X(x) = \Pr(X \leq x)$. For a complex random variable Z , since there is no natural ordering in the complex plane, the event $Z \leq z$ does not make sense. Instead, we specify a complex random variable by giving the joint distribution of its real and imaginary parts $F_{\Re\{Z\}, \Im\{Z\}}(x, y) = \Pr(\Re\{Z\} \leq x, \Im\{Z\} \leq y)$. Since the pair of real numbers (x, y) can be identified with a complex number $z = x + iy$, we will write the joint distribution $F_{\Re\{Z\}, \Im\{Z\}}(x, y)$ as $F_Z(z)$. Just as we do for real valued random variables, if the function $F_{\Re\{Z\}, \Im\{Z\}}(x, y)$ is differentiable in x and y , we will call the function

$$p_{\Re\{Z\}, \Im\{Z\}}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{\Re\{Z\}, \Im\{Z\}}(x, y)$$

the joint density of $(\Re\{Z\}, \Im\{Z\})$, and again associating with (x, y) the complex number $z = x + iy$, we will call the function

$$p_Z(z) = p_{\Re\{Z\}, \Im\{Z\}}(\Re\{z\}, \Im\{z\})$$

the density of the random variable Z .

A complex random vector $\mathbf{Z} = (Z_1, \dots, Z_n)$ is specified by the joint distribution of $(\Re\{Z_1\}, \dots, \Re\{Z_n\}, \Im\{Z_1\}, \dots, \Im\{Z_n\})$, and we define the distribution of Z as

$$F_{\mathbf{Z}}(\mathbf{z}) = \Pr(\Re\{Z_1\} \leq \Re\{z_1\}, \dots, \Re\{Z_n\} \leq \Re\{z_n\}, \Im\{Z_1\} \leq \Im\{z_1\}, \dots, \Im\{Z_n\} \leq \Im\{z_n\}),$$

and if this function is differentiable in $\Re\{z_1\}, \dots, \Re\{z_n\}, \Im\{z_1\}, \dots, \Im\{z_n\}$, then we define the density of \mathbf{Z} as

$$p_{\mathbf{Z}}(x_1 + iy_1, \dots, x_n + iy_n) = \frac{\partial^{2n}}{\partial x_1 \cdots \partial x_n \partial y_1 \cdots \partial y_n} F_{\mathbf{Z}}(x_1 + iy_1, \dots, x_n + iy_n).$$

The expectation of a real random vector \mathbf{x} is naturally generalized to the complex case

$$E[\mathbf{U}] = E[\mathbf{U}_R] + jE[\mathbf{U}_I].$$

Recall that the covariance matrix of two real-valued random vectors \mathbf{x} and \mathbf{y} is defined as

$$K_{\mathbf{x}\mathbf{y}} = \text{cov}[\mathbf{x}, \mathbf{y}] \triangleq E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{y} - E[\mathbf{y}])^T]. \quad (8.1)$$

To specify the ‘‘covariance’’ of the two complex random vectors $\mathbf{U} = \mathbf{U}_R + j\mathbf{U}_I$ and $\mathbf{V} = \mathbf{V}_R + j\mathbf{V}_I$ the four covariance matrices

$$\begin{aligned} K_{\mathbf{U}_R\mathbf{V}_R} &= \text{cov}[\mathbf{U}_R, \mathbf{V}_R] & K_{\mathbf{U}_R\mathbf{V}_I} &= \text{cov}[\mathbf{U}_R, \mathbf{V}_I] \\ K_{\mathbf{U}_I\mathbf{V}_R} &= \text{cov}[\mathbf{U}_I, \mathbf{V}_R] & K_{\mathbf{U}_I\mathbf{V}_I} &= \text{cov}[\mathbf{U}_I, \mathbf{V}_I] \end{aligned} \quad (8.2)$$

are needed. These four real-valued matrices are equivalent to the following two complex-valued matrices, each of which is a natural generalization of (8.1)

$$\begin{aligned} K_{\mathbf{U}\mathbf{V}} &\triangleq E[(\mathbf{U} - E[\mathbf{U}])(\mathbf{V} - E[\mathbf{V}])^\dagger] \\ J_{\mathbf{U}\mathbf{V}} &\triangleq E[(\mathbf{U} - E[\mathbf{U}])(\mathbf{V} - E[\mathbf{V}])^T] \end{aligned} \quad (8.3)$$

You are encouraged to verify that the following (straightforward) relationships hold:

$$\begin{aligned} K_{\mathbf{UV}} &= K_{\mathbf{U}_R\mathbf{V}_R} + K_{\mathbf{U}_I\mathbf{V}_I} + j(K_{\mathbf{U}_I\mathbf{V}_R} - K_{\mathbf{U}_R\mathbf{V}_I}) \\ J_{\mathbf{UV}} &= K_{\mathbf{U}_R\mathbf{V}_R} - K_{\mathbf{U}_I\mathbf{V}_I} + j(K_{\mathbf{U}_I\mathbf{V}_R} + K_{\mathbf{U}_R\mathbf{V}_I}). \end{aligned} \quad (8.4)$$

This system may be solved for $K_{\mathbf{U}_R\mathbf{V}_R}$, $K_{\mathbf{U}_I\mathbf{V}_I}$, $K_{\mathbf{U}_I\mathbf{V}_R}$, and $K_{\mathbf{U}_R\mathbf{V}_I}$ to obtain

$$\begin{aligned} K_{\mathbf{U}_R\mathbf{V}_R} &= \frac{1}{2}\Re\{K_{\mathbf{UV}} + J_{\mathbf{UV}}\} \\ K_{\mathbf{U}_I\mathbf{V}_I} &= \frac{1}{2}\Re\{K_{\mathbf{UV}} - J_{\mathbf{UV}}\} \\ K_{\mathbf{U}_I\mathbf{V}_R} &= \frac{1}{2}\Im\{K_{\mathbf{UV}} + J_{\mathbf{UV}}\} \\ K_{\mathbf{U}_R\mathbf{V}_I} &= \frac{1}{2}\Im\{-K_{\mathbf{UV}} + J_{\mathbf{UV}}\} \end{aligned} \quad (8.5)$$

proving that indeed the four real-valued covariance matrices in (8.2) are in one-to-one relationship with the two complex-valued covariance matrices in (8.3).

In the literature $K_{\mathbf{UV}}$ is widely used and it is called *covariance matrix* (of the complex random vectors \mathbf{U} and \mathbf{V}). Hereafter $J_{\mathbf{UV}}$ will be called the *pseudo-covariance matrix* (of \mathbf{U} and \mathbf{V}). For notational convenience we will write $K_{\mathbf{U}}$ instead of $K_{\mathbf{UU}}$ and $J_{\mathbf{U}}$ instead of $J_{\mathbf{UU}}$.

DEFINITION 49. \mathbf{U} and \mathbf{V} are said to be uncorrelated if all four covariances in (8.2) vanish.

From (8.3), we now obtain the following.

LEMMA 50. The complex random vectors \mathbf{U} and \mathbf{V} are uncorrelated iff $K_{\mathbf{UV}} = J_{\mathbf{UV}} = 0$.

Proof. The “if” part follows from (8.5) and the “only if” part from (8.4). \square

8.3 Complex-Valued Random Processes

We focus on discrete-time random processes since corresponding results for continuous-time random processes follow in a straightforward fashion.

A discrete-time *complex* random process is defined as a random process of the form

$$U[n] = U_R[n] + jU_I[n]$$

where $U_R[n]$ and $U_I[n]$ are a pair of *real* discrete-time random processes.

DEFINITION 51. A complex random process is wide-sense stationary (w.s.s.) if its real and imaginary parts are jointly w.s.s.

DEFINITION 52. We define

$$\begin{aligned} r_{\mathbf{U}}[m, n] &\triangleq E[U[n+m]U^*[n]] \\ s_{\mathbf{U}}[m, n] &\triangleq E[U[n+m]U[n]] \end{aligned}$$

as the autocorrelation and pseudo-autocorrelation functions of $U[n]$.

LEMMA 53. A complex random process $U[n]$ is w.s.s. if and only if $E[U[n]]$, $r_{\mathbf{U}}[m, n]$, and $s_{\mathbf{U}}[m, n]$ are independent of n .

Proof. The proof is left as an exercise. \square

8.4 Proper Complex Random Variables

Proper random variables are of interest to us since they arise in practical applications and since they are mathematically easier to deal with than their non-proper counterparts.¹

DEFINITION 54. A complex random vector \mathbf{U} is called proper if its pseudo-covariance $J_{\mathbf{U}}$ vanishes. The complex random vectors \mathbf{U}_1 and \mathbf{U}_2 are called jointly proper if the composite random vector $\begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix}$ is proper.

LEMMA 55. Two jointly proper, complex random vectors \mathbf{U} and \mathbf{V} are uncorrelated, if and only if their covariance matrix $K_{\mathbf{UV}}$ vanishes.

Proof. The proof easily follows from the definition of joint properness and Lemma 50. \square

Note that any subvector of a proper random vector is also proper. By this we mean that if $\begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ is proper, then U_1 and U_2 are proper. However, two individual proper random vectors are not necessarily jointly proper.

Using the fact that (by definition) $K_{\mathbf{U}_R\mathbf{U}_I} = K_{\mathbf{U}_I\mathbf{U}_R}^T$, the pseudo-covariance matrix $J_{\mathbf{U}}$ may be written as

$$J_{\mathbf{U}} = (K_{\mathbf{U}_R} - K_{\mathbf{U}_I}) + j(K_{\mathbf{U}_I\mathbf{U}_R} + K_{\mathbf{U}_I\mathbf{U}_R}^T).$$

Thus:

LEMMA 56. A complex random vector \mathbf{U} is proper iff

$$K_{\mathbf{U}_R} = K_{\mathbf{U}_I} \quad \text{and} \quad K_{\mathbf{U}_I\mathbf{U}_R} = -K_{\mathbf{U}_I\mathbf{U}_R}^T,$$

i.e. $J_{\mathbf{U}}$ vanishes, iff \mathbf{U}_R and \mathbf{U}_I have identical auto-covariance matrices and their cross-covariance matrix is skew-symmetric.²

¹Proper Gaussian random vectors also maximize entropy among all random vectors of a given covariance matrix. Among the many nice properties of Gaussian random vectors, this is arguably the most important one in information theory.

²A matrix A is skew-symmetric if $A^T = -A$.

Notice that the skew-symmetry of $K_{U_I U_R}$ implies that $K_{U_I U_R}$ has a zero main diagonal, which means that the real and imaginary part of each component U_k of \mathbf{U} are uncorrelated. The vanishing of $J_{\mathbf{U}}$ does not, however, imply that the real part of U_k and the imaginary part of U_l are uncorrelated for $k \neq l$.

Notice that a *real* random vector is a proper complex random vector, if and only if it is constant (with probability 1), since $K_{U_I} = 0$ and Lemma 56 imply $K_{U_R} = 0$.

LEMMA 57 (CLOSURE UNDER AFFINE TRANSFORMATIONS). *Let \mathbf{U} be a proper n -dimensional random vector, i.e., $J_{\mathbf{U}} = 0$. Then any vector obtained from \mathbf{U} by an affine transformation, i.e. any vector \mathbf{V} of the form $\mathbf{V} = A\mathbf{U} + \mathbf{b}$, where $A \in \mathbb{C}^{m \times n}$ and $\mathbf{b} \in \mathbb{C}^m$ are constant, is also proper.*

Proof. From

$$E[\mathbf{V}] = AE[\mathbf{U}] + \mathbf{b}$$

it follows

$$\mathbf{V} - E[\mathbf{V}] = A(\mathbf{U} - E[\mathbf{U}])$$

Hence we have

$$\begin{aligned} J_{\mathbf{V}} &= E[(\mathbf{V} - E[\mathbf{V}])(\mathbf{V} - E[\mathbf{V}])^T] \\ &= E\{A(\mathbf{U} - E[\mathbf{U}])(\mathbf{U} - E[\mathbf{U}])^T A^T\} \\ &= AJ_{\mathbf{U}}A^T = 0 \end{aligned}$$

□

COROLLARY 58. *Let \mathbf{U} and \mathbf{V} be as in the previous Lemma. Then \mathbf{U} and \mathbf{V} are jointly proper.*

Proof. The vector having \mathbf{U} and \mathbf{V} as subvectors is obtained by the affine transformation

$$\begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} = \begin{bmatrix} I_n \\ A \end{bmatrix} \mathbf{U} + \begin{bmatrix} \mathbf{0} \\ \mathbf{b} \end{bmatrix}.$$

The claim now follows from Lemma 57. □

LEMMA 59. *Let \mathbf{U} and \mathbf{V} be two independent complex random vectors and let \mathbf{U} be proper. Then the linear combination $\mathbf{W} = a_1\mathbf{U} + a_2\mathbf{V}$, $a_1, a_2 \in \mathbb{C}$, $a_2 \neq 0$, is proper iff \mathbf{V} is also proper.*

Proof. The independence of \mathbf{U} and \mathbf{V} and the properness of \mathbf{U} imply

$$J_{\mathbf{W}} = a_1^2 J_{\mathbf{U}} + a_2^2 J_{\mathbf{V}} = a_2^2 J_{\mathbf{V}}.$$

Thus $J_{\mathbf{W}}$ vanishes iff $J_{\mathbf{V}}$ vanishes. □

8.5 Relationship Between Real and Complex-Valued Operations

Consider now an arbitrary vector $\mathbf{u} \in \mathbb{C}^n$ (not necessarily a *random* vector), let $A \in \mathbb{C}^{m \times n}$, and suppose that we would like to implement the operation that maps \mathbf{u} to $\mathbf{v} = A\mathbf{u}$. Suppose also that we implement this operation on a DSP which is programmed at a level at which we can't rely on routines that handle complex-valued operations. A natural question is: how do we implement $\mathbf{v} = A\mathbf{u}$ using real-valued operations? More generally, what is the relationship between complex-valued variables and operations with respect to their real-valued counterparts? We need this knowledge in the next section to derive the probability density function of proper Gaussian random vectors.

A natural approach is to define the operation that maps a general complex vector \mathbf{u} into a real vector $\hat{\mathbf{u}}$ according to

$$\hat{\mathbf{u}} = \begin{bmatrix} \mathbf{u}_R \\ \mathbf{u}_I \end{bmatrix} \triangleq \begin{bmatrix} \Re[\mathbf{u}] \\ \Im[\mathbf{u}] \end{bmatrix} \quad (8.6)$$

and hope for the existence of a real-valued matrix \hat{A} such that

$$\hat{\mathbf{v}} = \hat{A}\hat{\mathbf{u}}.$$

From $\hat{\mathbf{v}}$ we can then immediately obtain \mathbf{v} . Fortunately such a matrix exists and it is straightforward to verify that

$$\hat{A} = \begin{bmatrix} A_R & -A_I \\ A_I & A_R \end{bmatrix} \triangleq \begin{bmatrix} \Re[A] & -\Im[A] \\ \Im[A] & \Re[A] \end{bmatrix}. \quad (8.7)$$

A set of operations on complex-valued vectors and matrices and the corresponding real-valued operations are described in the following Lemma.

LEMMA 60. *The following properties hold:*

$$\widehat{AB} = \hat{A}\hat{B} \quad (8.8a)$$

$$\widehat{A+B} = \hat{A} + \hat{B} \quad (8.8b)$$

$$\widehat{A^\dagger} = \hat{A}^\dagger \quad (8.8c)$$

$$\widehat{A^{-1}} = \hat{A}^{-1} \quad (8.8d)$$

$$\det(\hat{A}) = |\det(A)|^2 = \det(AA^\dagger) \quad (8.8e)$$

$$\widehat{\mathbf{u} + \mathbf{v}} = \hat{\mathbf{u}} + \hat{\mathbf{v}} \quad (8.8f)$$

$$\widehat{A\mathbf{u}} = \hat{A}\hat{\mathbf{u}} \quad (8.8g)$$

$$\Re(\mathbf{u}^\dagger \mathbf{v}) = \hat{\mathbf{u}}^\dagger \hat{\mathbf{v}} \quad (8.8h)$$

Proof. The properties (8.8a), (8.8b) and (8.8c) are immediate. For instance, property (8.8a) is verified as follows:

$$\begin{aligned}\widehat{AB} &= \begin{bmatrix} (AB)_R & -(AB)_I \\ (AB)_I & (AB)_R \end{bmatrix} \\ &= \begin{bmatrix} A_R B_R - A_I B_I & -A_R B_I - A_I B_R \\ A_R B_I + A_I B_R & A_R B_R - A_I B_I \end{bmatrix} \\ &= \begin{bmatrix} A_R & -A_I \\ A_I & A_R \end{bmatrix} \begin{bmatrix} B_R & -B_I \\ B_I & B_R \end{bmatrix} \\ &= \widehat{A}\widehat{B}\end{aligned}$$

Property (8.8d) follows from (8.8a) and the fact that $\widehat{I}_n = I_{2n}$. To prove (8.8e) we use the fact that the determinant of a product is the product of the determinant and the determinant of a block triangular matrix is the product of the determinants of the diagonal blocks. Hence:

$$\det(\widehat{A}) = \det \left(\begin{bmatrix} I & jI \\ 0 & I \end{bmatrix} \widehat{A} \begin{bmatrix} I & -jI \\ 0 & I \end{bmatrix} \right) = \det \left(\begin{bmatrix} A & 0 \\ \Im(A) & A^* \end{bmatrix} \right) = \det(A) \det(A)^*.$$

Properties (8.8f), (8.8g) and (8.8h) are immediate. \square

COROLLARY 61. *If $U \in \mathbb{C}^{n \times n}$ is unitary then $\widehat{U} \in \mathbb{R}^{2n \times 2n}$ is orthonormal.*

Proof. $U^\dagger U = I_n \iff (\widehat{U})^\dagger \widehat{U} = \widehat{I}_n = I_{2n}$. \square

COROLLARY 62. *If $Q \in \mathbb{C}^{n \times n}$ is non-negative definite, then so is $\widehat{Q} \in \mathbb{R}^{2n \times 2n}$. Moreover, $\mathbf{u}^\dagger Q \mathbf{u} = \widehat{\mathbf{u}}^\dagger \widehat{Q} \widehat{\mathbf{u}}$.*

Proof. Assume that Q is non-negative definite. Then $\mathbf{u}^\dagger Q \mathbf{u}$ is a non-negative real-valued number for all $\mathbf{u} \in \mathbb{C}^n$. Hence,

$$\begin{aligned}\mathbf{u}^\dagger Q \mathbf{u} &= \Re\{\mathbf{u}^\dagger (Q \mathbf{u})\} = \widehat{\mathbf{u}}^\dagger (\widehat{Q \mathbf{u}}) \\ &= \widehat{\mathbf{u}}^\dagger \widehat{Q} \widehat{\mathbf{u}}\end{aligned}$$

where in the last two equalities we used (8.8h) and (8.8g), respectively. \square

EXERCISE 63. *A random vector \mathbf{U} is proper iff $2K_{\widehat{\mathbf{U}}} = \widehat{K}_{\mathbf{U}}$.*

8.6 Complex-Valued Gaussian Random Variables

A complex-valued Gaussian random vector \mathbf{U} is defined as a vector with jointly Gaussian real and imaginary parts. Following Feller³, we consider Gaussian distributions to include

³Feller, *An Introduction to Probability Theory and its Applications*, vol. II. New York: Wiley, 1966, p. 86.

degenerate distributions concentrated on a lower-dimensional manifold, i.e., when the $2n \times 2n$ -covariance matrix

$$\text{cov} \left(\begin{bmatrix} \mathbf{U}_R \\ \mathbf{U}_I \end{bmatrix}, \begin{bmatrix} \mathbf{U}_R \\ \mathbf{U}_I \end{bmatrix} \right) = \begin{bmatrix} K_{\mathbf{U}_R} & K_{\mathbf{U}_I \mathbf{U}_R} \\ K_{\mathbf{U}_I \mathbf{U}_R} & K_{\mathbf{U}_I} \end{bmatrix}$$

is singular and the pdf does not exist unless one admits generalized functions.

Hence, by definition, a complex-valued random vector $\mathbf{U} \in \mathbb{C}^n$ with nonsingular covariance matrix $K_{\hat{\mathbf{U}}}$ is Gaussian iff

$$f_{\mathbf{U}}(\mathbf{u}) = f_{\hat{\mathbf{U}}}(\hat{\mathbf{u}}) = \frac{1}{[\det(2\pi K_{\hat{\mathbf{U}}})]^{1/2}} e^{-\frac{1}{2}(\hat{\mathbf{u}} - \hat{\mathbf{m}})^T K_{\hat{\mathbf{U}}}^{-1}(\hat{\mathbf{u}} - \hat{\mathbf{m}})}. \quad (8.9)$$

THEOREM 64. Let $\mathbf{U} \in \mathbb{C}^n$ be a proper Gaussian random vector with mean \mathbf{m} and nonsingular covariance matrix $K_{\mathbf{U}}$. Then the pdf of \mathbf{U} is given by

$$f_{\mathbf{U}}(\mathbf{u}) = f_{\hat{\mathbf{U}}}(\hat{\mathbf{u}}) = \frac{1}{\det(\pi K_{\mathbf{U}})} e^{-(\mathbf{u} - \mathbf{m})^\dagger K_{\mathbf{U}}^{-1}(\mathbf{u} - \mathbf{m})}. \quad (8.10)$$

Conversely, let the pdf of a random \mathbf{U} be given by (8.10) where $K_{\mathbf{U}}$ is some Hermitian and positive definite matrix. Then \mathbf{U} is proper and Gaussian with covariance matrix $K_{\mathbf{U}}$ and mean \mathbf{m} .

Proof. If \mathbf{U} is proper then by Exercise 63

$$\sqrt{\det 2\pi K_{\hat{\mathbf{U}}}} = \sqrt{\det \pi \hat{K}_{\mathbf{U}}} = |\det \pi K_{\mathbf{U}}| = \det \pi K_{\mathbf{U}},$$

where the last equality holds since the determinant of an Hermitian matrix is always real. Moreover, letting $\hat{\mathbf{v}} = \hat{\mathbf{u}} - \hat{\mathbf{m}}$, again by Exercise 63

$$\hat{\mathbf{v}}^\dagger (2K_{\hat{\mathbf{U}}})^{-1} \hat{\mathbf{v}} = \hat{\mathbf{v}}^\dagger (\hat{K}_{\mathbf{U}})^{-1} \hat{\mathbf{v}} = \mathbf{v}^\dagger (K_{\mathbf{U}})^{-1} \mathbf{v}$$

where for last equality we used Corollary 62 and the fact that if a matrix is positive definite, so is its inverse. Using the last two relationships in (8.9) yields the direct part of the theorem. The converse follows similarly. \square

Notice that two jointly proper Gaussian random vectors \mathbf{U} and \mathbf{V} are independent, iff $K_{\mathbf{UV}} = 0$, which follows from Lemma 55 and the fact that uncorrelatedness and independence are equivalent for Gaussian random variables.

8.6.1 Densities after Linear transformations of complex random variables

We know that \mathbf{X} is a real random vector with density $p_{\mathbf{X}}$, and if A is a non-singular matrix, then the density of $\mathbf{Y} = A\mathbf{X}$ is given by

$$p_{\mathbf{Y}}(\mathbf{y}) = |\det(A)|^{-1} p_{\mathbf{X}}(A^{-1}\mathbf{y}).$$

Now, if \mathbf{Z} is a complex random vector with density $p_{\mathbf{Z}}$ and if A is a complex non-singular matrix, then $W = AZ$ is again a complex random vector with

$$\begin{bmatrix} \Re\{W\} \\ \Im\{W\} \end{bmatrix} = \begin{bmatrix} \Re\{A\} & -\Im\{A\} \\ \Im\{A\} & \Re\{A\} \end{bmatrix} \begin{bmatrix} \Re\{Z\} \\ \Im\{Z\} \end{bmatrix}$$

and thus the density of W will be given by

$$p_{\mathbf{W}}(\mathbf{w}) = \left| \det \left(\begin{bmatrix} \Re\{A\} & -\Im\{A\} \\ \Im\{A\} & \Re\{A\} \end{bmatrix} \right) \right|^{-1} p_{\mathbf{Z}}(A^{-1}\mathbf{w}).$$

From the earlier lecture notes we know that

$$\det \left(\begin{bmatrix} \Re\{A\} & -\Im\{A\} \\ \Im\{A\} & \Re\{A\} \end{bmatrix} \right) = |\det(A)|^2,$$

and thus the transformation formula becomes

$$p_{\mathbf{W}}(\mathbf{w}) = |\det(A)|^{-2} p_{\mathbf{Z}}(A^{-1}\mathbf{w}). \quad (8.11)$$

8.7 Circular Symmetry

We say that a complex valued random variable Z is *circularly symmetric* if for any $\theta \in [0, 2\pi)$, the distribution of $Ze^{j\theta}$ is the same as the distribution of Z .

Using the linear transformation formula (8.11), we see that the density of Z must satisfy

$$p_Z(z) = p_Z(z \exp(j\theta))$$

for all θ , and thus, p_Z must not depend on the phase of its argument, i.e.,

$$p_Z(z) = p_Z(|z|).$$

We can also conclude that, if Z is circularly symmetric,

$$E[Z] = E[e^{j\theta}Z] = e^{j\theta}E[Z],$$

and taking $\theta = \pi$, we conclude that $E[Z] = 0$. Similarly, $E[Z^2] = 0$.

For (complex) random *vectors*, the definition of circular symmetry is that the distribution of \mathbf{Z} should be the same as the distribution of $e^{j\theta}\mathbf{Z}$. In particular, by taking $\theta = \pi$, we see that

$$E[\mathbf{Z}] = 0,$$

and by taking $\theta = \pi/2$, we see that the pseudo covariance

$$J_{\mathbf{Z}} = E[\mathbf{Z}\mathbf{Z}^T] = 0.$$

We have shown that if \mathbf{Z} is circularly symmetric, then it is also zero mean and proper.

If \mathbf{Z} is a zero-mean Gaussian random vector, then the converse is also true, i.e., properness implies circular symmetry. To see this let \mathbf{Z} be zero-mean proper and Gaussian. Then $e^{-j\theta}\mathbf{Z}$ is also zero-mean and Gaussian. Hence \mathbf{Z} and $e^{-j\theta}\mathbf{Z}$ have the same density iff they have the same covariance and pseudo-covariance matrices. The pseudo-covariance matrices vanish in both cases (\mathbf{Z} is proper and $e^{-j\theta}\mathbf{Z}$ is also proper since it is the linear transformation of a proper random vector). Using the definition, one immediately sees that \mathbf{Z} and $e^{-j\theta}\mathbf{Z}$ have the same covariance matrix. Hence they have the same density.

Appendix 8.A A Few Facts of Linear Algebra

DEFINITION 65. A matrix $U \in \mathbb{C}^{n \times n}$ is said to be unitary if $U^\dagger U = I$. If, in addition, $U \in \mathbb{R}^{n \times n}$, U is said to be orthogonal.

The following theorem lists a number of handy facts about unitary matrices. Most of them are straightforward.

THEOREM 66. If $U \in \mathbb{C}^{n \times n}$, the following are equivalent:

- (a) U is unitary;
- (b) U is nonsingular and $U^\dagger = U^{-1}$;
- (c) $UU^\dagger = I$;
- (d) U^\dagger is unitary;
- (e) The columns of U form an orthonormal set;
- (f) The rows of U form an orthonormal set; and
- (g) For all $x \in \mathbb{C}^n$ the Euclidean length of $y = Ux$ is the same as that of x ; that is, $y^\dagger y = x^\dagger x$.

LEMMA 67. (Schur) For any square matrix $A \in \mathbb{C}^{n \times n}$ there exists a unitary V and upper triangular R such that

$$A = VRV^\dagger.$$

Proof. See Homework. □

DEFINITION 68. A matrix $A \in \mathbb{C}^{n \times n}$ is said to be Hermitian if $A = A^\dagger$. It is said to be Skew-Hermitian if $A = -A^\dagger$.

Recall that an $n \times n$ matrix has exactly n eigenvalues in \mathbb{C} .

LEMMA 69. Let $H \in \mathbb{C}^{n \times n}$ be Hermitian. Then

- (i) All n eigenvalues of H are real.
- (ii) H has a set of eigenvectors $\{\mathbf{u}_i : i = 1, \dots, n\}$ that form an orthonormal basis of \mathbb{C}^n .

Proof. Using the Schur lemma there exists a unitary V and upper triangular R such that $H = VRV^\dagger$. Since H is Hermitian,

$$VRV^\dagger = (VRV^\dagger)^\dagger = VR^\dagger V^\dagger$$

from which we obtain $R = R^\dagger$. But since R is upper triangular, this implies that R is real and diagonal. It is now easy to verify that the columns of V are the eigenvectors, with the i th column of V being an eigenvector with eigenvalue R_{ii} . \square

Notice that all covariance matrices are Hermitian.

EXERCISE 70. Show that if $A \in \mathbb{C}^{n \times n}$ is Hermitian, then $\mathbf{u}^\dagger A \mathbf{u}$ is real for all $\mathbf{u} \in \mathbb{C}^n$.

A class of Hermitian matrices with a special positivity property arises naturally in many applications, including in communication theory. They provide one generalization to matrices of the notion of positive numbers.

DEFINITION 71. An Hermitian matrix $A \in \mathbb{C}^{n \times n}$ is said to be positive definite if

$$\mathbf{u}^\dagger A \mathbf{u} > 0 \quad \text{for all non zero } \mathbf{u} \in \mathbb{C}^n.$$

If the above strict inequality is weakened to $\mathbf{u}^\dagger A \mathbf{u} \geq 0$, then A is said to be positive semidefinite. Implicit in these defining inequalities is the observation that if A is Hermitian, the left hand side is always a real number.

Appendix 8.B Linear Transformations (*)

The Fourier transform is a useful tool in dealing with linear time-invariant (LTI) systems. This is so since the input/output relationship of a LTI system is easily described in the Fourier domain. In this section we learn that this is just a special case of a more general principle that applies to linear transformations (not necessarily time-invariant). Key ingredients are the eigenvectors.

8.B.1 Linear Transformations, Toeplitz, and Circulant Matrices

A linear transformation from \mathbb{C}^n to \mathbb{C}^n can be described by an $n \times n$ matrix H . If the matrix is *Toeplitz*, meaning that $H_{ij} = h_{i-j}$, then the transformation which sends $\mathbf{u} \in \mathbb{C}^n$ to $\mathbf{v} = H\mathbf{u}$ can be described by the *convolution* sum

$$v_i = \sum_k h_{i-k} u_k.$$

A Toeplitz matrix is a matrix which is constant along its diagonals.

In this section we focus attention on Toeplitz matrices of a special kind called circulant. A matrix H is *circulant* if $H_{ij} = h_{[i-j]}$ where here and hereafter the operator $[\cdot]$ applied to an index denotes the index taken modulo n . When H is circulant, the operation that maps \mathbf{u} to $\mathbf{v} = H\mathbf{u}$ may be described by the *circulant convolution*

$$v_i = \sum_k h_{[i-k]} u_k.$$

EXAMPLE 72.

$$H = \begin{bmatrix} 3 & 1 & 5 \\ 5 & 3 & 1 \\ 1 & 5 & 3 \end{bmatrix} \text{ is a circulant matrix.}$$

A circulant matrix H is completely described by its first column \mathbf{h} (or any column or row for that matter).

8.B.2 The DFT

The discrete Fourier transform of a vector $\mathbf{u} \in \mathbb{C}^n$ is the vector $\mathbf{U} \in \mathbb{C}^n$ defined by

$$\begin{aligned} \mathbf{U} &= F^\dagger \mathbf{u} \\ F &= (\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n) \\ \mathbf{f}_i &= \frac{1}{\sqrt{n}} \begin{bmatrix} \beta^{i0} \\ \beta^{i1} \\ \vdots \\ \beta^{i(n-1)} \end{bmatrix} \quad i = 1, 2, \dots, n, \end{aligned} \tag{8.12}$$

where $\beta = e^{j\frac{2\pi}{n}}$ is the primitive n -th root of unity in \mathbb{C} . Notice that $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ is an orthonormal basis for \mathbb{C}^n .

Usually, the DFT is defined without the \sqrt{n} in (8.12) and with a factor $\frac{1}{n}$ (instead of $1/\sqrt{n}$) in the inverse transform. The resulting transformation is not orthonormal, and a factor n must be inserted in Parseval's identity when it is applied to the DFT. In this class we call $F^\dagger \mathbf{u}$ the DFT of \mathbf{u} .

8.B.3 Eigenvectors of Circulant Matrices

LEMMA 73. Any circulant matrix $H \in \mathbb{C}^{n \times n}$ has exactly n (normalized) eigenvectors which may be taken as $\mathbf{f}_1, \dots, \mathbf{f}_n$. Moreover, the vector of eigenvalues $(\lambda_1, \dots, \lambda_n)^T$ equals \sqrt{n} times the DFT of the first column of H , namely $\sqrt{n}F^\dagger \mathbf{h}$.

EXAMPLE 74. Consider the matrix

$$H = \begin{bmatrix} h_0 & h_1 \\ h_1 & h_0 \end{bmatrix} \in \mathbb{C}^{2 \times 2}.$$

This is a circulant matrix. Hence

$$f_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad f_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

are eigenvectors and the eigenvalues are

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \sqrt{2} F^\dagger \mathbf{h} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} h_0 - h_1 \\ h_0 + h_1 \end{bmatrix}$$

indeed

$$H \mathbf{f}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} h_0 - h_1 \\ h_1 - h_0 \end{bmatrix} = \frac{h_0 - h_1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \lambda_1 \mathbf{f}_1$$

and

$$H \mathbf{f}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} h_0 + h_1 \\ h_1 + h_0 \end{bmatrix} = \frac{h_0 + h_1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda_2 \mathbf{f}_2$$

Proof.

$$\begin{aligned} (H \mathbf{f}_i)_k &= \frac{1}{\sqrt{n}} \sum_{e=0}^{n-1} h_{k-e} \beta^{ie} \\ &= \left(\sum_{m=0}^{n-1} h_m \beta^{-im} \right) \frac{1}{\sqrt{n}} \beta^{ik} \\ &= \sqrt{n} \mathbf{f}_i^\dagger \mathbf{h} \frac{1}{\sqrt{n}} \beta^{ik} = \lambda_i \frac{1}{\sqrt{n}} \beta^{ik}, \end{aligned}$$

where $\lambda_i = \sqrt{n} \mathbf{f}_i^\dagger \mathbf{h}$. Going to vector notation we obtain $H \mathbf{f}_i = \lambda_i \mathbf{f}_i$. □

8.B.4 Eigenvectors to Describe Linear Transformations

When the eigenvectors of a transformation $H \in \mathbb{C}^{n \times n}$ (not necessarily Toeplitz) span \mathbb{C}^n , both the vectors and the transformation can be represented with respect to a basis of eigenvectors. In that new basis the transformation takes the form $H' = \text{diag}(\lambda_1, \dots, \lambda_n)$, where $\text{diag}(\cdot)$ denotes a matrix with the arguments on the main diagonal and 0s elsewhere, and λ_i is the eigenvalue of the i -th eigenvector. In the new basis the input/output relationship is

$$\mathbf{v}' = H' \mathbf{u}'$$

or equivalently, $v'_i = \lambda_i u'_i$, $i = 1, 2, \dots, n$. To see this, let $\varphi_i, i = 1 \dots n$, be n eigenvectors of H spanning \mathbb{C}^n . Letting $\mathbf{u} = \sum_i \varphi_i u'_i$ and $\mathbf{v} = \sum_i \varphi_i v'_i$ and plugging into $H \mathbf{u}$ we obtain

$$H \mathbf{u} = H \left(\sum_i \varphi_i u'_i \right) = \sum_i H \varphi_i u'_i = \sum_i \varphi_i \lambda_i u'_i$$

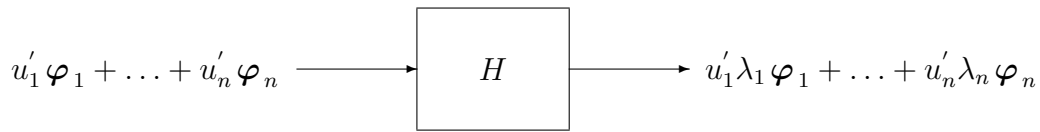


Figure 8.1: Input/output representation via eigenvectors.

showing that $v'_i = \lambda_i u'_i$.

Notice that the key aspects in the proof are the linearity of the transformation and the fact that $\varphi_i u'_i$ is sent to $\varphi_i \lambda_i u'_i$, as shown in Fig. 8.1.

It is often convenient to use matrix notation. To see how the proof goes with matrix notation we define $\Phi = (\varphi_1, \dots, \varphi_n)$ as the matrix whose columns span \mathbb{C}^n . Then $\mathbf{u} = \Phi \mathbf{u}'$ and the above proof in matrix notation is

$$\mathbf{v} = H\mathbf{u} = H\Phi\mathbf{u}' = \Phi H'\mathbf{u}',$$

showing that $\mathbf{v}' = H'\mathbf{u}'$.

For the case where H is circulant, $\mathbf{u} = F\mathbf{u}'$ and $\mathbf{v} = F\mathbf{v}'$. Hence $\mathbf{u}' = F^\dagger\mathbf{u}$ and $\mathbf{v}' = F^\dagger\mathbf{v}$ are the DFT of \mathbf{u} and \mathbf{v} , respectively. Similarly, the diagonal elements of H' are \sqrt{n} times the DFT of the first column of H . Hence the above representation via the new basis says (the well-known result) that a circular convolution corresponds to a multiplication in the DFT domain.

Appendix 8.C Karhunen-Loève Expansion (*)

In Appendix 8.B, we have seen that the eigenvectors of a linear transformation H can be used as a basis and in the new basis the linear transformation of interest becomes a componentwise multiplication.

A similar idea can be used to describe a random vector \mathbf{u} as a linear combination of deterministic vectors with orthogonal random coefficient. Now the eigenvectors are those of the correlation matrix $r_{\mathbf{u}}$. The procedure, that we now describe, is the Karhunen-Loève expansion.

Let $\varphi_1, \dots, \varphi_n$ be a set of eigenvectors of $r_{\mathbf{u}}$ that form an orthonormal basis of \mathbb{C}^n . Such a set exists since $r_{\mathbf{u}}$ is Hermitian. Hence

$$\lambda_i \varphi_i = r_{\mathbf{u}} \varphi_i, i = 1, 2, \dots, n$$

or, in matrix notation,

$$\Phi\Lambda = r_{\mathbf{u}}\Phi$$

where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ and $\Phi = [\varphi_1, \dots, \varphi_n]$ is the matrix whose columns are the eigenvectors. Since the eigenvectors are orthonormal, Φ is unitary (i.e. $\Phi^\dagger \Phi = I$). Solving for Λ we obtain

$$\Lambda = \Phi^\dagger r_{\mathbf{u}} \Phi.$$

Notice that if we solve for $r_{\mathbf{u}}$ we obtain $r_{\mathbf{u}} = \Phi \Lambda \Phi^\dagger$ which is the well known result that an Hermitian matrix can be diagonalized.

Since Φ forms a basis of \mathbb{C}^n we can write

$$\mathbf{u} = \Phi \mathbf{u}' \tag{8.13}$$

for some vector of coefficient \mathbf{u}' with correlation matrix

$$\begin{aligned} r_{\mathbf{u}'} &= E[\mathbf{u}'(\mathbf{u}')^\dagger] = \Phi^\dagger E[\mathbf{u}\mathbf{u}^\dagger] \Phi = \Phi^\dagger r_{\mathbf{u}} \Phi \\ &= \Lambda \end{aligned}$$

Hence (8.13) expresses \mathbf{u} as a linear combination of deterministic vectors $\varphi_1, \dots, \varphi_n$ with orthogonal random coefficients u'_1, \dots, u'_n . This is the Karhunen-Loève expansion of \mathbf{u} .

If $r_{\mathbf{u}}$ is circulant, then $\Phi = F$ and $\mathbf{u}' = \Phi^\dagger \mathbf{u}$ is the DFT of \mathbf{u} .

REMARK 75. $\|\mathbf{u}\|^2 = \|\mathbf{u}'\|^2 = \sum |u'_i|^2$. Also $E\|\mathbf{u}\|^2 = \sum \lambda_i$.

Appendix 8.D Circularly Wide-Sense Stationary Random Vectors (*)

We consider random vectors in \mathbb{C}^n . We will continue using the notation that \mathbf{u} and \mathbf{U} denotes DFT pairs. Observe that if \mathbf{U} is random then \mathbf{u} is also random. This forces us to abandon the convention that we use capital letters for random variables.

The following definitions are natural.

DEFINITION 76. A random vector $\mathbf{u} \in \mathbb{C}^n$ is circularly wide sense stationary (c.w.s.s.) if

$$\begin{aligned} m_{\mathbf{u}} &\stackrel{\Delta}{=} E[\mathbf{u}] \text{ is a constant vector} \\ r_{\mathbf{u}} &\stackrel{\Delta}{=} E[\mathbf{u}\mathbf{u}^\dagger] \text{ is a circulant matrix} \\ s_{\mathbf{u}} &\stackrel{\Delta}{=} E[\mathbf{u}\mathbf{u}^T] \text{ is a circulant matrix} \end{aligned}$$

DEFINITION 77. A random vector \mathbf{u} is uncorrelated if $K_{\mathbf{u}}$ and $J_{\mathbf{u}}$ are diagonal.

We will call $r_{\mathbf{u}}$ and $s_{\mathbf{u}}$ the circular correlation matrix and circular pseudo-correlation matrix, respectively.

THEOREM 78. Let $\mathbf{u} \in \mathbb{C}^n$ be a zero-mean proper random vector and $\mathbf{U} = F^\dagger \mathbf{u}$ be its DFT. Then \mathbf{u} is c.w.s.s. iff \mathbf{U} is uncorrelated. Moreover,

$$r_{\mathbf{u}} = \text{circ}(\mathbf{a}) \quad (8.14)$$

if and only if

$$r_{\mathbf{U}} = \sqrt{n} \text{diag}(\mathbf{A}) \quad (8.15)$$

for some \mathbf{a} and its DFT \mathbf{A} .

Proof. Let \mathbf{u} be a zero-mean proper random vector. If \mathbf{u} is c.w.s.s. then we can write $r_{\mathbf{u}} = \text{circ}(\mathbf{a})$ for some vector \mathbf{a} . Then, using Lemma 73,

$$\begin{aligned} r_{\mathbf{U}} &\triangleq E[F^\dagger \mathbf{u} \mathbf{u}^\dagger F] = F^\dagger r_{\mathbf{u}} F \\ &= F^\dagger \sqrt{n} F \text{diag}(F^\dagger \mathbf{a}) \\ &= \sqrt{n} \text{diag}(\mathbf{A}), \end{aligned}$$

proving (8.15). Moreover, $m_{\mathbf{U}} = 0$ since $m_{\mathbf{u}} = 0$ and therefore $s_{\mathbf{U}} = J_{\mathbf{U}}$. But $J_{\mathbf{U}} = 0$ since the properness of \mathbf{u} and Lemma 57 imply the properness of \mathbf{U} . Conversely, let $r_{\mathbf{U}} = \text{diag}(\mathbf{A})$. Then

$$r_{\mathbf{u}} = E[\mathbf{u} \mathbf{u}^\dagger] = F r_{\mathbf{U}} F^\dagger.$$

Due to the diagonality of $r_{\mathbf{U}}$, the element (k, l) of $r_{\mathbf{u}}$ is

$$\begin{aligned} \sqrt{n} \sum_m F_{k,m} A_m (F^\dagger)_{m,l} &= \sum_m F_{k,m} F_{l,m}^* A_m \sqrt{n} \\ &= \frac{1}{\sqrt{n}} \sum_m A_m e^{j \frac{2\pi}{n} m(k-l)} \\ &= a_{k-l} \end{aligned}$$

□

Chapter 9

The Baseband Equivalent Channel Model

We have already derived the impulse response of the baseband-equivalent channel (see (7.1)). What remains to be derived are the properties of the baseband-equivalent additive noise $N_E(t)$. Then the picture of the baseband-equivalent channel model shown in Fig. 9.1 will be complete.

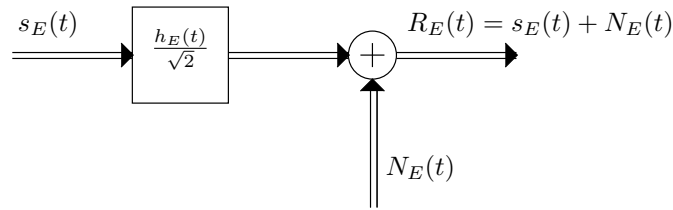


Figure 9.1: Baseband-equivalent communication system.

First, $N_E(t)$ is clearly a zero-mean (complex-valued) Gaussian random process since it is obtained from linear (complex-valued) operations on Gaussian noise. Furthermore:

- (a) $\hat{N}(t)$ is a Gaussian process (since it is obtained by filtering a Gaussian noise process). Its power spectral density is

$$\mathcal{S}_{\hat{N}}(f) = \mathcal{S}_N(f) \left| \sqrt{2} h_{>, \mathcal{F}}(f) \right|^2 = \begin{cases} 2\mathcal{S}_N(f), & f > 0 \\ \frac{1}{2}\mathcal{S}_N(f), & f = 0 \\ 0, & f < 0. \end{cases} \quad (9.1)$$

- (b) Let $N_E(t) = \hat{N}(t) e^{-j2\pi f_0 t}$ be the baseband-equivalent noise. The autocorrelation of $N_E(t)$ is given by:

$$\begin{aligned} \mathcal{R}_{N_E}(\tau) &= E \left[\hat{N}(t + \tau) e^{-j2\pi f_0(t+\tau)} \hat{N}^*(t) e^{j2\pi f_0 t} \right] \\ &= R_{\hat{N}}(\tau) e^{-j2\pi f_0 \tau} \end{aligned} \quad (9.2)$$

where we have used the fact that $\hat{N}(t)$ is WSS (since it is obtained from filtering a WSS process). We see that $N_E(t)$ is itself WSS. Its power spectral density is given by:

$$\mathcal{S}_{N_E}(f) = \mathcal{S}_{\hat{N}}(f + f_0) = \begin{cases} 2\mathcal{S}_N(f + f_0), & f > -f_0 \\ \frac{1}{2}\mathcal{S}_N(f + f_0), & f = -f_0 \\ 0, & f < -f_0. \end{cases} \quad (9.3)$$

(c) We now show that $\hat{N}(t)$ is proper.

$$\begin{aligned} E[\hat{N}(t)\hat{N}(s)] &= E \left[\int_{-\infty}^{+\infty} \sqrt{2}h_{>}(\alpha)N(t - \alpha) d\alpha \int_{-\infty}^{+\infty} \sqrt{2}h_{>}(\beta)N(s - \beta) d\beta \right] \\ &= 2 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h_{>}(\alpha)h_{>}(\beta)\mathcal{R}_N(t - \alpha - s + \beta) d\alpha d\beta \\ &= 2 \int_{\alpha} \int_{\beta} h_{>}(\alpha)h_{>}(\beta) d\alpha d\beta \int_{-\infty}^{+\infty} \mathcal{S}_N(f) e^{j2\pi f(t - \alpha - s + \beta)} df \\ &= 2 \int_f \mathcal{S}_N(f) e^{j2\pi f(t - s)} h_{>,\mathcal{F}}(f)h_{>,\mathcal{F}}(-f) df \\ &= 0 \end{aligned} \quad (9.4)$$

since $h_{>,\mathcal{F}}(f)h_{>,\mathcal{F}}(f) = 0$ for all frequencies except for $f = 0$. Hence the integral vanishes. Thus $\hat{N}(t)$ is proper.

(d) $N_E(t)$ is also proper since

$$\begin{aligned} E[N_E(t)N_E(s)] &= E \left[\hat{N}(t) e^{-j2\pi f_0 t} \hat{N}(s) e^{-j2\pi f_0 s} \right] \\ &= e^{-j2\pi f_0(t+s)} E \left[\hat{N}(t)\hat{N}(s) \right] \\ &= 0 \end{aligned} \quad (9.5)$$

(We could have simply argued that $N_E(t)$ is proper since it is obtained from the proper process $\hat{N}(t)$ via a linear transformation.)

(e) The real and imaginary components of $N_E(t)$ have the same autocorrelation function. Indeed,

$$\begin{aligned} 0 = E[N_E(t)N_E(s)] &= E [(\Re \{N_E(t)\} \Re \{N_E(s)\} - \Im \{N_E(t)\} \Im \{N_E(s)\}) \\ &\quad + j (\Re \{N_E(t)\} \Im \{N_E(s)\} + \Im \{N_E(t)\} \Re \{N_E(s)\})] \end{aligned} \quad (9.6)$$

implies

$$E [(\Re \{N_E(t)\} \Re \{N_E(s)\})] = E [\Im \{N_E(t)\} \Im \{N_E(s)\}]$$

As claimed.

(f) Furthermore, if $\mathcal{S}_z(f_0 - f) = \mathcal{S}_z(f_0 + f)$ then the real and imaginary parts of $N_E(t)$ are uncorrelated, hence they are independent. To see this we expand as follows

$$\begin{aligned} E[N_E(t)N_E^*(s)] &= E[(\Re\{N_E(t)\}\Re\{N_E(s)\} + \Im\{N_E(t)\}\Im\{N_E(s)\}) \\ &\quad - j(\Re\{N_E(t)\}\Im\{N_E(s)\} - \Im\{N_E(t)\}\Re\{N_E(s)\})]. \end{aligned} \quad (9.7)$$

and observe that if the power spectral density of $Z_E(t)$ is symmetric (that is $\mathcal{S}_Z(f_0 - f) = \mathcal{S}_Z(f_0 + f)$), the autocorrelation of $N_E(t)$ is real-valued. Thus

$$E[\Re\{N_E(t)\}\Im\{N_E(s)\} - \Im\{N_E(t)\}\Re\{N_E(s)\}] = 0. \quad (9.8)$$

On the other hand, from (9.6) we have

$$E[\Re\{N_E(t)\}\Im\{N_E(s)\} + \Im\{N_E(t)\}\Re\{N_E(s)\}] = 0. \quad (9.9)$$

The last two expressions imply

$$E[\Re\{N_E(t)\}\Im\{N_E(s)\}] = E[\Im\{N_E(t)\}\Re\{N_E(s)\}] = 0, \quad (9.10)$$

which is what we have claimed.

We summarize: We have shown that $N_E(t)$ is a proper zero-mean Gaussian random process. Hence it is circularly symmetric. Furthermore, from (9.3) we see that the power spectral density of $N_E(t)$ equals that of $N(t)$ translated towards baseband by f_0 and scaled by a factor 2. To remember where the factor 2 goes, it suffices to keep in mind that the variance of the noise within the band of interest is the same for both processes. To find the variance of $N(t)$ in the band of interest we have to integrate its power spectral density over $2B$ Hz. For that $N_E(t)$ we have to integrate over B Hz. Hence the power spectral density of $N_E(t)$ must be twice that of $N(t)$.

The real and imaginary parts of $N_E(t)$ have the same autocorrelation functions hence the same power spectral densities. If $\mathcal{S}(f)$ is symmetric with respect to f_0 , then the real and imaginary parts of $N_E(t)$ are uncorrelated, and since they are Gaussian they are independent. In this case their power spectral density must be half that of $N_E(t)$, i.e., equal that of $N(t)$ around f_0 .

Chapter 10

A Case Study (to be written)

Consider a passband AWGN channel of bandwidth B . The corresponding baseband-equivalent channel is a lowpass AWGN channel of bandwidth $B/2$ (where the additive white Gaussian noise is complex-valued).

From the sampling theorem, a lowpass signal $s_E(t)$ may always be written as

$$s_E(t) = \sum_{j=0}^{N-1} s_j \psi(t - jT_s), \quad (10.1)$$

where $\psi(t)$ is a sinc and $s_j = s_E(jT)/\sqrt{T}$. More generally, $\psi(t)$ is a Nyquist pulse.

Motivated by this, let us agree that our baseband equivalent form for a general transmit signal is as in the above expression with a real-valued $\psi(t)$ that fulfills Nyquist criterion and complex-valued symbols s_j . Then

$$\begin{aligned} S(t) &= \sqrt{2} \Re\{s_E(t) e^{j2\pi f_0 t}\} \\ &= \sqrt{2} \cos(2\pi f_0 t) \Re\{s_E(t)\} - \sqrt{2} \sin(2\pi f_0 t) \Im\{s_E(t)\} \\ &= \sqrt{2} \cos(2\pi f_0 t) \sum_{j=0}^{N-1} \Re\{s_j\} \psi(t - jT_s) \\ &\quad - \sqrt{2} \sin(2\pi f_0 t) \sum_{j=0}^{N-1} \Im\{s_j\} \psi(t - jT_s). \end{aligned} \quad (10.2)$$

This is called QAM (quadrature amplitude modulation). Fig. 10.1 shows a conceptual block diagram for the implementation of the transmitter and the maximum likelihood receiver (front-end). For no particular reason other than to show various implementation possibilities, the figure shows the transmitter in a form that uses only real-valued operations and signals, whereas the receiver uses complex-valued notation.

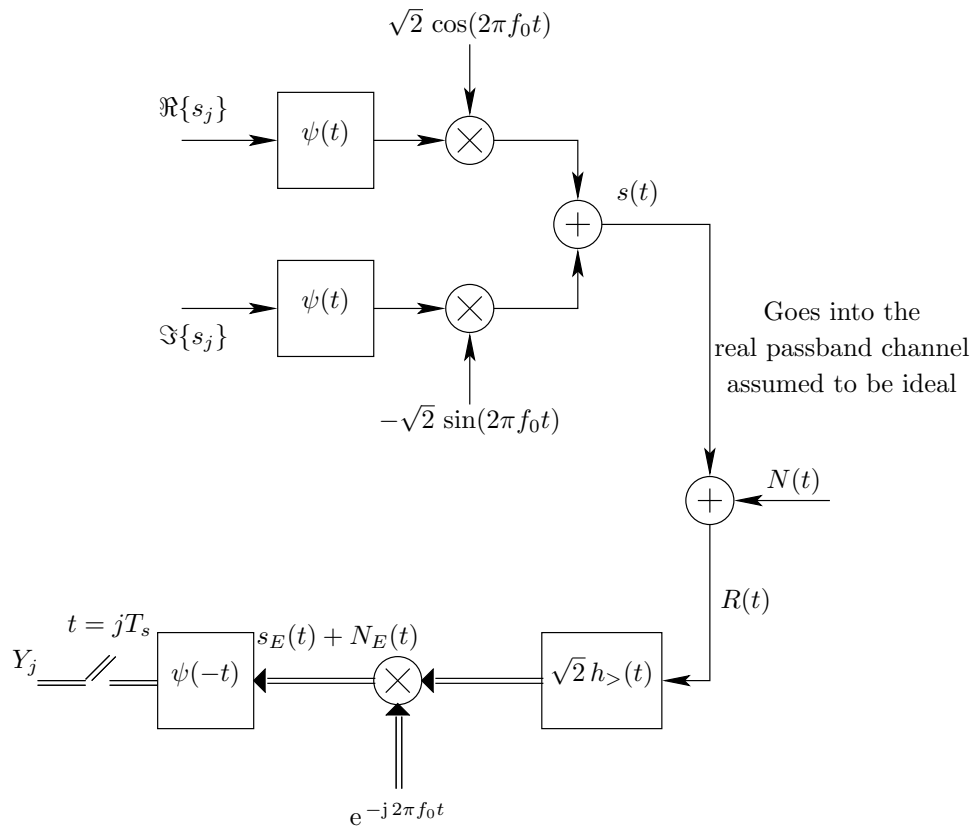


Figure 10.1: Bandpass system.

Fig 10.2 give a specific example of the signal constellation (per complex dimension), namely when s_j takes on the four possible values $1 + j$, $1 - j$, $-1 - j$, $-1 + j$. The result is 4-ary QAM.

When we compare the compactness of the baseband equivalent signal (10.1) to the form of the actual transmit signal (10.2) we understand why we like to work with the former, or even with the n -tuple \mathbf{s} . In designing and analyzing a system, we should use the simplest possible model we have. For instance, if we are studying the power spectral density of the transmit signal we should work with (10.1). If we are analyzing the error probability (under ideal conditions as we have assumed so far, i.e. no timing uncertainty) we should use the vector (n -tuple) model

$$\mathbf{Y} = \mathbf{s}_i + \mathbf{Z}$$

where \mathbf{Z} is a complex-valued zero-mean proper Gaussian random vector which independent of the transmit signal. If $N(t)$ is white, then Z_j is i.i.d. and its real and imaginary components are independent. We should use this model also to design the signal constellation, the encoder (if there is one), and the structure of the ML receiver (beyond the part which has to do with acquiring the sufficient statistic $\dots Y_{j-1}, Y_j, Y_{j+1} \dots$ i.e., the receiver front-end).

We use the description of the actual transmit signal when we are dealing with the im-

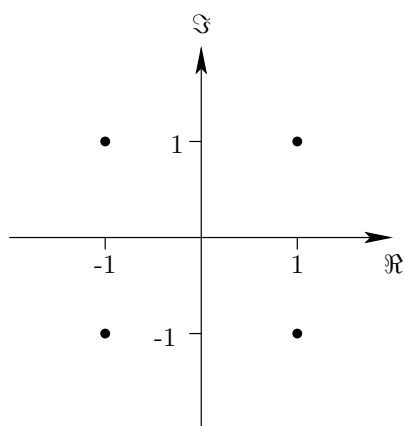


Figure 10.2: 4-ary QAM constellation.

plementation of the up/down converters. From (10.2) we see how to implement the up/converter using only real-valued operations. This is shown in Fig. 10.1. The receiver front end in Fig. 10.1 is still based on the corresponding complex-valued operations. It can also be translated into corresponding real-valued operations.

Towards this, we observe that the analytic filter with impulse response $h_>(t)$ does nothing to the signal of interest. In fact, the matched filter with impulse response $\psi(-t)$ puts to zero all frequency components that are outside the lowpass frequency range of interest $[-\frac{B}{2}, \frac{B}{2}]$. But the analytic filter affects the signal only in the frequency range $f < -f_0$. This has no effect at the output of the matched filter since $-f_0 \ll -\frac{B}{2}$. Fig. 10.3 shows the receiver front-end without the analytic filter. It is still based on complex-valued notation, but using the relationship $e^{-j2\pi f_0 t} = \cos(2\pi f_0 t) - j \sin(2\pi f_0 t)$ and the fact that $R(t)$ is real-valued, you should see immediately how to split the complex-valued part in two real-valued, paths (like in the up-converter).

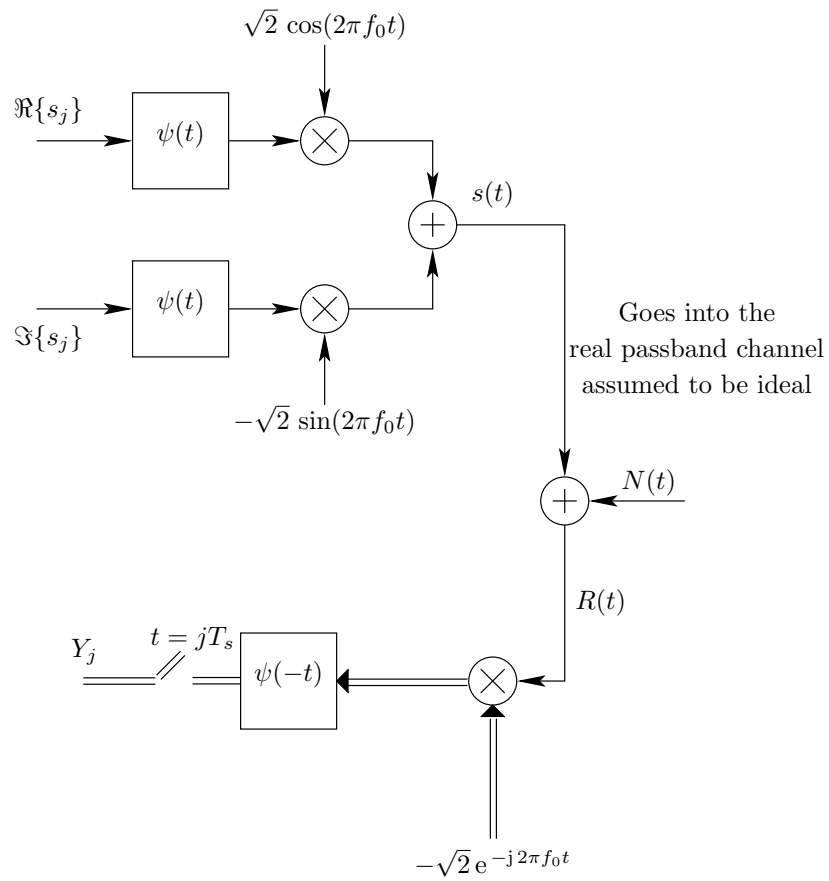


Figure 10.3: Bandpass system.

Index

signal space, 79

signal space constellation, 109

Bibliography

- [1] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge: Cambridge University Press, 1999.