Applied and Numerical Harmonic Analysis

$$\widehat{f}(\gamma) = \int f(x)e^{-2\pi i x\gamma}\,dx$$

Jeffrey A. Hogan
Joseph D. Lakey

Duration and Bandwidth Limiting

Prolate Functions, Sampling, and Applications

Birkhäuser

Birkhäuser

Applied and Numerical Harmonic Analysis

Jeffrey A. Hogan • Joseph D. Lakey

Duration and Bandwidth Limiting

Prolate Functions, Sampling, and Applications

Birkhäuser

Jeffrey A. Hogan
School of Mathematical
and Physical Sciences
University of Newcastle
Callaghan, NSW 2308
Australia
Jeff.Hogan@newcastle.edu.au

Joseph D. Lakey
Department of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003-8001
USA
jlakey@nmsu.edu

ISBN 978-0-8176-8306-1 e-ISBN 978-0-8176-8307-8
DOI 10.1007/978-0-8176-8307-8
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011942155

Mathematics Subject Classification (2010): 42A10, 42C05, 65T50, 94A12, 94A20

Printed on acid-free paper

Birkhäuser Boston is part of Springer Science+Business Media (www.birkhauser.com)

In memory of our fathers, Ron Hogan and Frank Lakey

ANHA Series Preface

The *Applied and Numerical Harmonic Analysis (ANHA)* book series aims to provide the engineering, mathematical, and scientific communities with significant developments in harmonic analysis, ranging from abstract harmonic analysis to basic applications. The title of the series reflects the importance of applications and numerical implementation, but richness and relevance of applications and implementation depend fundamentally on the structure and depth of theoretical underpinnings. Thus, from our point of view, the interleaving of theory and applications and their creative symbiotic evolution is axiomatic.

Harmonic analysis is a wellspring of ideas and applicability that has flourished, developed, and deepened over time within many disciplines and by means of creative cross-fertilization with diverse areas. The intricate and fundamental relationship between harmonic analysis and fields such as signal processing, partial differential equations (PDEs), and image processing is reflected in our state-of-the-art *ANHA* series.

Our vision of modern harmonic analysis includes mathematical areas such as wavelet theory, Banach algebras, classical Fourier analysis, time–frequency analysis, and fractal geometry, as well as the diverse topics that impinge on them.

For example, wavelet theory can be considered an appropriate tool to deal with some basic problems in digital signal processing, speech and image processing, geophysics, pattern recognition, biomedical engineering, and turbulence. These areas implement the latest technology from sampling methods on surfaces to fast algorithms and computer vision methods. The underlying mathematics of wavelet theory depends not only on classical Fourier analysis, but also on ideas from abstract harmonic analysis, including von Neumann algebras and the affine group. This leads to a study of the Heisenberg group and its relationship to Gabor systems, and of the metaplectic group for a meaningful interaction of signal decomposition methods. The unifying influence of wavelet theory in the aforementioned topics illustrates the justification for providing a means for centralizing and disseminating information from the broader, but still focused, area of harmonic analysis. This will be a key role of *ANHA*. We intend to publish the scope and interaction that such a host of issues demands.

Along with our commitment to publish mathematically significant works at the frontiers of harmonic analysis, we have a comparably strong commitment to publish major advances in the following applicable topics in which harmonic analysis plays a substantial role:

Antenna theory	*Prediction theory*
Biomedical signal processing	*Radar applications*
Digital signal processing	*Sampling theory*
Fast algorithms	*Spectral estimation*
Gabor theory and applications	*Speech processing*
Image processing	*Time–frequency and*
Numerical partial differential equations	*time-scale analysis*
	Wavelet theory

The above point of view for the *ANHA* book series is inspired by the history of Fourier analysis itself, whose tentacles reach into so many fields.

In the last two centuries, Fourier analysis has had a major impact on the development of mathematics, on the understanding of many engineering and scientific phenomena, and on the solution of some of the most important problems in mathematics and the sciences. Historically, Fourier series were developed in the analysis of some of the classical PDEs of mathematical physics; these series were used to solve such equations. In order to understand Fourier series and the kinds of solutions they could represent, some of the most basic notions of analysis were defined, e.g., the concept of "function". Since the coefficients of Fourier series are integrals, it is no surprise that Riemann integrals were conceived to deal with uniqueness properties of trigonometric series. Cantor's set theory was also developed because of such uniqueness questions.

A basic problem in Fourier analysis is to show how complicated phenomena, such as sound waves, can be described in terms of elementary harmonics. There are two aspects of this problem: first, to find, or even define properly, the harmonics or spectrum of a given phenomenon, e.g., the spectroscopy problem in optics; second, to determine which phenomena can be constructed from given classes of harmonics, as done, e.g., by the mechanical synthesizers in tidal analysis.

Fourier analysis is also the natural setting for many other problems in engineering, mathematics, and the sciences. For example, Wiener's Tauberian theorem in Fourier analysis not only characterizes the behavior of the prime numbers, but also provides the proper notion of spectrum for phenomena such as white light; this latter process leads to the Fourier analysis associated with correlation functions in filtering and prediction problems, and these problems, in turn, deal naturally with Hardy spaces in the theory of complex variables.

Nowadays, some of the theory of PDEs has given way to the study of Fourier integral operators. Problems in antenna theory are studied in terms of unimodular trigonometric polynomials. Applications of Fourier analysis abound in signal processing, whether with the fast Fourier transform (FFT), or filter design, or the adaptive modeling inherent in time–frequency-scale methods such as wavelet theory. The coherent states of mathematical physics are translated and modulated

Fourier transforms, and these are used, in conjunction with the uncertainty principle, for dealing with signal reconstruction in communications theory. We are back to the raison d'être of the *ANHA* series!

University of Maryland *John J. Benedetto*
College Park Series Editor

Preface

John von Neumann developed much of the rigorous formalism for quantum mechanics during the late 1920s [339], including a theory of unbounded operators in a Hilbert space that led to a mathematical justification of Heisenberg's uncertainty principle. In his autobiography [356, p. 97 ff.], Norbert Wiener gave a poignant account of his own role in advancing mathematical ideas parallel to von Neumann's during Wiener's visit to Göttingen in 1925. In contrast to Heisenberg, Wiener was motivated by limitations on simultaneous localization in time and frequency, or *duration and bandwidth limiting* pertaining to macroscopic physical systems. Meanwhile, across the Atlantic, Nyquist [246] and Hartley [137], working at Bell Labs, were laying the foundations for Shannon's communication theory in their accounts of transmission of information over physical communications channels. Hartley also quantified limitations on joint localization in time and frequency, perhaps in a more positive light than suggested by Wiener, by defining a bound on the rate at which information might be transmitted over a band-limited channel. Eventually this bound, expressed in terms of the *time–bandwidth product*, was formalized in the *Bell Labs theory*—a series of papers by Landau, Slepian, and Pollak [195, 196, 303, 309] appearing in the *Bell System Technical Journal* in the early 1960s, with additional follow-on work extending through 1980.

The Bell Labs theory identified those band-limited functions—the *prolate spheroidal wave functions* (PSWFs)—that have the largest proportions of their energies localized in a given time interval. It also quantified the eigenvalue behavior of the time- and frequency-localization operator "$P_\Omega Q_T$" that gives rise to these eigenfunctions. There are approximately $2\Omega T$ orthogonal functions that are band limited to the frequency band $[0,\Omega]$, while having most of their energies localized in the time interval $[-T,T]$. Perhaps the most satisfying account of this $2\Omega T$ principle, at least from a purely mathematical perspective, is due to Landau and Widom (Theorem 1.3.1 in this book). Their result quantified asymptotically, but precisely, the number of eigenvalues of $P_\Omega Q_T$ larger than any given $\alpha \in (0,1)$. The masterful use of the technical machinery of compact self-adjoint operators that went into this result arguably places the Bell Labs theory, as a foundational affirmation of what is observed in the physical universe, alongside von Neumann's work. However, the theory did not lead to immediate widespread use of prolates in signal processing.

This was partly due to the fact that the appropriate parallel development in the *discrete domain* had not yet been made. During the period from the late 1970s to the early 1980s, a number of contributions provided this discrete parallel theory of time and band limiting, e.g., [307], the basic tool being the *discrete prolate spheroidal sequences* (DPSSs). Applications then emerged. Thomson [323] used this discrete theory in his *multitaper* method for spectrum estimation, and Jain and Ranganath [162] formalized the use of DPSSs in signal extrapolation. Ideas of Grünbaum [126] and others continued to provide a firm foundation for further potential applications.

While Thomson's methods caught on quickly, particularly in the geosciences, applications of the theory of time and band limiting to the modeling of communications channels, numerical analysis, and other promising areas, such as tomography, had to wait. There were several reasons for this delay: wavelets attracted much of the attention in the 1980s, and mobile communications were still in their lag phase.

Applications of time and band limiting appear now to be entering a log phase. The IEEE literature on this topic has grown quickly, largely due to potential uses in communications, where multiband signals are also playing a wider role, and in compressed sensing. At the same time, potential uses in numerical analysis are also being identified, stemming from the observation that the PSWFs have zeros that are more evenly spaced than their (Chebyshev and Legendre) polynomial counterparts, making them attractive for spectral element methods. Extension and refinement of the mathematical ideas underlying the Bell Labs theory, particularly in the context of multiband signals, but also involving higher order approximations, will continue in the coming decades.

The need to have a common source for the Bell Labs theory and its extensions, together with a resource for related methods that have been developed in the mathematics and engineering literature since 2000, motivated us to put together this monograph. Here is an outline of the contents.

Chapter 1 lays out the fundamental results of the Bell Labs theory, including a discussion of properties of the PSWFs and properties of the eigenvalues of $P_\Omega Q_T$. A discussion of the parallel theory of DPSSs is also included. Further discussion in the *chapter notes* focuses primarily on the application of the DPSSs to signal extrapolation.

The PSWFs and their values were important enough in classical physics to devote large portions of lengthy monographs to tables of their values, e.g., [54, 231, 318]. Modern applications require not only accuracy, but also computational efficiency. Since the largest eigenvalues of $P_\Omega Q_T$ are very close to one, they are also close to one another. This means that standard numerical tools for estimating finite-dimensional analogues of projections onto individual eigenspaces, such as the singular value decomposition, are numerically unstable. This is one among a raft of issues that led several researchers to reconsider numerical methods for estimating and applying PSWFs. Chapter 2 addresses such numerical analytical aspects of PSWFs, including quadrature formulas associated with prolates and approximations of functions by prolate series. Along with classical references, sources for more recent advances here include [35, 36, 41–44, 123, 178, 280, 296, 331, 364].

Chapter 3 contains a detailed review of Thomson's multitaper method. Through multitapering, PSWFs can be used to model channel characteristics; see, e.g., [124, 179, 211, 294]. However, if such a channel model is intended to be used by a single mobile device under severe power usage constraints, it might be preferable to employ a fixed prior basis in order to encode channel information in a parsimonious fashion. Zemen and Mecklenbräuker [370, 372] proposed variants of the DPS sequences for various approaches to this problem. These variants will also be outlined in Chap. 3.

Chapter 4 contains a development of parallel extensions of the Bell Labs theory to the case of multiband signals. Multiband signals play an increasing role in radio frequency (RF) communications as RF spectrum usage increases. We outline Landau and Widom's [197] extension of the $2\Omega T$ theorem to the case in which the time- and frequency-localization sets S and Σ are finite unions of intervals. Other *non-asymptotic* estimates are given for the number of eigenvalues of $P_\Sigma Q_S$ that are larger than one-half, quantified in terms of the linear distributions of S and Σ.

The second part of Chap. 4 addresses a related problem in the theory of time and multiband limiting of finite signals (i.e., signals defined on a finite set of integers). Candès, Romberg, and Tao [57–60] proved *probabilistic* estimates on the norm of the corresponding time- and band-limiting operator. The main estimate states that if the time- and frequency-localization sets are *sparse* in a suitable sense, then, with overwhelming probability, the localization operator has norm smaller than one-half. The estimates were developed for applications to compressed sensing, in particular to probabilistic recovery of signals with sparse Fourier spectra from random measurements. The techniques involve intricate combinatorial extensions of the same approach that was employed by Landau and Widom in their $2\Omega T$ theorem.

Ever since Shannon's seminal work [293], sampling theory has played a central role in the theory of band-limited functions. Aspects of sampling theory related to numerical quadrature are developed in Sect. 2.3. Chapter 5 reviews other aspects of sampling of band-limited signals that are, conceptually, closely tied to time and band limiting. This includes Landau's necessary conditions for sampling and interpolation [188, 189], phrased in terms of the Beurling densities, whose proofs use the same techniques as those employed in Chap. 4 to estimate eigenvalues for time and band limiting. Results quantifying the structure of sets of sampling and interpolation for Paley–Wiener spaces of band-limited functions are surveyed. The latter part of Chap. 5 addresses methods for sampling multiband signals. Several such methods have been published since the 1990s. A few of these approaches will be addressed in detail, with other approaches outlined in the chapter notes.

Chapter 6 contains several results that represent steps toward a general theory that envelops sampling and time and band limiting, both for band-limited and for multiband signals. The core of the chapter builds on work of Walter and Shen [347] and Khare and George [177], who observed that the integer samples of suitably normalized PSWFs form eigenvectors of the discrete matrix whose entries are correlations of time localizations of shifts of the sinc kernel—the reproducing kernel for the space of band-limited signals. This fact permits a method for constructing approximate time- and band-limiting projections from samples. Other discrete methods for building time- and multiband-limiting projection operators from constituent components are also presented in Chap. 6.

This text is primarily a mathematical monograph. While applications to signal processing play a vital motivational role, those applications also involve technical details whose descriptions would detract from the emphasis on mathematical methods. However, we also want this book to be useful to practitioners who do not prove theorems for a living—so not all of the theorems that are discussed are proved in full. In choosing which results to present in finer detail, our goal has been to maintain focus on central concepts of time and band limiting. We also emphasize real-variable methods over those of complex function theory, in several cases favoring the more intuitive result or argument over the most complete result or rigorous proof. We hope that this approach serves to make the book more accessible to a broader audience, even if it makes the reading less satisfying to a stalwart mathematician.

The core material in each chapter is supplemented to some extent with *notes and auxiliary results* that outline further major mathematical contributions and primary applications. In some cases, proofs of results auxiliary to the core theorems are also included with this supplementary material. Even so, we have made choices regarding which contributions to mention. We easily could have cited over 1000 contributions without having mentioned every important theorem and application pertinent to time and band limiting.

A number of colleagues contributed to this work in substantial ways. We are particularly indebted to Chris Brislawn, Chuck Creusere, Scott Izu, Hans Feichtinger, Charly Gröchenig, Ned Haughton, Chris Heil, Tomasz Hrycak, Richard Laugesen, Kabe Moen, Xiaoping Shen, Adam Sikora and Mark Tygert for the generosity with which they gave advice and technical assistance.

This book would not have been possible without the hospitality and support of the Department of Mathematics at Washington University in St. Louis, where Lakey was on sabbatical in 2009-2010 and Hogan enjoyed an extended visit in early 2010. Both authors are especially grateful to Guido Weiss and Ed Wilson for arranging these visits and for their guidance on a number of matters. Hogan is appreciative of support provided by the University of Newcastle's Centre for Computer-Assisted Research Mathematics and its Applications as well as the Centre for Complex Dynamic Systems and Control.

The staff at Birkhäuser did a wonderful job of keeping us on track through all stages of this project. Special thanks to Ben Cronin and Tom Grasso.

Thanks to Roy and HG for determined drive from the backline. Thanks Max, for always being there. Finally, thanks to Mysie and to Ellen for encouragement and understanding. We owe you.

Newcastle, NSW, Australia *Jeff Hogan*
Las Cruces, New Mexico *Joe Lakey*

Contents

Chapter 1
The Bell Labs Theory

1.1 Information Theory of Time and Band Limiting

Duration limiting, or *time limiting*, refers to restricting a signal by setting its values equal to zero outside of a finite time interval or, more generally, outside of a compact set. Bandwidth limiting, or *band limiting*, refers to restricting a signal by setting its amplitudes equal to zero outside of a finite frequency interval or again, more generally, outside of a compact set. This book addresses primarily the theory of time and band limiting whose core was developed by Landau, Pollak, and Slepian in a series of papers [195, 196, 303, 309] appearing in the *Bell System Technical Journal* in the early to middle part of the 1960s, and a broader body of work that grew slowly but steadily out of that core up until around 1980, with a resurgence since 2000, due in large part to the importance of time and band limiting in wireless communications. The 1960s *Bell Labs theory* of time and band limiting is but one aspect of the Bell Labs information theory. The foundations of this encompassing theory were laid, in large part, in Nyquist's fundamental papers "Certain Topics in Telegraph Transmission Theory" [247], which appeared in the *Transactions of the American Institute of Electrical Engineers* in 1928, and "Certain Factors Affecting Telegraph Speed," published in April 1924 in the *Bell System Technical Journal*, along with Hartley's paper "Transmission of Information," which also appeared in the *Bell System Technical Journal* in 1928 [137]. These papers quantified general ideas that were in the air, though certain specific versions were attributed to Kelvin and Wiener among others. Of course, Claude Shannon's seminal work, "A Mathematical Theory of Communication," which appeared in the *Bell System Technical Journal* in July and October 1948 [293], is often cited as providing the basis for much of modern communications theory. His sampling theory plays a central role in Chap. 5 of this monograph. The works of Nyquist and Hartley however remain, in some ways, more germane to the study at hand.

Although we will comment briefly on other aspects of this early work, the emerging principle that is most fundamental here is summarized in the following quote from Hartley [137]:

J.A. Hogan and J.D. Lakey, *Duration and Bandwidth Limiting: Prolate Functions, Sampling, and Applications*, Applied and Numerical Harmonic Analysis,
DOI 10.1007/978-0-8176-8307-8_1,

> *the total amount of information which may be transmitted over such a* [band-limited, but otherwise distortionless] *system is proportional to the product of the frequency-range which it transmits by the time during which it is available for the transmission.*

Hartley also introduced the fundamental concept of *intersymbol interference* and considered its role in limiting the possible rate at which information could be intelligibly transmitted over a cable.

In his 1928 paper, Nyquist first laid the information-theoretic groundwork for what eventually became known as the Shannon sampling theorem. Therein, Nyquist showed that, in order to reconstruct the original signal, the sampling rate must be at least twice the highest frequency present in the sample. However, the bounds on the rate by which information could be transmitted across a *bandwidth constrained* channel were established in his 1924 paper. Specifically, Nyquist considered the problem of what waveform to use in order to transmit information over such a channel and what method might be optimal for encoding information to be transmitted via such a waveform. When phrased in these terms, the work of Landau, Slepian, and Pollak provided a precise quantitative bound on the capacity of a clear, band-limited channel defined in terms of an ideal lowpass filter to transmit a given amount of information per unit time by providing, in turn, a precise quantification of the dimension of the space of essentially time- and band-limited signals and an identification of what those signals are. As Slepian [308] made abundantly clear, this is an interpretation that applies only to an idealization of physical reality. Nevertheless, the mathematical Bell Labs theory of time and band limiting that will be laid out in this chapter is both elegant and satisfying. The information-theoretic details and consequences of this theory extend well beyond the brief comments just made. Some of the additional sources for these details and further general discussion include [80, 84, 108, 135, 161, 191, 194, 229, 302, 304, 305, 308, 310, 311, 359–361].

1.2 Time- and Frequency-localization Operators

1.2.1 Paley–Wiener Spaces on $\mathbb{R}$ and Localization Projections

We normalize the Fourier transform $\mathscr{F}f$ of $f \in L^2(\mathbb{R})$ as the limit, in $L^2(\mathbb{R})$,

$$(\mathscr{F}f)(\xi) = \widehat{f}(\xi) = \lim_{N\to\infty} \int_{-N}^{N} f(t)\,\mathrm{e}^{-2\pi i t\xi}\,\mathrm{d}t. \tag{1.1}$$

The Paley–Wiener (PW) spaces PW_I are subspaces of $L^2(\mathbb{R})$ consisting of functions whose Fourier transforms vanish off the interval I. If $f \in \mathrm{PW}_I$ then f has an extension to an entire function that can be expressed in terms of its absolutely convergent inverse Fourier integral,

$$f(z) = \int_{-\infty}^{\infty} \widehat{f}(\xi)\, e^{2\pi i z\xi}\, d\xi, \quad (z \in \mathbb{C}).$$

Although function theory underlies much of what follows, its role will not be emphasized. One can associate a Paley–Wiener space PW_Σ to any compact frequency support set Σ by defining it as the range of the *band-limiting* orthogonal projection operator P_Σ defined by

$$(P_\Sigma f)(t) = (\widehat{f}\, \mathbb{1}_\Sigma)^\vee(t)\,, \tag{1.2}$$

where $\mathbb{1}_\Sigma$ denotes the indicator function of the set Σ and $g^\vee$ denotes the inverse Fourier transform of g. In the special case in which $\Sigma = [-\Omega/2, \Omega/2]$ is an interval centered at the origin, we will abbreviate $P_{[-\Omega/2,\Omega/2]}$ as P_Ω and, when $\Omega = 1$, we will simply write $P_1 = P$.

The *time-limiting* operator Q_S is defined by the orthogonal projection

$$(Q_S f)(t) = \mathbb{1}_S(t)\, f(t). \tag{1.3}$$

When S is an interval centered about the origin of length $2T$, we will abbreviate $Q_{[-T,T]}$ to Q_T. This notation is slightly asymmetrical from the band-limiting case. In particular, the time–bandwidth product corresponding to the operator $P_\Omega Q_T$ is $2T\Omega$. There are two important differences between our normalizations and those used in the Bell Labs papers [195, 196, 303, 309]. The first is that Slepian et al. used the Fourier kernel $e^{-it\xi}/\sqrt{2\pi}$ as opposed to the kernel $e^{-2\pi i t\xi}$ used here. The second is that their band-limiting operator was defined in terms of multiplication by $\mathbb{1}_{[-\Omega,\Omega]}$ while their time-limiting operator was given by multiplication by $\mathbb{1}_{[-T/2,T/2]}$.

Time and band limiting then refers to a composition of the form $P_\Sigma Q_S$. Since the separate operators P_Σ and Q_S do not commute, except in trivial cases, $P_\Sigma Q_S$ itself is not a projection. However, it is still self-adjoint as an operator on PW_Σ. Sometimes we will consider the operator $P_\Sigma Q_S P_\Sigma$, which is self-adjoint on all of $L^2(\mathbb{R})$. When Σ is compact, PW_Σ is a reproducing kernel Hilbert space (RKHS) with kernel $K_\Sigma(t,s) = K_\Sigma(t-s)$ where $K_\Sigma(t) = (\mathbb{1}_\Sigma)^\vee(t)$. The kernel of $P_\Sigma Q_S$, as an operator on PW_Σ, is $K_{S,\Sigma}(t,s) = \mathbb{1}_S(s)(\mathbb{1}_\Sigma)^\vee(t-s)$. Since $P_\Sigma Q_S$ is positive definite, it follows from general theory (e.g., [277]) that if S and Σ are both compact then $P_\Sigma Q_S$ is a trace-class operator with

$$\mathrm{tr}\,(P_\Sigma Q_S) = \int_{\mathbb{R}} K_{S,\Sigma}(s,s)\, ds = \int_S |\Sigma|\, ds = |S||\Sigma|,$$

since $K_{S,\Sigma}(s,s) = \mathbb{1}_S(s)(\mathbb{1}_\Sigma)^\vee(0) = |\Sigma|\mathbb{1}_S(s)$. Let $\{\psi_n\}$ be an orthonormal basis of PW_Σ, that is, $\{\widehat{\psi_n}\}$ is an orthonormal basis of $L^2(\Sigma)$. The Hilbert–Schmidt norm $\|P_\Sigma Q_S\|_{HS}$ of $P_\Sigma Q_S$ satisfies

$$\|P_\Sigma Q_S\|_{\mathrm{HS}}^2 = \sum_n \|(P_\Sigma Q_S)\psi_n\|^2 = \sum_n \|\widehat{Q_S\psi_n}\|^2_{L^2(\Sigma)}$$
$$= \sum_n \int_\Sigma \Big| \int_\Sigma \widehat{\mathbb{1}_S}(\xi-\eta)\,\widehat{\psi_n}(\eta)\,\mathrm{d}\eta\Big|^2 \mathrm{d}\xi = \sum_n \int_\Sigma |\langle \widehat{\mathbb{1}_S}(\xi-\cdot), \widehat{\psi_n}\rangle|^2\,\mathrm{d}\xi$$
$$= \int_\Sigma \sum_n |\langle \widehat{\mathbb{1}_S}(\xi-\cdot), \widehat{\psi_n}\rangle|^2\,\mathrm{d}\xi = \int_\Sigma\int_\Sigma |\widehat{\mathbb{1}_S}(\xi-\eta)|^2\,\mathrm{d}\eta\,\mathrm{d}\xi\,.$$

When S is symmetric, $(P_S Q_\Sigma)^* = Q_\Sigma P_S$ is unitarily equivalent to the operator $\mathscr{F}^{-1} Q_\Sigma \mathscr{F}\mathscr{F}^{-1} P_S \mathscr{F} = P_\Sigma Q_S$, and

$$\|P_\Sigma Q_S\|_{\mathrm{HS}}^2 = \|K_{S,\Sigma}\|^2_{L^2(S\times S)} = \int_S\int_S |(\mathbb{1}_\Sigma)^\vee(t-s)|^2\,\mathrm{d}s\,\mathrm{d}t\,.$$

The Bell Labs papers studied the spectral decomposition of $P_\Omega Q_T$ and several mathematical and physical consequences of it. This theory and some extensions will be presented in the rest of this chapter. We begin with a brief review of the eigenfunctions of $P_\Omega Q_T$, the *prolate spheroidal wave functions*.

1.2.2 Prolate Spheroidal Wave Functions

Definition and Commentary

When used in the context of time and band limiting, the term *prolate spheroidal wave function* (PSWF) is not particularly resonant. The notion of a PSWF originated, rather, in the context of solving the wave equation in prolate spheroidal coordinates by means of separation of variables. The prolate spheroidal coordinates (ξ,η,ϕ) are related to the standard Euclidean (x,y,z)-coordinates by the equations

$$\begin{aligned} x &= a\sinh\xi\,\sin\eta\,\cos\phi \\ y &= a\sinh\xi\,\sin\eta\,\sin\phi \\ z &= a\cosh\xi\,\cos\eta\,. \end{aligned}$$

Upon substituting $u=\cosh\xi$ and $v=\cos\eta$, the Laplacian takes the form

$$\nabla^2 f = \frac{1}{a^2(u^2-v^2)}\Big\{\frac{\partial}{\partial u}(u^2-1)\frac{\partial f}{\partial u} + \frac{\partial}{\partial v}(v^2-1)\frac{\partial f}{\partial v} + \frac{u^2-v^2}{(u^2-1)(1-v^2)}\frac{\partial^2 f}{\partial\phi^2}\Big\}.$$

After multiplying by $a^2(u^2-v^2)$, the Helmholtz equation $(\nabla^2+k^2)f=0$ becomes

$$\Big\{\frac{\partial}{\partial u}(u^2-1)\frac{\partial}{\partial u} + \frac{\partial}{\partial v}(v^2-1)\frac{\partial}{\partial v} + \frac{u^2-v^2}{(u^2-1)(1-v^2)}\frac{\partial^2}{\partial\phi^2} + c^2(u^2-v^2)\Big\} f = 0 \tag{1.4}$$

where $c=ak$. Suppose that, for fixed c, one can express f as a product

$$f = R(u)S(v)\begin{cases}\cos m\phi \\ \sin m\phi\end{cases} \qquad m = 0, 1, \ldots.$$

Such an f can satisfy (1.4) if there exists an eigenvalue $\chi = \chi_{mn}$ such that

$$\frac{\mathrm{d}}{\mathrm{d}u}(u^2-1)\frac{\mathrm{d}R}{\mathrm{d}u} - \Big(\chi - c^2u^2 + \frac{m^2}{u^2-1}\Big)R(u) = 0 \quad \text{and} \tag{1.5a}$$

$$\frac{\mathrm{d}}{\mathrm{d}v}(v^2-1)\frac{\mathrm{d}S}{\mathrm{d}v} - \Big(\chi - c^2v^2 + \frac{m^2}{v^2-1}\Big)S(v) = 0. \tag{1.5b}$$

Solutions $R = R_{mn}(u;c)$ in (1.5) are called *radial* PSWFs and solutions $S = S_{mn}(v;c)$ are called *angular* PSWFs. The equations (1.5) are the same except that the variable $u = \cosh(\xi)$ is strictly defined on $[1,\infty)$ while $v = \cos\eta$ is strictly defined in $[-1,1]$. However, as solutions of (1.5), R_{mn} and S_{mn} can be defined on $\mathbb{R}$ and differ from one another by a scale factor. The parameter m is called the *order* of the PSWF. *For time and band limiting, we will be concerned only with the case of the angular functions and $m = 0$.* We will use the variable t for *time* and will be interested in solutions $S_n(t) = S_{0n}(t;c)$ that are eigenfunctions of the prolate operator $\mathscr{P}$,

$$\mathscr{P}S_n(t) = \chi_n S_n(t); \quad \mathscr{P} = \frac{\mathrm{d}}{\mathrm{d}t}(t^2-1)\frac{\mathrm{d}}{\mathrm{d}t} + c^2t^2. \tag{1.6}$$

Further basic properties of the PSWFs $S_{mn}(\cdot;c)$ and $R_{mn}(\cdot;c)$, especially those asymptotic properties useful for applications in mathematical physics, are developed in the monographs of Flammer [101], Meixner and Schäfke [231], Stratton et al. [318], and Morse and Feshbach [240], among other classical sources. To use Slepian and Pollak's phrase [309], "We will draw freely from this literature"—particularly regarding features of PSWFs that are incidental but not central to further discussion.

Various normalizations will be of fundamental importance. These normalizations include endpoint values of the PSWF solutions of (1.6) and normalizations of the eigenvalues. First, the solutions S_{0n} of (1.6) are defined, initially, on $[-1,1]$. These functions extend analytically both in t and in c. Because they are eigenfunctions of a positive definite, self-adjoint operator, basic Sturm–Liouville theory (e.g., [373]) shows that they are orthogonal over $[-1,1]$ and are also complete in $L^2[-1,1]$. It is worth just mentioning for now (see Chap. 2) that the eigenvalues $\chi_n(c)$ are nondegenerate and that S_{0n} is indexed in such a way that χ_n is strictly increasing. In the $c = 0$ limit, S_{0n} (suitably normalized) becomes the nth Legendre polynomial P_n, as will be discussed in more detail momentarily.

In what follows, we will let $\phi_n = \phi_n(t;c)$ be a multiple of $S_{0n}(t;c)$ normalized such that $\|\phi_n\|_{L^2(\mathbb{R})} = 1$. Plots of PSWFs ϕ_n normalized to $\|\phi_n\|_{L^2[-1,1]} = 1$ are shown in Fig. 1.1. We will usually employ the *variant phi*, $\varphi_n = \varphi_n(t;T,\Omega)$, to denote the nth eigenfunction of $P_\Omega Q_T$, also having L^2-norm one. When the values of Ω and T are clear from context or unspecified we will typically omit reference to Ω and T and simply write $\varphi_n(t)$. As will be seen momentarily, when $T = 1$, this notation gives $\varphi_n(t;1,c/\pi) = \phi_n(t;c)$.

Scaling and the Fourier Transform

Suppose that $f:\mathbb{R}\to\mathbb{C}$ is absolutely integrable and that the same applies to the image of f under any second-order, linear differential operator of the form $\rho_0(t)\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\rho_1(t)\frac{\mathrm{d}}{\mathrm{d}t}+\rho_2(t)$ in which the ρ_i are polynomials of degree at most two. Under this hypothesis, the Fourier transform (1.1) of the derivative of f is $2\pi\mathrm{i}\xi\widehat{f}(\xi)$. This follows upon integrating by parts. Interchanging the roles of f and $\widehat{f}$ and using the fact that the kernel of the inverse Fourier transform is the conjugate of that of the forward transform,

$$\left(\frac{\mathrm{d}\widehat{f}}{\mathrm{d}\xi}\right)^{\vee}=-2\pi\mathrm{i}t\,f(t)\quad\text{or}\quad\widehat{(tf)}(\xi)=\frac{\mathrm{i}}{2\pi}\frac{\mathrm{d}\widehat{f}}{\mathrm{d}\xi}.$$

Formally, these facts readily generalize to

$$\mathscr{F}P\left(t,\frac{\mathrm{d}}{\mathrm{d}t}\right)=P\left(\frac{\mathrm{i}}{2\pi}\frac{\mathrm{d}}{\mathrm{d}\xi},2\pi\mathrm{i}\xi\right)\mathscr{F}\tag{1.7}$$

whenever P is a polynomial function of two variables. Differentiation and multiplication by t do not commute. In particular, $(tf)'-tf'=f$ where $f'=\mathrm{d}f/\mathrm{d}t$. Therefore, $P(t,\mathrm{d}/\mathrm{d}t)$ has to be interpreted strictly as a linear combination of monomials in which the operator of the second argument is applied first, and that of the first argument is then applied.

The *Lucky Accident*

One of the fundamental observations made by Slepian and Pollak in [309] is that the differential operator $\mathscr{P}$ in (1.6) commutes with the time-localization operator $Q=Q_1$ and the frequency-localization operator P_a, for a an appropriate fixed multiple of c. Slepian and Pollak simply observed that this "lucky accident," as Slepian [308] called it, followed from general properties relating differential operators to corresponding integral operators. Considerably later, Walter [346] viewed this accident as a characterization of certain differential operators of the form

$$P\left(t,\frac{\mathrm{d}}{\mathrm{d}t}\right)=\rho_0(t)\frac{\mathrm{d}^2}{\mathrm{d}t^2}+\rho_1(t)\frac{\mathrm{d}}{\mathrm{d}t}+\rho_2(t)\tag{1.8}$$

that commute with multiplication by the indicator function of an interval.

Lemma 1.2.1. *A second-order linear differential operator (1.8) commutes with multiplication by the indicator function of $I=[a,b]$ in the sense that, for any $f\in C^2(\mathbb{R})$,*

$$P\left(t,\frac{\mathrm{d}}{\mathrm{d}t}\right)(\mathbb{1}_I f)=\mathbb{1}_I P\left(t,\frac{\mathrm{d}}{\mathrm{d}t}\right)(f),$$

if and only if ρ_0 vanishes at the endpoints of I and the values of ρ_1 and $\mathrm{d}\rho_0/\mathrm{d}t$ coincide at the endpoints of I.

Proof. The distributional derivative of $\mathbb{1}_{[-T,T]}$ is $\delta_{-T}-\delta_T$ and its second derivative is $\delta'_{-T}-\delta'_T$ where the prime denotes differentiation with respect to t. To prove the lemma, first it suffices to assume that I is symmetric, of the form $[-T,T]$ for some $T>0$, so multiplication by $\mathbb{1}_{[-T,T]}$ is Q_T. From the product rule for derivatives, therefore,

$$P\Big(t,\frac{\mathrm{d}}{\mathrm{d}t}\Big)(Q_T f)=\rho_0 f\,(\delta'_{-T}-\delta'_T)+(2\rho_0 f'+\rho_1 f)(\delta_{-T}-\delta_T)+Q_T\Big(P\Big(t,\frac{\mathrm{d}}{\mathrm{d}t}\Big)f\Big).$$

The commutation condition then takes the form

$$\rho_0 f\,(\delta'_{-T}-\delta'_T)+(2\rho_0 f'+\rho_1 f)(\delta_{-T}-\delta_T)=0,\quad (f\in C^2(\mathbb{R})).$$

Restricting to those $f\in C^2(\mathbb{R})$ that are compactly supported in $(0,\infty)$, the commutation condition implies that

$$\rho_0 f\,\delta'_T+(2\rho_0 f'+\rho_1 f)\delta_T=0\,.$$

As distributions, $f\,\delta_T=f(T)\delta_T$ and $f\,\delta'_T=f(T)\delta'_T-f'(T)\delta_T$. Therefore,

$$\rho_0(T)\,f(T)\,\delta'_T-\rho'_0(T)\,f(T)\,\delta_T+(\rho_0(T)\,f'(T)+\rho_1(T)\,f(T))\delta_T=0\,.$$

Taking f such that $f'(T)=0$ but $f(T)\neq 0$ gives

$$\rho_0(T)\,f(T)\,\delta'_T+(\rho_1(T)-\rho'_0(T))\,f(T)\,\delta_T=0\,.$$

Treating δ'_T and δ_T as independent then leads to $\rho_0(T)=0$ and $\rho'_0(T)=\rho_1(T)$. A similar argument gives the corresponding conditions at $-T$ also. □

For quadratic P, the criteria can be reformulated as follows.

Corollary 1.2.2. *If ρ_0 and ρ_1 are quadratic polynomials then $P\big(t,\frac{\mathrm{d}}{\mathrm{d}t}\big)$ in (1.8) commutes with $\mathbb{1}_{[-T,T]}$ if and only if there exist constants a, b such that*

$$\rho_0(t)=a(t^2-T^2)\quad\text{and}\quad\rho_1(t)=2at+b(t^2-T^2).\tag{1.9}$$

Consider now the special case in which P is quadratic in both of its arguments. Temporarily, we will use the shorthand $\partial_\xi=\frac{\mathrm{i}}{2\pi}\frac{\mathrm{d}}{\mathrm{d}\xi}$. Consider the action of $P\big(\frac{\mathrm{i}}{2\pi}\frac{\mathrm{d}}{\mathrm{d}\xi},2\pi\mathrm{i}\xi\big)$ when P is as in Corollary 1.2.2:

$$P(\partial_\xi,2\pi\mathrm{i}\xi)=a(\partial_\xi^2-T^2)\,(2\pi\mathrm{i}\xi)^2+\big(2a\partial_\xi+b(\partial_\xi^2-T^2)\big)(2\pi\mathrm{i}\xi)+c_1\partial_\xi^2+c_2\partial_\xi+c_3.$$

As operators,

$$\begin{aligned}
\partial_\xi^2 (2\pi \mathrm{i}\xi)^2 &= \Big(\frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\Big)\,\xi^2 = \frac{\mathrm{d}}{\mathrm{d}\xi}\Big(2\xi + \xi^2 \frac{\mathrm{d}}{\mathrm{d}\xi}\Big) = 2 + 4\xi\,\frac{\mathrm{d}}{\mathrm{d}\xi} + \xi^2\,\frac{\mathrm{d}^2}{\mathrm{d}\xi^2},\\
\partial_\xi (2\pi \mathrm{i}\xi) &= -\frac{\mathrm{d}}{\mathrm{d}\xi}\,\xi = -1 - \xi\,\frac{\mathrm{d}}{\mathrm{d}\xi}, \quad \text{and}\\
\partial_\xi^2 (2\pi \mathrm{i}\xi) &= -\frac{\mathrm{i}}{2\pi}\Big(\frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\Big)\,\xi = -\frac{\mathrm{i}}{2\pi}\frac{\mathrm{d}}{\mathrm{d}\xi}\Big(1 + \xi\,\frac{\mathrm{d}}{\mathrm{d}\xi}\Big) = -\frac{\mathrm{i}}{2\pi}\Big(2\frac{\mathrm{d}}{\mathrm{d}\xi} + \xi\,\frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\Big).
\end{aligned}$$

Substitution of the corresponding terms in $P(\partial_\xi, 2\pi\mathrm{i}\xi)$ with $c_3 = 0$ gives

$$\begin{aligned}
P(\partial_\xi, 2\pi\mathrm{i}\xi) &= a(\partial_\xi^2 - T^2)\,(2\pi\mathrm{i}\xi)^2 + \big(2a\partial_\xi + b(\partial_\xi^2 - T^2)\big)\,(2\pi\mathrm{i}\xi) + c_1\partial_\xi^2 + c_2\partial_\xi\\
&= a\Big(2 + 4\xi\,\frac{\mathrm{d}}{\mathrm{d}\xi} + \xi^2\,\frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\Big) - aT^2(2\pi\mathrm{i}\xi)^2 - 2a\Big(1 + \xi\,\frac{\mathrm{d}}{\mathrm{d}\xi}\Big)\\
&\quad - \frac{\mathrm{i}b}{2\pi}\Big(2\frac{\mathrm{d}}{\mathrm{d}\xi} + \xi\,\frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\Big) - bT^2(2\pi\mathrm{i}\xi) - \frac{c_1}{(2\pi)^2}\frac{\mathrm{d}^2}{\mathrm{d}\xi^2} + \frac{\mathrm{i}c_2}{2\pi}\frac{\mathrm{d}}{\mathrm{d}\xi}\\
&= \Big(a\xi^2 - \frac{\mathrm{i}b\xi}{2\pi} - \frac{c_1}{4\pi^2}\Big)\frac{\mathrm{d}^2}{\mathrm{d}\xi^2} + \Big(2a\xi + \mathrm{i}\frac{c_2 - 2b}{2\pi}\Big)\frac{\mathrm{d}}{\mathrm{d}\xi} + T^2(4\pi^2 a\xi^2 - 2\pi\mathrm{i}b\xi)\\
&\equiv Q\Big(\xi, \frac{\mathrm{d}}{\mathrm{d}\xi}\Big)
\end{aligned}$$

where $Q\big(\xi, \frac{\mathrm{d}}{\mathrm{d}\xi}\big) = \sigma_0(\xi)\frac{\mathrm{d}^2}{\mathrm{d}\xi^2} + \sigma_1(\xi)\frac{\mathrm{d}}{\mathrm{d}\xi} + \sigma_2(\xi)$ with

$$\begin{aligned}
\sigma_0(\xi) &= a\xi^2 - \frac{\mathrm{i}b}{2\pi}\xi - \frac{c_1}{4\pi^2},\\
\sigma_1(\xi) &= 2a\xi + \mathrm{i}\frac{c_2 - 2b}{2\pi}, \quad \text{and}\\
\sigma_2(\xi) &= 4\pi^2 a\xi^2 T^2 - 2\pi\mathrm{i}b\,\xi T^2.
\end{aligned}$$

Applying the criterion of Corollary 1.2.2, one finds that in order for $Q\big(\xi, \mathrm{d}/\mathrm{d}\xi\big)$ to commute with the characteristic function of an interval $[-\Omega, \Omega]$, it is necessary that $\sigma_0(\xi) = \alpha(\xi^2 - \Omega^2)$ and then that $\sigma_1(\xi) = 2\alpha\xi + \beta(\xi^2 - \Omega)^2$. Since σ_1 has no quadratic term, it follows that $\beta = 0$ and, consequently, that $\alpha = a$ and $c_2 - 2b = 0$. The required form of σ_0 implies that $b = 0$ (so $c_2 = 0$ also) and then that $c_1 = 4\pi^2 a\Omega^2$. Putting these observations together and allowing for an additive constant term in σ_2, we conclude the following.

Theorem 1.2.3. *Suppose that $P(t, \mathrm{d}/\mathrm{d}t)$ is a differential operator of the form $\rho_0\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \rho_1\frac{\mathrm{d}}{\mathrm{d}t} + \rho_2$ with quadratic coefficients ρ_i such that $P(t, \mathrm{d}/\mathrm{d}t)$ commutes with multiplication by the characteristic function of $[-T, T]$ and such that $\mathscr{F}(P(t, \mathrm{d}/\mathrm{d}t)) = P(\partial_\xi, 2\pi\mathrm{i}\xi)$ commutes with multiplication by the characteristic function of $[-\Omega, \Omega]$. Then there exist constants a and b such that*

$$\rho_0(t) = a(t^2 - T^2) \quad \text{and} \quad \rho_1(t) = 2at \quad \text{while} \quad \rho_2(t) = 4\pi^2 a\Omega^2 t^2 + b. \qquad (1.10)$$

Returning to (1.6) and writing

$$\mathscr{P} = \frac{\mathrm{d}}{\mathrm{d}t}(t^2-1)\frac{\mathrm{d}}{\mathrm{d}t} + b^2t^2 = (t^2-1)\frac{\mathrm{d}^2}{\mathrm{d}t^2} + 2t\frac{\mathrm{d}}{\mathrm{d}t} + b^2t^2,$$

one recovers the particular case of (1.10) in which $a=1$, $T=1$, and $b=2\pi\Omega$. We conclude that *any second-order differential operator with quadratic coefficients that commutes with the time- and band-limiting operators is a multiple of a rescaling of the differential operator $\mathscr{P}$ for prolate spheroidal functions of order zero, plus a multiple of the identity*. In particular, PSWFs are eigenfunctions of time- and band-limiting operators.

The PSWF Parameter c and the Time–Bandwidth Product $\mathfrak{a}$

The parameter c in the operator $\mathscr{P}$ in (1.6) is closely related to the time–bandwidth product, and is often used in the literature to index quantities that depend on this product, but c is not equal to this product. The operator $\mathscr{P}$ commutes with both Q and $P_{2\Omega}$ when $c = 2\pi\Omega$. Since the time–bandwidth product of $[-1,1]$ and $[-\Omega,\Omega]$ is 4Ω, this tells us that *c equals the time–bandwidth product multiplied by $\pi/2$*. The area of the product of the time-localization and frequency-localization sets plays such a vital role in what follows that we introduce here the special symbol "$\mathfrak{a}$" in order to reference this quantity. Thus, to the operator $P_\Sigma Q_S$, with P_Σ in (1.2) and Q_S in (1.3), one associates the area $\mathfrak{a}(S,\Sigma) = |S||\Sigma|$. Abusing notation as before, to the operator $P_\Omega Q_T$, where $P_\Omega = P_{[-\Omega/2,\Omega/2]}$ and $Q_T = Q_{[-T,T]}$, we associate $\mathfrak{a} = \mathfrak{a}(T,\Omega) = 2T\Omega$. When ϕ is a unitary dilate of $S_n(t;c)$ in (1.6) that also happens to be an eigenfunction of $P_\Omega Q_T$, one will have $c = \pi\mathfrak{a}/2$. In the Bell Labs papers, Landau, Slepian and Pollak defined the Fourier transform as the integral operator with kernel $\mathrm{e}^{-\mathrm{i}t\xi}$. We will denote this Fourier operator by $\sqrt{2\pi}\mathscr{F}_{2\pi}$. The conversion from $\mathscr{F}$ to $\mathscr{F}_{2\pi}$ can be regarded as a units conversion from hertz to radians per second. The *Bell Labs* inverse Fourier transform has kernel $\mathrm{e}^{\mathrm{i}t\xi}/2\pi$. Also, the Bell Labs papers denoted by $Df(t) = \mathbb{1}_{[-T/2,T/2]}(t)f(t)$ the time-limiting operator of duration T, and by $B = \mathscr{F}_{2\pi}^{-1}\mathbb{1}_{[-\Omega,\Omega]}\mathscr{F}_{2\pi}$ the band-limiting operation of bandwidth 2Ω. The basic Bell Labs time- and band-limiting operator then was BD. Thus, in the Bell Labs notation, with c as in (1.6), $2c = \Omega T$, whereas $2c = \pi\Omega T$ in our notation.

1.2.3 Time–Frequency Properties of PSWFs

In this section we discuss certain analytical properties of eigenfunctions of $P_\Omega Q_T$. Specifically, the eigenvalues are simple, the eigenfunctions are orthogonal in L^2 over $[-T,T]$ as well as over $\mathbb{R}$, the eigenfunctions are complete in $L^2[-T,T]$, and the eigenfunctions are covariant under the Fourier transform, meaning that their Fourier