A FOUNDATION IN DIGITAL COMMUNICATION

This intuitive but rigorous introduction derives the core results and engineering schemes of digital communication from first principles. Theory, rather than industry standards, motivates the engineering approaches, and key results are stated with all the required assumptions.

The book emphasizes the geometric view, opening with the inner product, the matched filter for its computation, Parseval's theorem, the sampling theorem as an orthonormal expansion, the isometry between passband signals and their baseband representation, and the spectral-efficiency optimality of quadrature amplitude modulation (QAM). Subsequent chapters address noise, with a comprehensive study of hypothesis testing, Gaussian stochastic processes, the sufficiency of the matched filter outputs, and some coding theory.

New is a treatment of white noise without generalized functions and a presentation of the power spectral density without artificial random jitters and random phases in the analysis of QAM.

This systematic and insightful book – with over 300 exercises – is ideal for graduate courses in digital communication, and for anyone asking "why" and not just "how."

AMOS LAPIDOTH received his Ph.D. in electrical engineering from Stanford University. He was an assistant and associate professor at the Massachusetts Institute of Technology, and is currently Professor of Information Theory at ETH Zürich, the Swiss Federal Institute of Technology. He is a Fellow of the IEEE.

A FOUNDATION IN DIGITAL COMMUNICATION

AMOS LAPIDOTH
ETH Zürich, Swiss Federal Institute of Technology

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521193955

First published 2009

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-19395-5 hardback

Additional resources for this publication at www.cambridge.org/9780521193955

To my family

Contents

Cambridge University Press
978-0-521-19395-5 - A Foundation in Digital Communication
Amos Lapidoth
Frontmatter
More information

Preface

Claude Shannon, the father of Information Theory, described the fundamental problem of point-to-point communications in his classic 1948 paper as "that of reproducing at one point either exactly or approximately a message selected at another point." How engineers solve this problem is the subject of this book. But unlike Shannon's general problem, where the message can be an image, a sound clip, or a movie, here we restrict ourselves to bits. We thus envision that the original message is either a binary sequence to start with, or else that it was described using bits by a device outside our control and that our job is to reproduce the describing bits with high reliability. The issue of how images or text files are converted efficiently into bits is the subject of lossy and lossless data compression and is addressed in texts on information theory and on quantization.

The engineering solutions to the point-to-point communication problem greatly depend on the available resources and on the channel between the points. They typically bring together beautiful techniques from Fourier Analysis, Hilbert Spaces, Probability Theory, and Decision Theory. The purpose of this book is to introduce the reader to these techniques and to their interplay.

The book is intended for advanced undergraduates and beginning graduate students. The key prerequisites are basic courses in Calculus, Linear Algebra, and Probability Theory. A course in Linear Systems is a plus but not a must, because all the results from Linear Systems that are needed for this book are summarized in Chapters 5 and 6. But more importantly, the book requires a certain mathematical maturity and patience, because we begin with first principles and develop the theory before discussing its engineering applications. The book is for those who appreciate the views along the way as much as getting to the destination; who like to "stop and smell the roses;" and who prefer fundamentals to acronyms. I firmly believe that those with a sound foundation can easily pick up the acronyms and learn the jargon on the job, but that once one leaves the academic environment, one rarely has the time or peace of mind to study fundamentals.

In the early stages of the planning of this book I took a decision that greatly influenced the project. I decided that every key concept should be unambiguously defined; that every key result should be stated as a mathematical theorem; and that every mathematical theorem should be correct. This, I believe, makes for a solid foundation on which one can build with confidence. But it is also a tall order. It required that I scrutinize each "classical" result before I used it in order to be sure that I knew what the needed qualifiers were, and it forced me to include background material to which the reader may have already been exposed, because I needed the results "done right." Hence Chapters 5 and 6 on Linear Systems and Fourier Analysis. This is also partly the reason why the book is so long. When I started out my intention was to write a much shorter book. But I found that to do justice to the beautiful mathematics on which Digital Communications is based I had to expand the book.

Most physical layer communication problems are at their core of a continuous-time nature. The transmitted physical waveforms are functions of time and not sequences synchronized to a clock. But most solutions first reduce the problem to a discrete-time setting and then solve the problem in the discrete-time domain. The reduction to discrete-time often requires great ingenuity, which I try to describe. It is often taken for granted in courses that open with a discrete-time model from Lecture 1. I emphasize that most communication problems are of a continuous-time nature, and that the reduction to discrete-time is not always trivial or even possible. For example, it is extremely difficult to translate a peak-power constraint (stating that at no epoch is the magnitude of the transmitted waveform allowed to exceed a given constant) to a statement about the sequence that is used to represent the waveform. Similarly, in Wireless Communications it is often very difficult to reduce the received waveform to a sequence without any loss in performance.

The quest for mathematical precision can be demanding. I have therefore tried to precede the statement of every key theorem with its gist in plain English. Instructors may well choose to present the material in class with less rigor and direct the students to the book for a more mathematical approach. I would rather have textbooks be more mathematical than the lectures than the other way round. Having a rigorous textbook allows the instructor in class to discuss the intuition knowing that the students can obtain the technical details from the book at home.

The communication problem comes with a beautiful geometric picture that I try to emphasize. To appreciate this picture one needs the definition of the inner product between energy-limited signals and some of the geometry of the space of energy-limited signals. These are therefore introduced early on in Chapters 3 and 4. Chapters 5 and 6 cover standard material from Linear Systems. But note the early introduction of the matched filter as a mechanism for computing inner products in Section 5.8. Also key is Parseval's Theorem in Section 6.2.2 which relates the geometric pictures in the time domain and in the frequency domain.

Chapter 7 deals with passband signals and their baseband representation. We emphasize how the inner product between passband signals is related to the inner product between their baseband representations. This elegant geometric relationship is often lost in the haze of various trigonometric identities. While this topic is important in wireless applications, it is not always taught in a first course in Digital Communications. Instructors who prefer to discuss baseband communication only can skip Chapters 7, 9, 16, 17, 18, 24 27, and Sections 26.10 and 28.5. But it would be a shame.

Chapter 8 presents the celebrated Sampling Theorem from a geometric perspective. It is inessential to the rest of the book but is a striking example of the geometric approach. Chapter 9 discusses the Sampling Theorem for passband signals.

Chapter 10 discusses modulation. I have tried to motivate Linear Modulation and Pulse Amplitude Modulation and to minimize the use of the "that's just how it is done" argument. The use of the Matched Filter for detecting (here in the absence of noise) is emphasized. This also motivates the Nyquist Theory, which is treated in Chapter 11. I stress that the motivation for the Nyquist Theory is not to avoid inter-symbol interference at the sampling points but rather to guarantee the orthogonality of the time shifts of the pulse shape by integer multiples of the baud period. This ultimately makes more engineering sense and leads to cleaner mathematics: compare Theorem 11.3.2 with its corollary, Corollary 11.3.4.

The result of modulating random bits is a stochastic process, a concept which is first encountered in Chapter 10; formally defined in Chapter 12; and revisited in Chapters 13, 17, and 25. It is an important concept in Digital Communications, and I find it best to first introduce man-made synthesized stochastic processes (as the waveforms produced by an encoder when fed random bits) and only later to introduce the nature-made stochastic processes that model noise. Stationary discrete-time stochastic processes are introduced in Chapter 13 and their complex counterparts in Chapter 17. These are needed for the analysis in Chapter 14 of the power in Pulse Amplitude Modulation and for the analysis in Chapter 17 of the power in Quadrature Amplitude Modulation.

I emphasize that power is a physical quantity that is related to the time-averaged energy in the continuous-time transmitted power. Its relation to the power in the discrete-time modulating sequence is a nontrivial result. In deriving this relation I refrain from adding random timing jitters that are often poorly motivated and that turn out to be unnecessary. (The transmitted power does not depend on the realization of the fictitious jitter.) The Power Spectral Density in Pulse Amplitude Modulation and Quadrature Amplitude Modulation is discussed in Chapters 15 and 18. The discussion requires a definition for Power Spectral Density for nonstationary processes (Definitions 15.3.1 and 18.4.1) and a proof that this definition coincides with the classical definition when the process is wide-sense stationary (Theorem 25.14.3).

Chapter 19 opens the second part of the book, which deals with noise and detection. It introduces the univariate Gaussian distribution and some related distributions. The principles of Detection Theory are presented in Chapters 20–22. I emphasize the notion of Sufficient Statistics, which is central to Detection Theory. Building on Chapter 19, Chapter 23 introduces the all-important multivariate Gaussian distribution. Chapter 24 treats the complex case.

Chapter 25 deals with continuous-time stochastic processes with an emphasis on stationary Gaussian processes, which are often used to model the noise in Digital Communications. This chapter also introduces white Gaussian noise. My approach to this topic is perhaps new and is probably where this text differs the most from other textbooks on the subject.

I define **white Gaussian noise of double-sided power spectral density** $\mathsf{N}_0/2$ **with respect to the bandwidth** W as any measurable,[1] stationary, Gaussian stochastic process whose power spectral density is a nonnegative, symmetric, inte-

[1]This book does not assume any Measure Theory and does not teach any Measure Theory. (I do define sets of Lebesgue measure zero in order to be able to state uniqueness theorems.) I

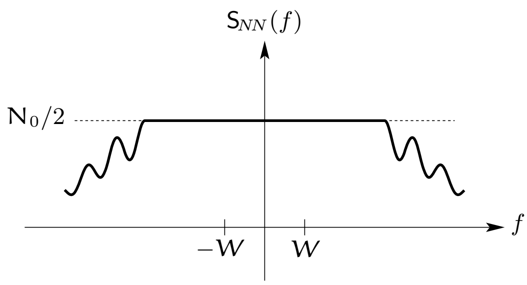


Figure 1: The power spectral density of a white Gaussian noise process of double-sided power spectral density $N_0/2$ with respect to the bandwidth W.

grable function of frequency that is equal to $N_0/2$ at all frequencies f satisfying $|f| \leq W$. The power spectral density at other frequencies can be arbitrary. An example of the power spectral density of such a process is depicted in Figure 1. Adopting this definition has a number of advantages. The first is, of course, that such processes exist. One need not discuss "generalized processes," Gaussian processes with infinite variances (that, by definition, do not exist), or introduce the Itô calculus to study stochastic integrals. (Stochastic integrals with respect to the Brownian motion are mathematically intricate and physically unappealing. The idea of the noise having infinite power is ludicrous.) The above definition also frees me from discussing Dirac's Delta, and, in fact, Dirac's Delta is never used in this book. (A rigorous treatment of Generalized Functions is beyond the engineering curriculum in most schools, so using Dirac's Delta always gives the reader the unsettling feeling of being on unsure footing.)

The detection problem in white Gaussian noise is treated in Chapter 26. No course in Digital Communications should end without Theorem 26.4.1. Roughly speaking, this theorem states that if the mean-signals are bandlimited to W Hz and if the noise is white Gaussian noise with respect to the bandwidth W, then the inner products between the received signal and the mean-signals form a sufficient statistic. Numerous examples as well as a treatment of colored noise are also discussed in this chapter. Extensions to noncoherent detection are addressed in Chapter 27 and implications for Pulse Amplitude Modulation and for Quadrature Amplitude Modulation in Chapter 28.

The book concludes with Chapter 29, which introduces Coding. It emphasizes how the code design influences the transmitted power, the transmitted power spectral density, the required bandwidth, and the probability of error. The construction of good codes is left to texts on Coding Theory.

use Measure Theory only in stating theorems that require measurability assumptions. This is in line with my attempt to state theorems together with all the assumptions that are required for their validity. I recommend that students ignore measurability issues and just make a mental note that whenever measurability is mentioned there is a minor technical condition lurking in the background.

Basic Latin

Mathematics sometimes reads like a foreign language. I therefore include here a short glossary for such terms as "*i.e.*," "*that is*," "*in particular*," "*a fortiori*," "*for example*," and "*e.g.*," whose meaning in Mathematics is slightly different from the definition you will find in your English dictionary. In mathematical contexts these terms are actually logical statements that the reader should verify. Verifying these statements is an important way to make sure that you understand the math.

What are these logical statements? First note the synonym "*i.e.*" = "*that is*" and the synonym "*e.g.*" = "*for example*." Next note that the term "*that is*" often indicates that the statement following the term is equivalent to the one preceding it: "We next show that p is a prime, *i.e.*, that p is a positive integer that is not divisible by any number other than one and itself." The terms "*in particular*" or "*a fortiori*" indicate that the statement following them is implied by the one preceding them: "Since $g(\cdot)$ is differentiable and, *a fortiori*, continuous, it follows from the Mean Value Theorem that the integral of $g(\cdot)$ over the interval $[0, 1]$ is equal to $g(\xi)$ for some $\xi \in [0, 1]$." The term "*for example*" can have its regular day-to-day meaning but in mathematical writing it also sometimes indicates that the statement following it implies the one preceding it: "Suppose that the function $g(\cdot)$ is monotonically nondecreasing, *e.g.*, that it is differentiable with a nonnegative derivative."

Another important word to look out for is "*indeed*," which in this book typically signifies that the statement just made is about to be expanded upon and explained. So when you read something that is unclear to you, be sure to check whether the next sentence begins with the word "indeed" before you panic.

The Latin phrases "*a priori*" and "*a posteriori*" show up in Probability Theory. The former is usually associated with the unconditional probability of an event and the latter with the conditional. Thus, the "*a priori*" probability that the sun will shine this Sunday in Zurich is 25%, but now that I know that it is raining today, my outlook on life changes and I assign this event the *a posteriori* probability of 15%.

The phrase "*prima facie*" is roughly equivalent to the phrase "before any further mathematical arguments have been presented." For example, the definition of the projection of a signal $\mathbf{v}$ onto the signal $\mathbf{u}$ as the vector $\mathbf{w}$ that is collinear with $\mathbf{u}$ and for which $\mathbf{v} - \mathbf{w}$ is orthogonal to $\mathbf{u}$, may be followed by the sentence: "*Prima facie*, it is not clear that the projection always exists and that it is unique. Nevertheless, as we next show, this is the case."

Syllabuses or Syllabi

The book can be used as a textbook for a number of different courses. For a course that focuses on deterministic signals one could use Chapters 1–9 & Chapter 11. A course that covers Stochastic Processes and Detection Theory could be based on Chapter 12 and Chapters 19–26 with or without discrete-time stochastic processes (Chapter 13) and with or without complex random variables and processes

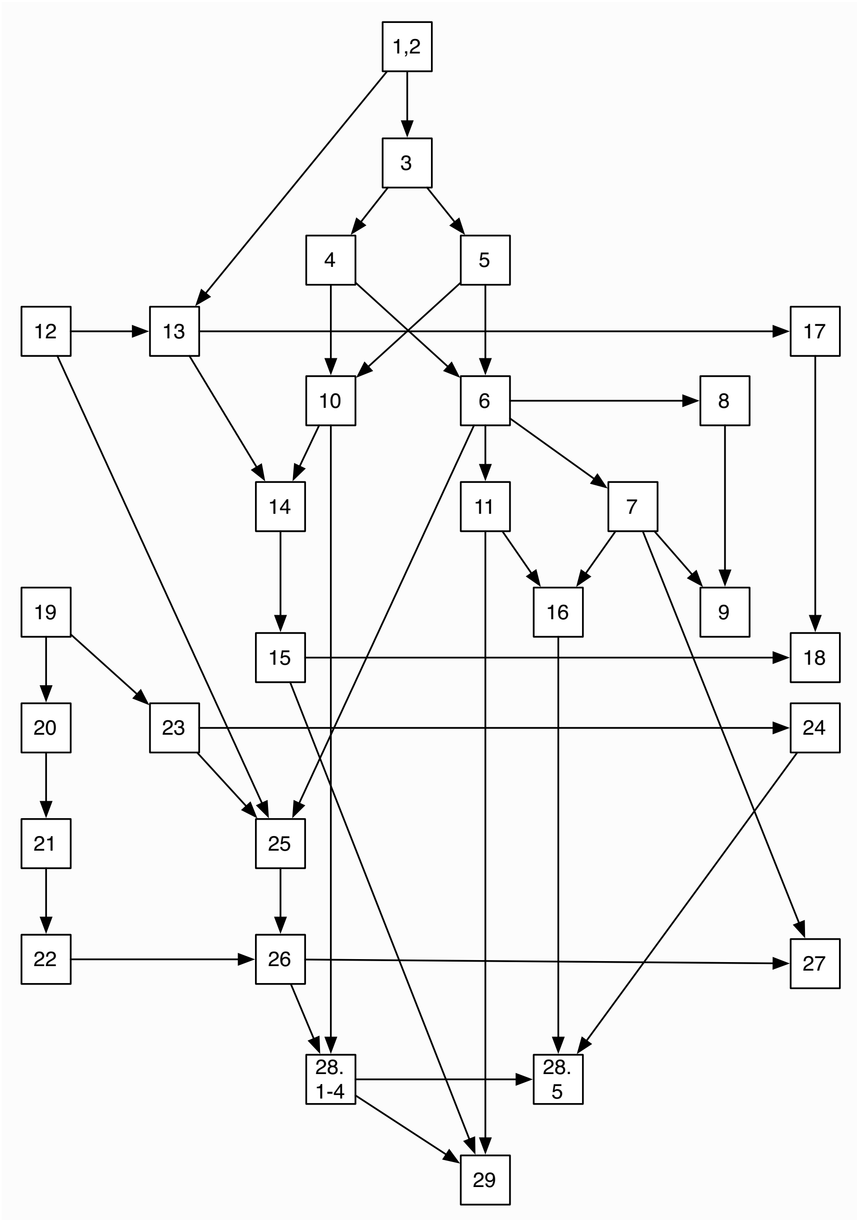
(Chapters 17 & 24).

For a course on Digital Communications one could use the entire book or, if time does not permit it, discuss only baseband communication. In the latter case one could omit Chapters 7, 9, 16, 17, 18, 24, 27, and Section 28.5,

The dependencies between the chapters are depicted on Page xxiii.

A web page for this book can be found at

`www.afoundationindigitalcommunication.ethz.ch`



A Dependency Diagram.

Acknowledgments

This book has a long history. Its origins are in a course entitled "Introduction to Digital Communication" that Bob Gallager and I developed at the Massachusetts Institute of Technology (MIT) in the years 1997 (course number 6.917) and 1998 (course number 6.401). Assisting us in these courses were Emre Koksal and Poompat Saengudomlert (Tengo) respectively. The course was first conceived as an advanced undergraduate course, but at MIT it has since evolved into a first-year graduate course leading to the publication of the textbook (Gallager, 2008). At ETH the course is still an advanced undergraduate course, and the lecture notes evolved into the present book. Assisting me at ETH were my former and current Ph.D. students Stefan Moser, Daniel Hösli, Natalia Miliou, Stephan Tinguely, Tobias Koch, Michèle Wigger, and Ligong Wang. I thank them all for their enormous help. Marion Brändle was also a great help.

I also thank Bixio Rimoldi for his comments on an earlier draft of this book, from which he taught at École Polytechnique Fédérale de Lausanne (EPFL) and Thomas Mittelholzer, who used a draft of this book to teach a course at ETH during my sabbatical.

Extremely helpful were discussions with Amir Dembo, Sanjoy Mitter, Alain-Sol Sznitman, and Ofer Zeitouni about some of the more mathematical aspects of this book. Discussions with Ezio Biglieri, Holger Boche, Stephen Boyd, Young-Han Kim, and Sergio Verdú are also gratefully acknowledged.

Special thanks are due to Bob Gallager and Dave Forney with whom I had endless discussions about the material in this book both while at MIT and afterwards at ETH. Their ideas have greatly influenced my thinking about how this course should be taught.

I thank Helmut Bölcskei, Andi Loeliger, and Nikolai Nefedov for having tolerated my endless ramblings regarding Digital Communications during our daily lunches. Jim Massey was a huge help in patiently answering my questions regarding English usage. I should have asked him much more!

A number of dear colleagues read parts of this manuscript. Their comments were extremely useful. These include Helmut Bölcskei, Moritz Borgmann, Samuel Braendle, Shraga Bross, Giuseppe Durisi, Yariv Ephraim, Minnie Ho, Young-Han Kim, Yiannis Kontoyiannis, Nick Laneman, Venya Morgenshtern, Prakash Narayan, Igal Sason, Brooke Shrader, Aslan Tchamkerten, Sergio Verdú, Pascal Vontobel, and Ofer Zeitouni. I am especially indebted to Emre Telatar for his enormous help in all aspects of this project.

I would like to express my sincere gratitude to the Rockefeller Foundation at whose Study and Conference Center in Bellagio, Italy, this all began.

Finally, I thank my wife, Danielle, for her encouragement, her tireless editing, and for making it possible for me to complete this project.