Xiaoping Shen · Ahmed I. Zayed *Editors*

Multiscale Signal Analysis and Modeling

Springer

Multiscale Signal Analysis and Modeling

Xiaoping Shen • Ahmed I. Zayed
Editors

Multiscale Signal Analysis and Modeling

Springer

Editors
Xiaoping Shen
Department of Mathematics
Ohio University
Ohio University 1
Athens, Ohio
Morton Hall
USA

Ahmed I. Zayed
Department of Mathematical Sciences
DePaul University
North Kenmore Avenue 2320
Chicago, Illinois
Schmitt Academic Center (SAC) 524
USA

ISBN 978-1-4614-4144-1 ISBN 978-1-4614-4145-8 (eBook)
DOI 10.1007/978-1-4614-4145-8
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012945404

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

In recognition of Dr. Gilbert G. Walter having outlived his prime of 79.

From Xiaoping A. Shen and Ahmed I. Zayed

Preface

This monograph is a collection of chapters authored or coauthored by friends and colleagues of Professor Gilbert Walter in celebration of his 80th birthday. The authors represent a spectrum of disciplines, mathematics, applied mathematics, electrical engineering, and statistics; yet, the monograph has one common theme: multiscale analysis.

Multiscale analysis has recently become a topic of increasing interest because of its important applications, in particular, in analyzing complex systems in which the data behave differently depending upon which scale the data are looked at. The advent of wavelets has given an impetus to multiscale analysis, but other techniques such as sampling and subsampling have been used successfully as a tool in analyzing multiscale signals. For this reason we have decided to include a variety of chapters covering different aspects and applications of multiscale analysis.

The monograph is divided into three main parts: Part I is a collection of chapters on sampling theory while Parts II and III contain chapters on multiscale analysis and statistical analysis, respectively. The level of presentation varies. Few chapters are very specialized, while others are self-contained or of expository nature, but most chapters should be accessible to graduate students in mathematics or electrical engineering.

Part I, which consists of eight chapters, has chapters on sampling, interpolation, and approximation in the space of bandlimited functions, shift-invariant and reproducing-kernel Hilbert spaces, and the Hardy space $H^2(\mathbf{D})$ of analytic functions in the open unit disk. Part II contains four chapters on a unified theory for multiscale analysis, multiscale signal processing, developing algorithms for signal and image classification problems in large data sets, and wavelet analysis of ECG (electrocardiogram) signals. Part III is comprised of three chapters on characterization of continuous probability distributions, Bayesian wavelet shrinkage methods, and multiparameter regularization for the construction of estimators in statistical learning theory.

In Chapter 1, W. Madych revisits the classical cardinal series and considers its symmetric partial sums. He derives necessary and sufficient conditions for its convergence under growth conditions on the coefficients that imply analogous

asymptotic behavior of the function represented by the series. Several corollaries, including sampling type theorems, are obtained.

Chapter 2, by F. Stenger, M. Youssef, and J. Niebsch, is also related to the cardinal series and the Sinc function , but in the context of function interpolation and approximation. Function interpolation may be carried out using algebraic polynomials, splines, Fourier polynomials, rational functions, wavelets, or Sinc methods. Function interpolation by one of the above methods frequently has one of the following features: (1) the modulus of the error of approximation near one endpoint of an interval differs considerably from the modulus of the error near the other end-point; (2) the moduli of the errors near the two endpoints of the interval are roughly the same, but differ appreciably from the modulus of the error in the mid-range of the interval. The authors call the former the (E-E) case and the latter the (E-M) case.

In this chapter the authors describe methods for getting a more uniform approximation throughout the interval of approximation in the two aforementioned cases (E-E) and (E-M). They also discuss improving approximation of the derivative obtained by differentiating the constructed interpolation approximations.

In Chapter 3, H. R. Fernández-Morales, A. G. Garcia, and G. Pérez-Villalón consider sampling in a general shift-invariant space $V_2(\phi)$ of $L^2(\mathbb{R}^d)$ with a set Φ of r stable generators and in which the data are samples of some filtered versions of the signal itself taken at a sub-lattice of $\mathbb{Z}^d$. The authors call this problem the problem of generalized sampling in shift-invariant spaces. Assuming that the ℓ^2-norm of the generalized samples of any $f \in V_2(\phi)$ is stable with respect to the $L^2(\mathbb{R}^d)$-norm of the signal f, the authors derive frame expansions in the shift-invariant subspace $V_2(\phi)$ allowing the recovery of signals in this space from the available data.

A similar sampling problem is considered by M. Nashed and Q. Sun in Chapter 4, where they consider a variety of Hilbert and Banach spaces of functions that admit sampling expansions of the form $f(t) = \sum_{n=1}^{\infty} f(t_n) S_n(t)$, where $\{S_n(t)\}_{n=1}^{\infty}$ is a family of functions that depend on the sampling points $\{t_n\}$ but not on the function f. Those function spaces, which arise in connection with sampling expansions, include reproducing-kernel spaces, Sobolev spaces, shift-invariant spaces, translation-invariant spaces and spaces of signals with finite rate of innovation. The authors first discuss the engineering approach to the Shannon sampling theorem which is based on trains of delta functions and then try to provide rigorous justification to the engineering approach using distribution theory and generalized functions. They also discuss sampling in some reproducing-kernel Banach spaces.

Chapter 5 by P. Vaidyanathan and P. Pal gives an overview of the concept of coprime sampling and its applications. Coprime sampling was recently introduced by the authors first for the case of one-dimensional signals and then extended to multidimensional signals. The basic idea is that a continuous-time (or spatial) signal is sampled simultaneously by two sets of samplers, with sampling rates $1/NT$ and $1/MT$ where M and N are coprime integers and $T > 0$. One of the main results is that it is possible to estimate the autocorrelation of the signal at a much higher rate $1/T$ than the total sampling rate. Thus, any application which is based on autocorrelation will benefit from such sampling and reconstruction. An interesting mathematical

problem that comes up when coprime sampling is extended to higher dimensions is how to generate a pair of integer matrices M and N which are commuting and coprime.

Chromatic derivatives and series expansions have recently been introduced as an alternative representation to Taylor series for bandlimited functions and they have been shown to be more useful in practical applications than Taylor series. Chromatic series have similar properties to those of the Whittaker–Shannon–Kotel'nikov sampling series. In Chapter 6, A. Zayed gives an overview of chromatic derivatives and series in one and several variables and then use the Bargmann transform to show that functions, in the Bargmann–Segal–Foch space $\mathfrak{F}$, which is a reproducing-kernel Hilbert Space of entire functions, can be represented by chromatic series expansions. As a result, some properties of the Bargmann–Segal–Foch space can be deduced from those of chromatic series.

In Chapter 7, by D. Alpay, P. Jorgensen, I. Lewkowicz, and I. Marziano, the authors use functional analysis techniques to solve interpolation problems not in the context of bandlimited functions or shift-invariant spaces but in the setting of the Hardy space $H^2(\mathbf{D})$ of functions analytic in the open unit disk $\mathbf{D}$. The space $H^2(\mathbf{D})$ plays an important role in complex analysis, signal processing, and linear dynamical systems. Recently the Cuntz semigroups of C^*-algebras and the Cuntz relations for positive elements in a C^*-algebra have attracted some attention because of their newly discovered connections with applications in signal processing, sub-band filters, and wavelets, which all fall within a larger context of multiscale analysis.

In this work the authors study the Cuntz relations in a different context. They introduce connections between the Cuntz relations and the Hardy space $H^2(\mathbf{D})$ and then use a decomposition of elements in $H^2(\mathbf{D})$ associated with certain isometries which satisfy the Cuntz relations, to solve a new kind of multipoint interpolation problem in $H^2(\mathbf{D})$ where for instance only a linear combination of the values of a function at given points is preassigned, rather than the values at the points themselves.

Chapter 8 by J. Benedetto and S. Datta deals with the autocorrelation of sequences and the construction of constant amplitude zero autocorrelation (CAZAC) sequences x on the integers $\mathbb{Z}$ by means of Hadamard matrices. Recall that a real Hadamard matrix is a square matrix whose entries are either $+1$ or -1 and whose rows are mutually orthogonal. First, the authors explain from a practical point of view why constant amplitude and zero autocorrelation sequences are important. The zero autocorrelation property ensures minimum interference between signals sharing the same channel.

The authors review properties and problems related to Hadamard matrices and then establish the relation between CAZAC sequences on $\mathbb{Z}/N\mathbb{Z}$, Hadamard matrices, and the discrete Fourier transform. They proceed to construct CAZAC sequences on $\mathbb{Z}$ by means of Hadamard matrices and construct unimodular functions on $\mathbb{Z}$ whose autocorrelations are triangles.

Part II, Multiscale Analysis, consists of Chapters 9–12. The authors in Chapter 9 explain that chaos and random fractal theories, which have been used in the analysis of complex data, are fundamentally two different theories. Chaos theory

shows that irregular behaviors in a complex system may be generated by nonlinear deterministic processes, while noise or randomness does not play any role. On the other hand, random fractal theory assumes that the dynamics of the system are inherently random.

Since the two theories are different, different conclusions may be drawn depending upon which theory is used to analyze the data. A great deal of research has been devoted to determining whether a complex time series is generated by a chaotic or a random system. The authors discuss the scale-dependent Lyapunov exponent (SDLE) and use it to develop a unified multiscale analysis theory of complex data.

The goal of Chapter 10, by E. Lin, M. Haske, M. Smith, and D. Sowards, is to determine the optimal wavelet, order, level, and threshold for denoising and compressing an ECG (electrocardiogram) signal while smoothing out and maintaining the integrity of the original signal. The wavelets used are: Daubechies, Biorthogonal Spline, Coiflet, and Symlet. Various thresholds have been utilized, such as soft, hard, global, rigorous SURE, heuristic SURE, universal, and minimax. But, the two kinds of thresholding that are used extensively in this chapter are hard and soft thresholding.

It is well known that the Hermite functions are an orthogonal basis for $L^2(\mathbb{R})$. They are also eigenfunctions of a Sturm–Liouville differential operator, as well as the Fourier transform. D. Mugler and A. Mahadevan in Chapter 11 call these functions the continuous Hermite functions (CHF) to distinguish them from another set of functions that they introduced in a previous work and called the discrete Hermite functions (DHF).

The DHF have the analogous property that they are eigenvectors of a shifted (centered) Fourier matrix and they also form an orthonormal set in a vector space. The authors discuss the notion of Gaussian derivatives and their relationship to the Hermite functions. Because of their relationship with the Gaussian derivatives, the CHF have been used for the multiscale Hermite transform. Multiscale analysis in this chapter refers to the ability to zoom in on features in a signal, moving from a coarse approximation to include details at several different levels. In particular, multiscale analysis provides a decomposition of the input signal into an approximation signal and detail signals at several different levels. The main goal of this chapter is to extend these results to the discrete case. The discrete Hermite transform analysis of an input signal is then compared to the wavelet analysis of the same signal at three different levels.

Local discriminant basis (LDB), which was developed by Saito and Coifman for the purpose of classification, is a tool to extract useful features in signal and image classification problems. It works by decomposing training signals into a time–frequency dictionary. The dictionaries decompose signals into a redundant set of orthogonal subspaces. Each subspace contains basis vectors localized in time and frequency. The goal is, given a dictionary, to find the signal representation within the dictionary that is most useful for classification.

In Chapter 12, B. Marchand and N. Saito propose the use of signatures and earth mover's distance (EMD) to provide data adaptive statistic that is more descriptive

than the distribution of energies and more robust than an epdf (empirical probability density function)-based approach.

The authors first review LDB and EMD and then outline how they can be incorporated into a fast EMD-based LDB algorithm and then compare its performance with different LDB algorithms. They also demonstrate the capabilities of their new algorithm, in comparison with both energy distribution and epdf-based LDB algorithms, by using four different classification problems made of synthetic data sets.

Part III, Statistical Analysis, is comprised of Chapters 13–15. The problem of characterizing a probability distribution is an important problem in various fields. Various systems of distributions have been constructed to provide approximations to a wide variety of distributions. These systems are designed with the requirements of computational ease and feasibility of algebraic manipulation.

In Chapter 13, G. Hamadani focuses on the characterization of the Amoroso distribution, which is a four-parameter, continuous, univariate, unimodel pdf, with semi-infinite range. Many well-known and important distributions are special cases or limiting cases of the Amoroso distribution. The author gives characterizations of the Amoroso distribution in two separate cases based on the truncated moment of a function of first-order statistic and of a function of nth order statistic. He also presents similar characterizations of other distributions, such as SSK, SKS, SK, and SKS-type distributions.

Bayesian paradigm is popular in wavelet data processing because Bayes rules are shrinkers. The Bayes rules can be constructed to mimic the thresholding rules for wavelets, i.e., to slightly shrink the large coefficients and heavily shrink the small coefficients. A paradigmatic task in which wavelets are typically applied is the recovery of an unknown signal f from noisy measurements.

In Chapter 14 by N.Reményi and B. Vidakovic, the authors review some of these concepts and discuss different Bayesian wavelet regression models and methods for wavelet shrinkage. As an illustration of the Bayesian approach, they present BAMS (Bayesian adaptive multiresolution shrinkage) method.

The subject of Chapter 15, the last chapter of the monograph, is the so-called statistical learning theory. One of the central problems in the statistical learning theory is this: given some empirical data $Z = \{(x_i, y_i), i = 1, 2, \ldots n\}$, construct the estimator $f : X \to Y$ that approximates best the relationship between the input x and the output y of a system, i.e., $y \approx f(x)$. The data are seen as the realizations of random variables $(x, y) \in X \times Y$ with a probability density $p(x, y)$. The theory suggests an approach for constructing an estimator that is based on an operator equation for the estimator. The authors, S. Lu, S. Pereverzyev Jr, and S. Sampath, discuss this operator equation and show how it can be treated by the recently developed multiparameter regularization methods, the dual regularized total least squares (DRTLS) and the multi-penalty regularization (MPR).

Finally, the editors would like to express their gratitude to Professor Gilbert Walter for his support and guidance over the years and more importantly for his kindness and friendship. We also wish to thank the Springer-Verlag editors for their support, in particular, Mr. Steven Elliot, who initiated the project, and Merry Stubber

and Dr. Alison Waldron for their help throughout the project. The first editor would like to thank the Battlefield Visualization Branch at the US Air Force Research Lab (AFRL) and the Ohio University Faculty Fellowship Leave program for their generous financial support. Her particular thanks are given to Dr. Jeffery Connor (Ohio University), Dr. Paul Havigs (AFRL), and Ms. Kathryn Farris (AFRL) for their encouragement and support.

USA Xiaoping Shen
USA Ahmed I. Zayed

Contents

Part I Sampling

1 Convergence of Classical Cardinal Series 3
W.R. Madych

2 Improved Approximation via Use of Transformations 25
Frank Stenger, Maha Youssef, and Jenny Niebsch

3 Generalized Sampling in $L^2(\mathbb{R}^d)$ Shift-Invariant Subspaces with Multiple Stable Generators 51
H.R. Fernández-Morales, A. G. García, and G. Pérez-Villalón

4 Function Spaces for Sampling Expansions 81
M. Zuhair Nashed and Qiyu Sun

5 Coprime Sampling and Arrays in One and Multiple Dimensions 105
P.P. Vaidyanathan and Piya Pal

6 Chromatic Expansions and the Bargmann Transform 139
Ahmed I. Zayed

7 Representation Formulas for Hardy Space Functions Through the Cuntz Relations and New Interpolation Problems 161
Daniel Alpay, Palle Jorgensen, Izchak Lewkowicz, and Itzik Marziano

8 Constructions and a Generalization of Perfect Autocorrelation Sequences on $\mathbb{Z}$ 183
John J. Benedetto and Somantika Datta

Part II Multiscale Analysis

9 On the Application of the SDLE to the Analysis of Complex Time Series 211
Jianbo Gao, Jing Hu, and Wen-wen Tung

10 Wavelet Analysis of ECG Signals 233
En-Bing Lin, Megan Haske, Marilyn Smith, and Darren Sowards

11 Multiscale Signal Processing with Discrete Hermite Functions 257
Dale H. Mugler and Anandi Mahadevan

12 Earth Mover's Distance-Based Local Discriminant Basis 275
Bradley Marchand and Naoki Saito

Part III Statistical Analysis

13 Characterizations of Certain Continuous Distributions 297
G.G. Hamedani

14 Bayesian Wavelet Shrinkage Strategies: A Review 317
Norbert Reményi and Brani Vidakovic

15 Multiparameter Regularization for Construction of Extrapolating Estimators in Statistical Learning Theory 347
Shuai Lu, Sergiy Pereverzyev Jr., and Sivananthan Sampath

Index 367

Contributors

Daniel Alpay Department of Mathematics, Ben Gurion University of the Negev, Be'er Sheva, Israel

John J. Benedetto Department of Mathematics, Norbert Wiener Center, University of Maryland, College Park, MD, USA

Somantika Datta Department of Mathematics, University of Idaho, Moscow, ID, USA

Héctor R. Fernández-Morales Departamento de Matemáticas, Universidad Carlos III de Madrid, Madrid, Spain

Jianbo Gao PMB Intelligence LLC, West Lafayette, IN, USA

Antonio G. García Departamento de Matemáticas, Universidad Carlos III de Madrid, Madrid, Spain

G.G. Hamedani Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, WI, USA

Megan Haske Department of Mathematics, Central Michigan University, Mt. Pleasant, MI, USA

W. R. Madych Department of Mathematics, University of Connecticut, Storrs, CT, USA

Anandi Mahadevan Department of Biomedical Engineering, University of Akron, Akron, OH 44325, USA

Philips Home Healthcare Solutions, Murraysville, PA 15668, USA

Dale H. Mugler Departments of Biomedical Engineering and Applied Mathematics, University of Akron, Akron, OH, 44325, USA

Jing Hu Affymetrix, Inc., Santa Clara, CA 95051, USA

Palle Jorgensen Department of Mathematics, The University of Iowa, Iowa City, IA, USA

Izchak Lewkowicz Department of Electrical Engineering, Ben Gurion University of the Negev, Be'er, Sheva, Israel

En-Bing Lin Department of Mathematics, Central Michigan University, Mt. Pleasant, MI, USA

Shuai Lu School of Mathematical Science, Fudan University, Shanghai, P. R. China

Bradley Marchand Panama City Division, Naval Surface Warfare Center, FL 32407, USA

Itzik Marziano Department of Mathematics, Ben Gurion University of the Negev, Be'er Sheva, Israel

M. Zuhair Nashed Department of Mathematics, University of Central Florida, Orlando, FL, USA

Jenny Niebsch Radon Institute (RICAM), Austrian Academy of Science, Linz, Austria

Piya Pal Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

Sergiy Pereverzyev Jr Industrial Mathematics Institute, Johannes Kepler University of Linz, Linz, Austria

Gerardo Pérez-Villalón Departamento de Matemática Aplicada, EUITT, UPM, Madrid, Spain

Norbert Remènyi School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA

Naoki Saito Department of Mathematics, University of California-Davis, Davis, CA, USA

Sivananthan Sampath Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenberger Strasse 69, 4040 Linz, Austria

Marilyn Smith Department of Mathematics, Central Michigan University, Mt. Pleasant, MI, USA

Darren Sowards Department of Mathematics, Central Michigan University, Mt. Pleasant, MI, USA

Frank Stenger Manager, SINC, LLC, School of Computing, University of Utah, Salt Lake City, UT, USA

Qiyu Sun Department of Mathematics, University of Central Florida, Orlando, FL, USA

Wen-wen Tung Department of Earth & Atmospheric Sciences, Purdue University, West Lafayette, IN, USA

P. P. Vaidyanathan Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

Brani Vidakovic School of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, USA

Maha Youssef Mathematics Department, German University in Cairo, New Cairo City, Egypt

Ahmed I. Zayed Department of Mathematical Sciences, DePaul University, Schmitt Academic Center (SAC) 524 2320 North Kenmore Avenue, Chicago, IL 60614, USA

Part I
Sampling

Chapter 1
Convergence of Classical Cardinal Series

W.R. Madych

Abstract We consider symmetric partial sums of the classical cardinal series and record necessary and sufficient conditions for convergence. Included are growth conditions on the coefficients that imply analogous asymptotic behavior of the function represented by the series. Several relatively immediate corollaries are also recorded, including sampling-type theorems.

Mathematics subject classification (2000): 40A30; 94A20

1.1 Introduction

The classical cardinal series with coefficients $\{a(n) : n = 0, \pm 1, \pm 2, \ldots\}$ is defined by

$$f(z) = \sum_{n=-\infty}^{\infty} a(n) \frac{\sin \pi(z-n)}{\pi(z-n)}, \tag{1.1}$$

where the variable z is often restricted to the real line but, in general, can take on complex values. The coefficients of course are in general complex.

Under suitable restrictions on the coefficients $\{a(n) : n = 0, \pm 1, \pm 2, \ldots\}$ the series (1.1) provides a solution to the interpolation problem of finding an entire function $f(z)$ of exponential type no greater than π that satisfies

$$f(n) = a(n), \qquad n = 0, \pm 1, \pm 2, \ldots . \tag{1.2}$$

W.R. Madych (✉)
Department of Mathematics, 196 Auditorium Road, University of Connecticut,
Storrs, CT 06269-3009, USA
e-mail: madych@math.uconn.edu

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_1,

The cardinal series (1.1) is a well-known and highly celebrated solution to the interpolation problem (1.2). Indeed the list [2–8, 10, 11, 15–17, 19, 20] is but a small sampling of the many articles and books that are devoted to or significantly treat the subject. We assume that the reader is familiar with what are now relatively widely well-known facts concerning the cardinal series (1.1) that are associated with the theory commonly referred to as the W-K-S sampling theorem and that can be found, for example, in [11, Lecture 20] or [16, Chap. 9].

In this note we are concerned with the convergence of the symmetric partial sums of (1.1), more specifically, with conditions on the coefficients $\{a(n)\}$ that insure the convergence of the sequence $\{f_N(z) : N = 1,2,\ldots\}$ of symmetric partial sums

$$f_N(z) = \sum_{n=-N}^{N} a(n)\frac{\sin\pi(z-n)}{\pi(z-n)}. \tag{1.3}$$

We use standard notation and only alert the reader to the fact that E_π denotes the class of entire functions of exponential type no greater than π that have no greater than polynomial growth along the real axis. In view of the distributional variant of the Paley–Wiener theorem, for example, see [9, Theorem 1.7.7], E_π consists of the Fourier transforms of distributions with support in the interval $[-\pi,\pi]$.

The main results, including some explanatory material, are given in Sect. 1.2. All the details, including necessary technical lemmas, are given in Sect. 1.3. Section 1.4 is devoted to certain miscellany that is a relatively immediate consequence of the development in Sects. 1.2 and 1.3; Corollary 6 here is an example of a sampling-type theorem mentioned in the introduction.

1.2 Results

We make use of the fact that the partial sums $f_N(z)$ defined by (1.3) can be re-expressed as

$$f_N(z) = \frac{\sin\pi z}{\pi}\sum_{n=-N}^{N}\frac{(-1)^n a(n)}{(z-n)}. \tag{1.4}$$

It follows from (1.4) that when the sequence of coefficients $\{a(n)\}$ is even, namely $a(-n) = a(n)$ for $n = 1,2,\ldots$, we may write

$$f_N(z) = \frac{\sin\pi z}{\pi}\left\{\frac{a(0)}{z} + 2z\sum_{n=1}^{N}\frac{(-1)^n a(n)}{(z^2-n^2)}\right\}. \tag{1.5}$$

From (1.5) it is clear that when $\{a(n)\}$ is an even sequence then the convergence of $\sum(-1)^n a(n)/n^2$ is a sufficient condition for the convergence of the partial sums $\{f_N(z)\}$. This condition is also necessary. Furthermore, the limiting function $f(z)$

is a solution within a certain class of entire functions to the interpolation problem $f(n) = a(n)$, $n = 0, \pm 1, \pm 2, \ldots$. We formulate this more precisely as follows.

Theorem 1. *Suppose the sequence of coefficients* $\{a(n)\}$ *is even, namely* $a(-n) = a(n)$ *for* $n = 1, 2, \ldots$.

1. *If*

$$\sum (-1)^n \frac{a(n)}{n^2} \quad \text{converges}, \tag{1.6}$$

then the partial sums $f_N(z)$, $N = 1, 2, \ldots$, *converge uniformly on compact subsets of the complex plane* $\mathbb{C}$. *The limiting function*

$$f(z) = \frac{\sin \pi z}{\pi} \left\{ \frac{a(0)}{z} + 2z \sum_{n=1}^{\infty} \frac{(-1)^n a(n)}{(z^2 - n^2)} \right\} \tag{1.7}$$

is even, is in E_π, *satisfies*

$$|f(z)| \mathrm{e}^{-\pi |\operatorname{Im} z|} = O(|z|^2 \log |z|) \quad \text{as} \quad |z| \to \infty, \tag{1.8}$$

and solves the interpolation problem (1.2).

2. *If* (1.6) *fails to hold then the sequence* $f_N(z)$ *fails to converge as* $N \to \infty$ *at every point* z *that is not an integer.*

The statement concerning convergence of the partial sums $f_N(z)$ can be reformulated as follows: There is an entire function $f(z)$ such that for every positive number R,

$$\lim_{N \to \infty} \sup_{|z| \le R} \left| f(z) - \sum_{n=-N}^{N} a(n) \frac{\sin \pi (z-n)}{\pi (z-n)} \right|.$$

In view of the function

$$z \sin \pi z,$$

any solution of the interpolation problem (1.2) that is even, is in E_π, and enjoys (1.8) cannot be unique. Additional restrictions on the coefficients $\{a(n)\}$ are required to ensure that the solution given by (1.7) is unique within an appropriate class of entire functions in analogy with the celebrated sampling theorem, for example, [11, Lecture 20, Theorem 1].

Theorem 2. *Suppose the sequence of coefficients* $\{a(n)\}$ *is even and satisfies property* (1.6). *If, in addition, for some* p *that satisfies* $0 \le p \le 2$ *we have*

$$a(n) = O(n^p) \quad \text{as} \quad n \to \infty, \tag{1.9}$$

then the limiting function $f(z)$ *defined by* (1.7) *satisfies*

$$|f(z)| \mathrm{e}^{-\pi |\operatorname{Im} z|} = O(|z|^p \log |z|) \quad \text{as} \quad |z| \to \infty. \tag{1.10}$$

If $0 \leq p < 1$ *then the limiting function* $f(z)$ *defined by* (1.7) *is the unique solution to the interpolation problem* (1.2) *that is in* E_π*, is even, and satisfies* (1.10)*.*

Note that condition (1.9) on the growth of the coefficients $\{a(n)\}$ does not imply that the solution (1.7) of the interpolation problem (1.2) has the same order of growth. However, an additional restriction, on what amounts to the oscillatory nature of the coefficients, will ensure that the solution (1.7) has the same order of growth as the coefficients (1.9).

Theorem 3. *Suppose the sequence of coefficients* $\{a(n)\}$ *is even and satisfies property* (1.6)*. If, in addition, for some* p *that satisfies* $0 \leq p \leq 2$ *we have*

$$a(n+1) - a(n) = O(n^{p-1}) \quad as \quad n \to \infty \quad when \quad 0 < p \leq 2 \tag{1.11}$$

and

$$\sum_{n=1}^{\infty} |a(n+1) - a(n)| < \infty \quad when \quad p = 0, \tag{1.12}$$

then the limiting function $f(z)$ *defined by* (1.7) *satisfies*

$$|f(z)|\mathrm{e}^{-\pi|\operatorname{Im} z|} = O(|z|^p) \quad as \quad |z| \to \infty. \tag{1.13}$$

When the sequence of coefficients $\{a(n)\}$ is odd, namely $a(-n) = -a(n)$ for $n = 1, 2, \ldots$, in view of (1.4) we have

$$f_N(z) = \frac{2\sin \pi z}{\pi} \sum_{n=1}^{N} \frac{(-1)^n n a(n)}{(z^2 - n^2)}. \tag{1.14}$$

From (1.14) it should be clear that the conditions required of $\{a(n)\}$ in this case will be somewhat more restrictive than in the even case. Nevertheless, with relatively minor modifications, the analogues of Theorems 1–3 remain valid and can be formulated as follows.

Theorem 4. *Suppose the sequence of coefficients* $\{a(n)\}$ *is odd, namely* $a(-n) = -a(n)$ *for* $n = 1, 2, \ldots$.

1. If

$$\sum (-1)^n \frac{a(n)}{n} \quad converges, \tag{1.15}$$

then the partial sums $f_N(z)$, $N = 1, 2, \ldots$, *converge uniformly on compact subsets of the complex plane* $\mathbb{C}$*. The limiting function*

$$f(z) = \frac{2\sin \pi z}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n n a(n)}{(z^2 - n^2)} \tag{1.16}$$

is odd, is in E_π*, satisfies*

$$|f(z)|\mathrm{e}^{-\pi|\operatorname{Im} z|} = O(|z|\log|z|) \quad as \quad |z| \to \infty, \tag{1.17}$$

and solves the interpolation problem (1.2).

2. *If* (1.15) *fails to hold, then the sequence* $f_N(z)$ *fails to converge as* $N \to \infty$ *at every point* z *that is not an integer.*

In view of the function

$$\sin \pi z,$$

any solution of the interpolation problem (1.2) that is odd, is in E_π, and enjoys (1.17) cannot be unique. In this case I am unaware of any conditions on the coefficients other than the decay conditions implied by the sampling-type theorems, for example, [11, Lecture 20, Theorems 1 and 2] or Corollary 6 in Sect. 1.4, that will ensure that the solution (1.16) is unique within some appropriate class of entire functions. The statements in Theorem 4 concerning convergence are also implied by [18, Theorem 1].

Theorem 5. *Suppose the sequence of coefficients* $\{a(n)\}$ *is odd and satisfies property* (1.15). *If, in addition, for some* p *that satisfies* $0 \le p \le 1$ *we have*

$$a(n) = O(n^p) \quad as \quad n \to \infty, \tag{1.18}$$

then the limiting function $f(z)$ *defined by* (1.16) *satisfies*

$$|f(z)|\mathrm{e}^{-\pi|\operatorname{Im} z|} = O(|z|^p \log|z|) \quad as \quad |z| \to \infty. \tag{1.19}$$

Note that condition (1.18) on the growth of the coefficients $\{a(n)\}$ does not imply that the solution (1.16) of the interpolation problem (1.2) has the same order of growth. However, as in the earlier case, an additional restriction, on what amounts to the oscillatory nature of the coefficients, will ensure that the solution (1.16) has the same order of growth as the coefficients (1.18).

Theorem 6. *Suppose the sequence of coefficients* $\{a(n)\}$ *is odd and satisfies property* (1.15). *If, in addition, for some* p *that satisfies* $0 \le p \le 1$ *we have*

$$a(n+1) - a(n) = O(n^{p-1}) \quad as \quad n \to \infty \quad when \quad 0 < p \le 1 \tag{1.20}$$

and

$$\sum_{n=1}^{\infty} |a(n+1) - a(n)| < \infty \quad when \quad p = 0, \tag{1.21}$$

then the limiting function $f(z)$ *defined by* (1.16) *satisfies*

$$|f(z)|\mathrm{e}^{-\pi|\operatorname{Im} z|} = O(|z|^p) \quad as \quad |z| \to \infty. \tag{1.22}$$

1.3 Details

In what follows the symbol C, with or without a subscript, is used to denote certain generic constants whose specific value can vary from one occurrence to another.

1.3.1 Proof of Theorem 1

In view of (1.5) we may re-express $f_N(z)$ as

$$\begin{aligned} f_N(z) &= \frac{\sin \pi z}{\pi}\left\{\frac{a(0)}{z} + 2z\sum_{n=1}^{N}\left[\left(\frac{1}{(z^2-n^2)}+\frac{1}{n^2}\right)(-1)^n a(n) - (-1)^n\frac{a(n)}{n^2}\right]\right\} \\ &= \frac{\sin \pi z}{\pi}\left\{\frac{a(0)}{z} + 2z\sum_{n=1}^{N}\left(\frac{(-1)^n z^2}{n^2(z^2-n^2)}\right)a(n) - 2z\sum_{n=1}^{N}(-1)^n\frac{a(n)}{n^2}\right\} \\ &= \phi_N(z) + \psi_N(z), \end{aligned}$$

where

$$\phi_N(z) = \frac{\sin \pi z}{\pi}\left\{\frac{a(0)}{z} + 2z\sum_{n=1}^{N}\left(\frac{(-1)^n z^2}{n^2(z^2-n^2)}\right)a(n)\right\}$$

and

$$\psi_N(z) = \frac{-2z\sin \pi z}{\pi}\sum_{n=1}^{N}(-1)^n\frac{a(n)}{n^2}.$$

In view of (1.6) the sequence $\psi_N(z)$ converges uniformly on compacta as $N \to \infty$.

Condition (1.6) also implies that $\lim_{n\to\infty} a(n)/n^2 = 0$ and hence $\sum_{n=1}^{\infty} a(n)/n^4$ converges absolutely. It follows that the series

$$\sum_{n=1}^{\infty}\left(\frac{2z^3\sin \pi z}{\pi(z^2-n^2)}\right)(-1)^n\frac{a(n)}{n^2}$$

converges absolutely and uniformly on compacta. This means, of course, that $\phi_N(z)$ converges uniformly on compacta as $N \to \infty$.

From the last expression for $f_N(z)$ it follows that $f_N(z)$ converges uniformly on compacta as $N \to \infty$ since both $\phi_N(z)$ and $\psi_N(z)$ do so.

We may express the limiting function $f(z)$ as

$$f(z) = \frac{\sin \pi z}{\pi}\left\{\frac{a(0)}{z} + 2z\sum_{n=1}^{\infty}\frac{(-1)^n a(n)}{(z^2-n^2)}\right\} \tag{1.7}$$

since the series in fact converges to the entire function $f(z)$. In view of the above development we may also express $f(z)$ as

$$f(z) = \phi(z) + cz\sin \pi z, \tag{1.23}$$

where

$$\phi(z) = \frac{\sin \pi z}{\pi}\left\{\frac{a(0)}{z} + 2z\sum_{n=1}^{\infty}\left(\frac{(-1)^n z^2}{n^2(z^2-n^2)}\right)a(n)\right\}$$

and

$$c = \frac{-2}{\pi}\sum_{n=1}^{N}(-1)^n\frac{a(n)}{n^2}.$$

An efficient way of arguing that $f(z)$ is in E_π is to observe that this is an immediate consequence of (1.8).

To see that $f(z)$ satisfies (1.8) use representation (1.7) of $f(z)$, assume $|z| \geq 100$, and break up the series into a sum over $n \geq 2|z|$ and another over $n < 2|z|$. Thus

$$f(z) = A_N(z) + B_N(z),$$

where N is the greatest integer $\leq 2|z|$,

$$A_N(z) = \frac{\sin \pi z}{\pi}\left\{\frac{a(0)}{z} + 2z\sum_{n=1}^{N}\frac{(-1)^n a(n)}{z^2-n^2}\right\},$$

and

$$B_N(z) = \frac{\sin \pi z}{\pi}\left\{2z\sum_{n=N+1}^{\infty}\frac{(-1)^n a(n)}{z^2-n^2}\right\}.$$

To estimate $A_N(z)$ assume that Im z is positive so that $|e^{i2\pi z} - 1| \leq 2$ and note that

$$\left|\frac{\sin \pi z}{z \pm n}\right| = |e^{-i\pi z}|\left|\frac{e^{2\pi i z}-1}{z\pm n}\right| = e^{\pi|\operatorname{Im} z|}\left|\frac{e^{2\pi i(z\pm n)}-1}{z\pm n}\right|$$

and

$$\left|\frac{e^{2\pi i(z\pm n)}-1}{z\pm n}\right| \leq \frac{C}{1+\big||n|-|z|\big|}.$$

Hence,

$$\frac{2z\sin \pi z}{z^2-n^2} = \left\{\frac{1}{z-n}+\frac{1}{z+n}\right\}\sin \pi z \leq \frac{Ce^{\pi|\operatorname{Im} z|}}{1+\big||n|-|z|\big|},$$

so that

$$|A_N(z)| \le Ce^{\pi|\operatorname{Im} z|} \sum_{n=0}^{N} \frac{|a(n)|}{1+|n-|z||}.$$

Now,

$$|a(n)| \le C|z|^2$$

in the above sum, since for $n \ge 1$ (1.6) implies that $|a(n)|/n^2$ is bounded which in turn implies $|a(n)| \le Cn^2 \le CN^2$. The last two displayed inequalities imply that

$$|A_N(z)| \le C|z|^2 e^{\pi|\operatorname{Im} z|} \sum_{n=0}^{N} \frac{1}{1+|n-|z||} \le C|z|^2 e^{\pi|\operatorname{Im} z|} \log|z|. \tag{1.24}$$

An analogous argument *mutatis mutandis* shows that (1.24) is still valid when $\operatorname{Im} z \le 0$ so that (1.24) holds whenever $|z|$ is sufficiently large.

To estimate $B_N(z)$ break it up into two terms analogous to $\phi_N(z)$ and $\psi_N(z)$ above. Namely, write

$$B_N(z) = \frac{\sin \pi z}{\pi}\left\{2\sum_{n=N+1}^{\infty}\left(\frac{(-1)^n z^3}{n^2(z^2-n^2)}\right)a(n) - 2z\sum_{n=N+1}^{\infty}(-1)^n\frac{a(n)}{n^2}\right\}$$

and note that

$$\left|\sum_{n=N+1}^{\infty}\left(\frac{(-1)^n z^3}{n^2(z^2-n^2)}\right)a(n)\right| \le |z|^3\sum_{n=N+1}^{\infty}\frac{C_1|a(n)|}{n^4}$$

$$\le |z|^3\sum_{n=N+1}^{\infty}\frac{C_2}{n^2} \le C_3|z|^3N^{-1} \le C|z|^2$$

and

$$\left|z\sum_{n=N+1}^{\infty}(-1)^n\frac{a(n)}{n^2}\right| \le C|z|.$$

The last expression for $B_N(z)$ together with the last two inequalities implies that

$$|B_N(z)| \le C|z|^2 e^{\pi|\operatorname{Im} z|}. \tag{1.25}$$

The desired result (1.8) follows from (1.24) and (1.25).

Now, suppose that the series in (1.6) diverges. The proof of item 2 can be reduced to two simple cases. (a) If the terms of the series in (1.6) are unbounded, then so are the terms of the series (1.7), and desired result follows. (b) If the terms of the series in (1.6) are bounded, then using representation (1.23) for $f(z)$, note that the series representing $\phi(z)$ converges while the series representing the constant c diverges, and the desired result follows.

1.3.2 Proof of Theorem 2

The proof of (1.10) is essentially analogous to the proof of (1.8) while making use of the additional restrictions on the coefficients $\{a(n)\}$. The only significant modification involves the estimation of $B_N(z)$ which requires the consideration of two cases depending on whether p is less than or ≥ 1.

Thus, use representation (1.7) of $f(z)$, assume $|z| \geq 100$, break up the series into a sum over $n \geq 2|z|$ and another over $n < 2|z|$, and write

$$f(z) = A_N(z) + B_N(z),$$

where N is the greatest integer $\leq 2|z|$ and both $A_N(z)$ and $B_N(z)$ are defined exactly the same as in the proof of (1.8). Then estimating $A_N(z)$ as before but using the fact that in this case $|a(n)| \leq C|z|^p$ results in

$$|A_N(z)| \leq C|z|^p e^{\pi|\operatorname{Im} z|} \log|z|. \tag{1.26}$$

As mentioned earlier the estimation of $B_N(z)$ depends on whether p is less than or ≥ 1.

If $0 \leq p < 1$, simply recall that

$$B_N(z) = \frac{\sin \pi z}{\pi} \left\{ 2z \sum_{n=N+1}^{\infty} \frac{(-1)^n a(n)}{z^2 - n^2} \right\}$$

and observe that

$$\left| z \sum_{n \geq 2|z|} \frac{(-1)^n a(n)}{(z^2 - n^2)} \right| \leq |z| \sum_{n \geq 2|z|} \frac{4|a(n)|}{3n^2} \leq C_1 |z| \sum_{n \geq 2|z|} n^{p-2} \leq C_2 |z|\, |z|^{p-1}.$$

to conclude that

$$|B_N(z)| \leq C|z|^p e^{\pi|\operatorname{Im} z|}. \tag{1.27}$$

If $1 \leq p \leq 2$ estimate exactly as in the derivation of (1.8) but use the bound $|a(n)| \leq Cn^p$. This leads to (1.27) for this case.

Bounds (1.26) and (1.27) together imply the desired result (1.10).

To see the uniqueness statement we argue as follows: If $g(z)$ is another solution of the interpolation problem (1.2), is in E_π, is even, and satisfies (1.10) for some $p < 1$, then $h(z) = \big(f(z) - g(z)\big) / \sin \pi z$ is an entire function that is $o(|z|)$ as $|z| \to \infty$. Hence Cauchy's estimate, [1, p 122, identity (25) with $n = 1$] implies that $h(z)$ is a constant. In view of the fact that $h(z)$ is odd this constant must be zero. Thus $g(z) = f(z)$, which implies the desired result.

1.3.3 A Technical Lemma

Let

$$S_n(z) = \frac{\sin \pi z}{\pi} \sum_{k=-n}^{n} \frac{(-1)^k}{(z-k)}. \tag{1.28}$$

Then in view of the uniqueness statement in Theorem 2, it follows that

$$\lim_{n\to\infty} S_n(z) = 1$$

uniformly on compacta. Our proof of Theorem 3 uses the fact that $S_n(z)$ is uniformly bounded in n on strips parallel to the real axis, $\{z : |\operatorname{Im} z| < R < \infty\}$, which however does not follow from Theorem 2 and requires an additional tweak.

Lemma. *There is a positive constant C, independent of z and n, such that*

$$|S_n(z)| \le C e^{\pi |\operatorname{Im} z|}. \tag{1.29}$$

To see the lemma note that for positive k

$$\frac{(-1)^{2k-1}}{z-(2k-1)} + \frac{(-1)^{2k}}{z-2k} = \frac{1}{(z-2k+1)(z-2k)}$$

and that

$$\left| \frac{\sin \pi z}{(z-2k+1)(z-2k)} \right| \le \frac{C e^{\pi |\operatorname{Im} z|}}{1+|z-2k|^2}$$

with a similar estimate valid for negative k. Hence

$$|S_{2k}(z) - S_{2(k-1)}(z)| \le C e^{\pi |\operatorname{Im} z|} \left\{ \frac{1}{1+|z-2k|^2} + \frac{1}{1+|z+2k|^2} \right\}.$$

If n is even, $n = 2m$, then

$$S_{2m}(z) = S_0(z) + \sum_{k=1}^{m} \{S_{2k}(z) - S_{2(k-1)}(z)\},$$

and if n is odd, $n = 2m+1$, then

$$S_{2m+1}(z) = S_{2m}(z) - \frac{2z \sin \pi z}{z^2 - (2m+1)^2}.$$

Hence

$$|S_{2m}(z)| = |S_0(z)| + \sum_{k=1}^{m} |S_{2k}(z) - S_{2(k-1)}(z)|$$
$$\leq Ce^{\pi|\operatorname{Im} z|} \left\{ \frac{1}{1+|z|} + \sum_{k=1}^{m} \left\{ \frac{1}{1+|z-2k|^2} + \frac{1}{1+|z+2k|^2} \right\} \right\}$$

which implies that (1.29) is valid when $n = 2m$ and since

$$\left| \frac{2z \sin \pi z}{z^2 - (2m+1)^2} \right| \leq Ce^{\pi|\operatorname{Im} z|}$$

inequality (1.29) follows for all n.

1.3.4 *Proof of Theorem 3*

As in the proof of Theorem 2, use representation of $f(z)$, assume $|z| \geq 100$, and break up the series into a sum over $n \geq 2|z|$ and another over $n < 2|z|$, and write

$$f(z) = A_N(z) + B_N(z),$$

where N is the greatest integer $\leq 2|z|$ and both $A_N(z)$ and $B_N(z)$ are defined exactly the same as before. Also note that the hypothesis on the coefficients $\{a(n)\}$ implies that $a(n) = O(n^p)$ as $n \to \infty$.

$B_N(z)$ can be estimated in exactly the same way as in the proof of Theorem 2 to get

$$|B_N(z)| \leq C|z|^p e^{\pi|\operatorname{Im} z|}.$$

To estimate $A_N(z)$ use summation by parts to write

$$A_N(z) = \sum_{n=0}^{N-1} S_n(z) \big(a(n) - a(n+1)\big) + S_N(z)\, a(N),$$

where

$$S_0(z) = \frac{\sin \pi z}{\pi z}$$

and

$$S_n(z) = S_0(z) + \frac{\sin \pi z}{\pi} \left\{ 2z \sum_{k=1}^{n} \frac{(-1)^k}{z^2 - k^2} \right\}, \quad n = 1, 2, \ldots .$$

In view of inequality (1.29) the last expression for $A_N(z)$ allows us to write

$$|A_N(z)| \leq Ce^{\pi|\operatorname{Im} z|}\left\{\sum_{n=0}^{N-1}|a(n)-a(n+1)|+|a(N)|\right\}$$

which together with the hypothesis on the coefficients $\{a(n)\}$ implies that

$$|A_N(z)| \leq C|z|^p e^{\pi|\operatorname{Im} z|}.$$

The bounds on $A_N(z)$ and $B_N(z)$ imply the desired result (1.22).

1.3.5 Proof of Theorems 4 and 5

The proofs of Theorems 4 and 5 are essentially the same as those of Theorems 1 and 2, *mutatis mutandis*.

The necessary modifications are evident by reexpressing (1.16) as

$$f(z)=\frac{2\sin\pi z}{\pi}\left\{\sum_{n=1}^{\infty}\left[\left(\frac{n}{(z^2-n^2)}+\frac{1}{n}\right)(-1)^n a(n)-(-1)^n\frac{a(n)}{n}\right]\right\}$$

which, in analogy with (1.23), can be written as

$$f(z)=\phi(z)+c\sin\pi z, \tag{1.30}$$

where

$$\phi(z)=\frac{2\sin\pi z}{\pi}\left\{\sum_{n=1}^{\infty}\left(\frac{(-1)^n z^2}{n(z^2-n^2)}\right)a(n)\right\}$$

and

$$c=\frac{-2}{\pi}\sum_{n=1}^{N}(-1)^n\frac{a(n)}{n}.$$

Also recall that

$$\left|\frac{2n\sin\pi z}{z^2-n^2}\right|=\left|\left\{\frac{1}{z-n}-\frac{1}{z+n}\right\}\sin\pi z\right|\leq\frac{Ce^{\pi|\operatorname{Im} z|}}{1+\big||n|-|z|\big|}.$$

1.3.6 Another Technical Lemma

In analogy with (1.28) let

$$\operatorname{Sgn}_n(z)=\frac{\sin\pi z}{\pi}\sum_{k=-n}^{n}\frac{(-1)^k\operatorname{sgn}(k)}{(z-k)}, \tag{1.31}$$

where

$$\operatorname{sgn}(z) = \begin{cases} z/|z| & \text{if } z \neq 0, \\ 0 & \text{when } z = 0. \end{cases}$$

Then in view of Theorem 5 it follows that there is an entire function $\operatorname{Sgn}(z)$ in E_π such that

$$\lim_{n\to\infty} \operatorname{Sgn}_n(z) = \operatorname{Sgn}(z)$$

uniformly on compacta. Our proof of Theorem 6 is analogous to that of Theorem 3 and uses the fact that $\operatorname{Sgn}_n(z)$ is uniformly bounded in n on strips parallel to the real axis, $\{z : |\operatorname{Im} z| < R < \infty\}$, which does not follow from Theorem 5. The proof of the following lemma is completely analogous to the proof of (1.29).

Lemma. *There is a positive constant C, independent of z and n, such that*

$$|\operatorname{Sgn}_n(z)| \le C e^{\pi|\operatorname{Im} z|}. \tag{1.32}$$

1.3.7 Proof of Theorem 6

Our proof of Theorem 6 is completely analogous to that of Theorem 3. Simply replace $S_n(z)$ with $\operatorname{Sgn}_n(z)$ and use (1.32) instead of (1.29).

1.4 Additional Remarks, Examples, and Corollaries

1.4.1 Specific Bounds

It should be evident from the above development that more specific bounds on the growth of the coefficients $\{a(n)\}$ will lead, via essentially the same calculations, to more specific bounds on the growth of the corresponding function (1.7) or (1.16).

For example, if in Theorem 2 we assume that $0 \le p \le 1$ and

$$\|\{a(n)\}\|_p = \sup_n \frac{|a(n)|}{(1+|n|)^p} < \infty,$$

then we may conclude that

$$|f(z)| \le C\|\{a(n)\}\|_p \, e^{\pi|\operatorname{Im} z|}(1+|z|)^p \log(e+|z|),$$

where C is a constant that may depend on p but is otherwise independent of $\{a(n)\}$. Similar results are valid in all the other cases.

1.4.2 Some Special Functions

If

$$a(n) = (-1)^n = \cos \pi n,$$

then the statement in Theorem 2 concerning uniqueness can be applied to conclude that

$$\cos \pi z = \frac{\sin \pi z}{\pi} \left\{ \frac{1}{z} + 2z \sum_{k=1}^{\infty} \frac{1}{(z^2 - k^2)} \right\}. \tag{1.33}$$

But the fact that the right-hand side of (1.33) is bounded when z is restricted to a strip about the real axis, $|\operatorname{Im} z| \leq \varepsilon < \infty$ does not follow from Theorem 3 since $|a(n+1) - a(n)| = 2$. On the other hand, unlike the partial sums of

$$1 = \frac{\sin \pi z}{\pi} \left\{ \frac{1}{z} + 2z \sum_{k=1}^{\infty} \frac{(-1)^k}{(z^2 - k^2)} \right\} \tag{1.34}$$

and

$$\operatorname{Sgn}(z) = \frac{2 \sin \pi z}{\pi} \left\{ \sum_{k=1}^{\infty} \frac{(-1)^k k}{(z^2 - k^2)} \right\}, \tag{1.35}$$

the partial sums

$$\cos_n \pi z = \frac{\sin \pi z}{\pi} \left\{ \frac{1}{z} + 2z \sum_{k=1}^{n} \frac{1}{(z^2 - k^2)} \right\}$$

are not uniformly bounded.

In fact, choosing $z = n + 1/2$, we have for sufficiently large n

$$\begin{aligned} |\pi \cos_n \pi z| &= \frac{1}{n+1/2} + 2(n+1/2) \sum_{k=1}^{n} \frac{1}{n+k+1/2} \, \frac{1}{n-k+1/2} \\ &\geq 2(n+1/2) \sum_{k=1}^{n} \frac{1}{2(n+1/2)} \, \frac{1}{n-k+1/2} \\ &= \sum_{m=1}^{n} \frac{1}{m-1/2} \geq \log n, \end{aligned}$$

where the first inequality above follows from $\frac{1}{n+1/2} > 0$ and $n + k + 1/2 \leq 2(n+1/2)$. This implies that on the strips $|\operatorname{Im} z| \leq \varepsilon < \infty$ and for sufficiently large $|z|$, the uniform bound

$$|\cos_n \pi z| \leq C \log |z|$$

guaranteed by Theorem 2 cannot be improved.

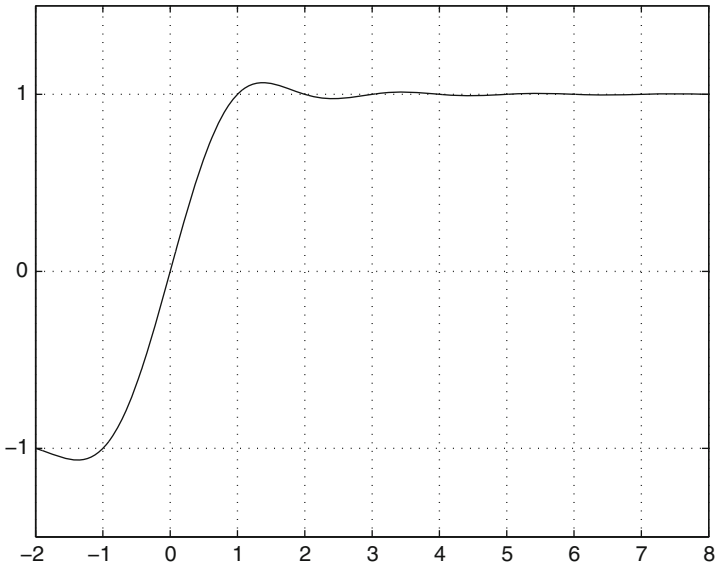

Fig. 1.1 Plot of $\mathrm{Sgn}(x)$ for $-2 \le x \le 8$

Formulas (1.33) and (1.34) are classical and well known, for example, see [1, formulas (11) and (13) on p 188]. But the fact that they also follow from Theorem 2, involving cardinal series expansions seems not to be so well known. We also bring attention to the elementary curiosity concerning the difference of behavior of their respective partial sums.

For the record we also mention the following which follows from the development in Sect. 1.3.6.

Corollary 1. *The function* $\mathrm{Sgn}(z)$ *defined by* (1.35) *is a member of* E_π *that is odd and satisfies both*

$$\mathrm{Sgn}(n) = \mathrm{sgn}(n), \quad n = 0, \pm 1, \pm 2, \ldots$$

and

$$|\mathrm{Sgn}(z)| \le C\mathrm{e}^{\pi|\operatorname{Im} z|},$$

where

$$\mathrm{sgn}(z) = \begin{cases} z/|z| & \text{if } z \ne 0 \\ 0 & \text{when } z = 0 \end{cases}$$

and C is a constant independent of z.

The above considerations suggest that reasonable candidates for a pair of odd functions in E_π that are analogous to the pair of even functions 1 and $\cos \pi z$ might be the pair $\mathrm{Sgn}(z)$ and an odd function $w(z)$ that satisfies (Figs. 1.1 and 1.2)

$$w(n) = (-1)^n \mathrm{sgn}(n), \quad n = 0, \pm 1, \pm 2 \ldots . \tag{1.36}$$

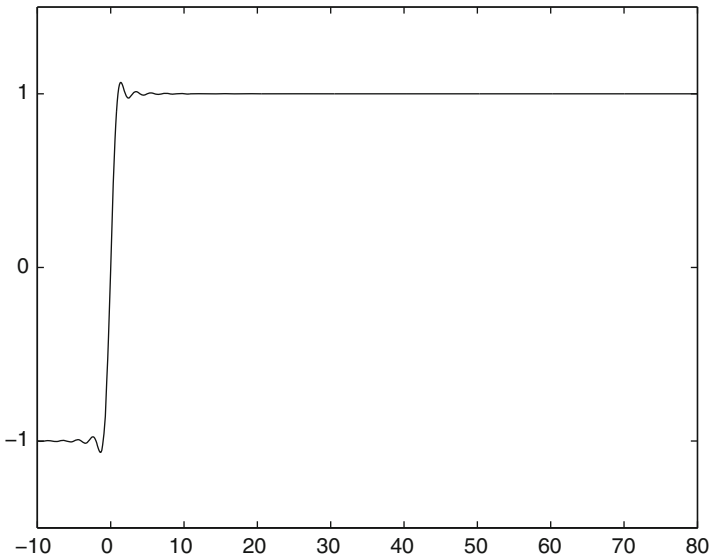

Fig. 1.2 Plot of $\mathrm{Sgn}(x)$ for $-10 \le x \le 80$

Note that in the case of the coefficients $\{w(n)\}$ given by (1.36)

$$\sum_{n=1}^{\infty} (-1)^n \frac{w(n)}{n} \quad \text{does not converge}$$

so in view of second item in Theorem 4, such a function w, unlike the case of the cosine, cannot be represented by a cardinal series (1.1). Nevertheless, if we ignore the second term on the right-hand side of (1.30) and use the coefficients $a(n) = w(n)$ in the first term, we may write

$$w(z) = \frac{2\sin \pi z}{\pi} \left\{ \sum_{n=1}^{\infty} \left(\frac{z^2}{n(z^2 - n^2)} \right) \right\}, \tag{1.37}$$

where the series converges uniformly on compacta and defines an odd entire function in E_π that satisfies (1.36). A calculation analogous to the one used to obtain a lower bound on $|\cos_n(z)|$ shows that

$$|w(N + 1/2)| \ge C \log(N) \quad \text{for sufficiently large } N$$

so that $w(z)$ is not bounded on the strips $|\mathrm{Im}\, z| \le \varepsilon < \infty$. But the function w defined by (1.37) does satisfy

$$|w(z)| \le C e^{\pi |\mathrm{Im}\, z|} \log |z| \quad \text{for sufficiently large } |z|$$

as can be verified by a calculation essentially identical to the one used to establish (1.17).

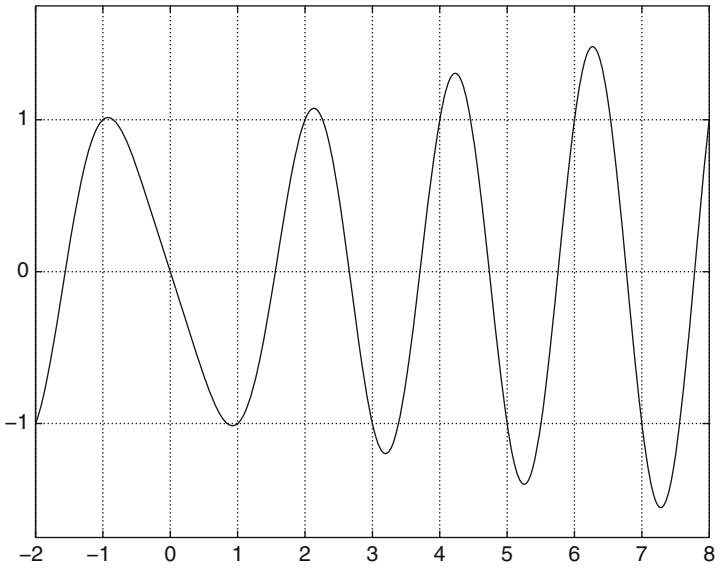

Fig. 1.3 Plot of $\mathrm{Cgn}(x)$ for $-2 \le x \le 8$

If w is the function defined by (1.37), then its derivative at $z = 0$ is 0, namely $w'(0) = 0$. This seems somewhat unnatural. A comparison with $\mathrm{Sgn}(z)$ suggests that the value of this derivative should be $-\mathrm{Sgn}'(0) = -\log 4$. This can be achieved without altering the values at the integers $z = 0, \pm 1, \pm 2, \dots$, by simply adding $-\frac{\log 4}{\pi} \sin \pi z$ to $w(z)$. Thus as an odd analogue of $\cos \pi z$ we propose the function (Figs. 1.3 and 1.4)

$$\mathrm{Cgn}(z) = w(z) - \frac{\log 4}{\pi} \sin \pi z.$$

1.4.3 Special Classes of Data

Here the term data is used to refer to the coefficients $\{a(n)\}$ in (1.1).

As mentioned in the introduction, the class E_π of entire functions $u(z)$ consists of Fourier transforms of distributions $\hat{u}$ with support in the interval $[-\pi, \pi]$. In other words, for every u in E_π there is a distribution $\hat{u}$ with support in the interval $[-\pi, \pi]$ such that $u(z)$ is the value of the distribution $\hat{u}$ evaluated at the test function $\varphi(\xi) = \frac{\mathrm{e}^{iz\xi}}{2\pi}$ that, in the standard notation of linear functionals, can be expressed as

$$u(z) = \langle \varphi, \hat{u} \rangle.$$

In the case that $\hat{u}$ is an integrable function, the last identity can be re-expressed as

$$u(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{iz\xi} \hat{u}(\xi) \mathrm{d}\xi.$$

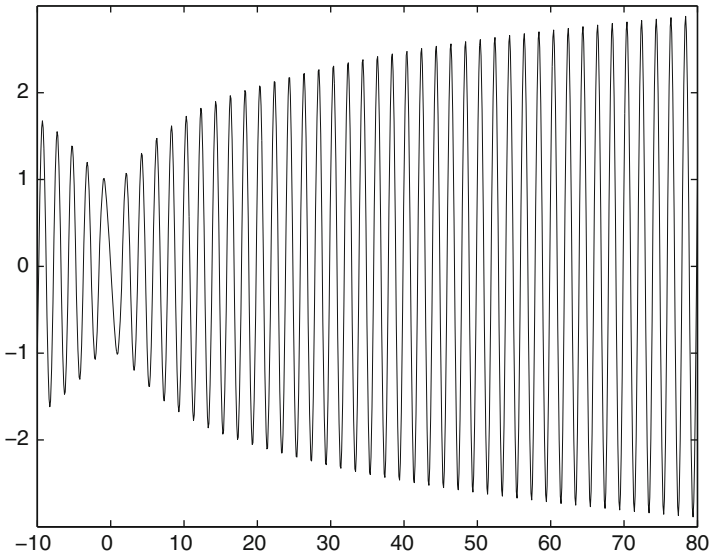

Fig. 1.4 Plot of $\mathrm{Cgn}(x)$ for $-10 \le x \le 80$

The Paley–Wiener class PW consists of those members u of E_π such that $\hat{u}$ is square integrable. This class is often referred to as the class of band-limited functions that plays a very prominent role in classical sampling theory. While it makes sense to refer to the members of E_π as being frequency band-limited, the term "band-limited" is so closely associated with the subclass PW in the literature that to avoid confusion, we have precluded its use in the wider sense.

An issue of interest in sampling theory are requirements on u or $\hat{u}$ that guarantee that the cardinal series $f(z)$ with coefficients $a(n) = u(n)$, $n = 0, \pm 1, \pm 2, \dots$ exists and satisfies the property that $f = u$. In what follows we give several such conditions that are consequences of the results in Sect. 1.2 and are somewhat less restrictive than those associated with classical sampling theory.

As an immediate consequence of the uniqueness statement in Theorem 2 we have

Corollary 2. *Suppose u is an even entire function in E_π such that for some value of $p < 1$*

$$u(x) = O(|x|^p) \quad as \quad x \to \pm\infty$$

on the real axis. Then the symmetric partial sums (1.3) *of the cardinal series with coefficients $a(n) = u(n)$, $n = 0, \pm 1, \pm 2, \dots$ converge uniformly on compacta to $u(z)$.*

If $\hat{u}$ is integrable or, more generally, a finite measure, then $u(z)$ is bounded on the real axis. Hence Corollary 2 can be applied in this case to get

Corollary 3. *Suppose u is an even entire function in E_π such that $\hat{u}$ is an integrable function or, more generally, a finite measure. Then the symmetric partial sums* (1.3)

of the cardinal series with coefficients $a(n) = u(n)$, $n = 0, \pm 1, \pm 2, \ldots$ *converge uniformly on compacta to* $u(z)$.

If k is a nonnegative integer, then PW^k denotes the class of those entire functions u whose derivative of order k, $u^{(k)}$, is in the Paley–Wiener class PW. In other words, $PW^k = \{u : u^{(k)} \in PW\}$. The class PW^k is endowed with the natural semi-norm

$$\|u\|_{PW^k} = \left\{ \int_{-\infty}^{\infty} |u^{(k)}(x)|^2 \mathrm{d}x \right\}^{1/2}.$$

Note that $PW^0 = PW$ and $PW^k \subset PW^{k+1}$ where the containment is proper.

The standard sampling theorem for PW does not apply to PW^k when $k \geq 1$. Nevertheless, it was shown in [14] that members u of PW^k can be recovered from their samples $\{u(n)\}$ via the spline summability method. Additional properties of PW^k can be found in [12].

The following facts concerning PW^k will be useful in what follows:

If u is in PW^k, $k \geq 1$, then

$$u(x) = O\left(|x|^{k-1/2}\right) \quad \text{as} \quad x \to \pm\infty \tag{1.38}$$

on the real axis, and the samples $\{u(n)\}$ enjoy

$$\sum_{n=-\infty}^{\infty} |\Delta^k u(n)|^2 \leq C \|u\|_{PW^k}^2, \tag{1.39}$$

where $\Delta^k u(n)$ are the forward differences of order k of $\{u(n)\}$ that can be defined recursively as

$$\Delta^1 u(n) = \Delta u(n) = u(n+1) - u(n), \quad \Delta^{k+1} u(n) = \Delta^k u(n+1) - \Delta^k u(n).$$

In view of (1.38) Corollary 2 implies the following.

Corollary 4. *Suppose* u *is an even function in* PW^1. *Then the symmetric partial sums* (1.3) *of the cardinal series with coefficients* $a(n) = u(n)$, $n = 0, \pm 1, \pm 2, \ldots$ *converge uniformly on compacta to* $u(z)$.

The corresponding results for odd functions are not quite so transparent. Nevertheless Theorem 4 can be used to show that the following is true.

Corollary 5. *Suppose* u *is an odd entire function in* E_π *such that* $\hat{u}$ *is an integrable function. Then the symmetric partial sums* (1.3) *of the cardinal series with coefficients* $a(n) = u(n)$, $n = 0, \pm 1, \pm 2, \ldots$ *converge uniformly on compacta to* $u(z)$.

To see this, note that $u(n)$ are the Fourier coefficients of $\hat{u}(\xi)$ while $\{(-1)^n u(n)\}$ are the Fourier coefficients of $\hat{u}(\xi - \pi)$. In other words

$$\hat{u}(\xi) \sim \sum_{n=-\infty}^{\infty} u(n)\mathrm{e}^{-in\xi} = -2i\sum_{n=1}^{\infty} u(n)\sin n\xi, \quad -\pi \le \xi \le \pi$$

with partial sums

$$\hat{u}_N(\xi) = -2i\sum_{n=1}^{N} u(n)\sin n\xi, \quad -\pi \le \xi \le \pi$$

and

$$\hat{u}(\xi-\pi) \sim -2i\sum_{n=1}^{\infty} (-1)^n u(n)\sin n\xi, \quad -\pi \le \xi \le \pi.$$

Since both $\hat{u}(\xi)$ and $\hat{u}(\xi-\pi)$ are integrable functions it follows that both

$$\sum_{n=1}^{\infty} \frac{u(n)}{n} \quad \text{and} \quad \sum_{n=1}^{\infty} (-1)^n \frac{u(n)}{n} \quad \text{converge,}$$

see, for example, [21, Theorem 8.7 and the remarks that follow on p 59]. In view of Theorem 4 the convergence of the second series implies that the symmetric partial sums (1.3) of the cardinal series with coefficients $a(n) = u(n)$, $n = 0, \pm 1, \pm 2, \ldots$ converge uniformly on compacta to an entire function $f(z)$.

To see that $f(z) = u(z)$ we argue as follows: the arithmetic means of the partial sums $\hat{u}_n(\xi)$ converge to $\hat{u}(\xi)$ in L^1, for example, see [21, Theorem 5.5(ii) on p 144]. That is

$$\lim_{N\to\infty} \int_{-\pi}^{\pi} \left| \hat{u}(\xi) - \frac{1}{N}\sum_{n=1}^{N} \hat{u}_n(\xi) \right| \mathrm{d}\xi = 0.$$

Hence the arithmetic means of the partial sums f_N of the corresponding cardinal series converge to $u(z)$ uniformly on strips, namely

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f_n(z) = u(z)$$

uniformly on the strips $|\operatorname{Im} z| \le \varepsilon < \infty$. Since the arithmetic means of a sequence converge to the same limit as the original sequence we may conclude that $f(z) = u(z)$.

Corollaries 3 and 5 can be combined to obtain

Corollary 6. *Suppose u is an entire function in E_π such that $\hat{u}$ is an integrable function. Then the symmetric partial sums (1.3) of the cardinal series with coefficients $a(n) = u(n)$, $n = 0, \pm 1, \pm 2, \ldots$ converge uniformly on compacta to $u(z)$.*

Versions of the statement of Corollary 6 have been recorded in [3, Theorem 1 and the cited references] and [7, Theorem 3 on p 70]. For alternate proofs see [2, p 124] and [13, p 499].

There is an analogue of Corollary 4 for odd functions $u(z)$, but its proof is significantly more complicated. For example, to see that (1.39) implies that (1.15) is valid for the coefficients $a(n) = u(n)$ we argue as follows:

$$\begin{aligned}\sum_{n=1}^{\infty}(-1)^n\frac{a(n)}{n} &= \sum_{k=1}^{\infty}\left\{\frac{a(2k)}{2k}-\frac{a(2k-1)}{2k-1}\right\}\\ &= \sum_{k=1}^{\infty}\left\{\frac{a(2k)}{2k}-\frac{a(2k-1)}{2k}\right\}+\sum_{k=1}^{\infty}\left\{\frac{1}{2k}-\frac{1}{2k-1}\right\}a(2k-1)\\ &= \sum_{k=1}^{\infty}\frac{a(2k)-a(2k-1)}{2k}-\sum_{k=1}^{\infty}\frac{a(2k-1)}{2k(2k-1)}.\end{aligned}$$

Now, the Schwarz inequality and (1.39) yield

$$\Big|\sum_{k=1}^{\infty}\frac{a(2k)-a(2k-1)}{2k}\Big|^2 \le \left\{\sum_{k=1}^{\infty}\frac{1}{(2k)^2}\right\}\sum_{k=1}^{\infty}|a(2k)-a(2k-1)|^2 \le C\|u\|^2_{PW^1}$$

while (1.38) yields

$$\sum_{k=1}^{\infty}\frac{|a(2k-1)|}{2k(2k-1)}<\infty.$$

Altogether the above identity and inequalities imply (1.15).

It now follows from Theorem 4 that if $a(n) = u(n)$ and u is in PW^1, then the symmetric partial sums for the cardinal series (1.3) converge to the entire function $f(z)$ given by (1.16). An argument analogous to the one used to prove the uniqueness portion of Theorem 3 shows that $f(z) = u(z) + c\sin\pi z$ where c is a constant. But our argument for the fact that the constant c is indeed 0 involves more intricate properties of PW^k and is too complicated to be included here.

However, let us bring attention to the fact that a variant of the above argument used to show that the coefficients $a(n) = u(n)$ satisfy (1.15) when $u(z)$ is an odd function in PW^1 can be used to show that such coefficients satisfy (1.6) when $u(z)$ is an even function in PW^2.

We summarize these observations as follows:

Corollary 7. *Suppose u is an odd function in PW^1. Then the symmetric partial sums (1.3) of the cardinal series with coefficients $a(n) = u(n)$, $n = 0, \pm1, \pm2, \ldots$ converge uniformly on compacta to $u(z) + c\sin\pi z$. If u is an even function in PW^2, then the symmetric partial sums (1.3) of the cardinal series with coefficients $a(n) = u(n)$, $n = 0, \pm1, \pm2, \ldots$ converge uniformly on compacta to $u(z) + cz\sin\pi z$.*

References

1. Ahlfors LV (1966) Complex analysis: an introduction of the theory of analytic functions of one complex variable, 2nd ed. McGraw-Hill Book Co, New York, Toronto, London
2. Boas RP Jr (1972) Summation formulas and band-limited signals. Tôhoku Math J 24(2):121–125
3. Boche H, Mónich UJ (2009) Behavior of Shannon's sampling series for Hardy spaces. J Fourier Anal Appl 15(3):404–412
4. Brown JL Jr (1968) On the error in reconstructing a non-bandlimited function by means of the bandpass sampling theorem. J Math Anal Appl 18(1967):75–84. (Erratum, ibid, 21, 699)
5. Butzer PL, Higgins JR, Stens RL (2000) Sampling theory of signal analysis. Development of mathematics 1950–2000, pp 193–234. Birkhuser, Basel
6. Butzer PL, Higgins JR, Stens RL (2005) Classical and approximate sampling theorems: studies in the $L^P(\mathbb{R})$ and the uniform norm, J Approx Theor 137(2):250–263
7. Higgins JR (1985) Five short stories about the cardinal series. Bull Am Math Soc (NS) 12(1):45–89
8. Higgins, JR (1996) Sampling theory in Fourier and signal analysis: foundations. Oxford Science Publications, Clarendon Press, Oxford
9. Hörmander L (1969) Linear partial differential operators, 3rd revised printing. Springer, New York
10. Jerri AJ (1977) The Shannon sampling theorem-its various extensions and applications: a tutorial review. Proc IEEE 65(11):1565–1596
11. Levin BYa (1996) Lectures on entire functions. In: Lyubarskii Yu, Sodin M, Tkachenko V (eds) Collaboration with and with a preface. Translated from the Russian manuscript by Tkachenko. Translations of mathematical monographs, vol 150. American Mathematical Society, Providence, RI
12. Madych WR (2001) Summability of Lagrange type interpolation series. J Anal Math 84:207–229
13. Pollard H, Shisha O (1972) Variations on the binomial series. Am Math Mon 79:495–499
14. Schoenberg IJ (1974) Cardinal interpolation and spline functions. VII. The behavior of cardinal spline interpolants as their degree tends to infinity. J Anal Math 27:205–229
15. Walter GG (1988) Sampling bandlimited functions of polynomial growth. SIAM J Math Anal 19(5):11981203
16. Walter GG, Shen X (2001) Wavelets and other orthogonal systems, 2nd edn. Studies in advanced mathematics. Chapman and Hall/CRC, Boca Raton, FL
17. Whittaker ET (1915) On the functions which are represented by the expansions of the interpolation theory. Proc Roy Soc Edinb 35:181–194
18. Whittaker JM (1929) On the cardinal function of interpolation theory. Proc Edinb Math Soc 1:41–46
19. Whittaker JM (1935) Interpolatory function theory. Cambridge tracts in mathematics and mathematical physics, vol 33. Cambridge University Press, Cambridge
20. Zayed AI (1993) Advances in Shannon's sampling theory. CRC Press, Boca Raton, FL
21. Zygmund A (1968) Trigonometric series, reprint of 2nd edn., Vol. I and II. Cambridge University Press, Cambridge

Chapter 2
Improved Approximation via Use of Transformations

Frank Stenger, Maha Youssef, and Jenny Niebsch

Abstract Function interpolation may be carried out using algebraic polynomial, splines, Fourier polynomial, rational functions, wavelets, or Sinc methods. In this chapter we describe methods for getting a more uniform approximation throughout the interval of approximation in the cases when the magnitude of the errors of interpolation is either much larger at one endpoint of the interval than the other, or when the magnitudes of the errors at endpoints are roughly the same, but differ considerably from those errors in the mid-range of the interval. We also discuss improving approximation of the derivative obtained by differentiating the constructed interpolation approximations. This chapter extends the recently obtained results of (Stenger, J Complex 25:292–302, 2009).

2.1 Introduction and Summary

Wavelets have become a powerful tool for solving computational problems of engineering and science and to this end, G.G. Walter has made many excellent contributions [12, 19, 20]. The majority of wavelet applications use Sinc functions [14, 17, 19, 21], although Walter and Shen have made important extensions to other types of wavelets [12,20,21]. We expect that the work of [20,21] will have important applications to other areas of electrical engineering, such as to the work of [7].

F. Stenger (✉)
School of Computing, University of Utah, Salt Lake City, UT 84112, USA
e-mail: sinc_f@msn.com

M. Youssef
Mathematics Department, German University in Cairo, New Cairo City 11835, Egypt
e-mail: Maha.Youssef@GUC.edu.eg.

J. Niebsch
Radon Institute (RICAM), Austrian Academy of Science, 4040 Linz, Austria
e-mail: jenny.niebsch@oeaw.ac.at

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_2,

It is often possible to improve a wavelet or other method of approximation by a very simple procedure, and this is the main purpose of the this chapter.

More specifically, we address in this chapter, the approximations of a function f on an interval or arc Γ, given a set of data points

$$X = \{(x_j, f_j)\}_{j=-M}^{N}, \tag{2.1}$$

where the x_j are distinct points of Γ. We assume that the data X is being interpolated in some way, by a polynomial, a rational function, Fourier polynomial, a spline method, a wavelet method, or a Sinc method.

A novel family of polynomial-like approximations that interpolate given Sinc data of the form $\{(x_j, f_j)\}_{j=-M}^{N}$ where the x_j are Sinc points was recently derived in [16]. Sinc interpolation to this data (see [17, Definition 1.5.12 and Theorem 1.5.13]) is of course accurate, provided that the function f with $f_j = f(x_j)$ belongs to a suitable space of functions. Frequently, we also desire derivative approximations of the function f, and one way of obtaining these is by differentiating the interpolant used to interpolate the data X. On the other hand, this type of approximation of the derivative may not be very accurate, as is the case of Chebyshev polynomial, or Sinc approximation. To this end, the main purpose of this chapter [16] was to be able to get more accurate method of obtaining an approximation for the derivative of the function f at the Sinc points X. This chapter [16] was thus directed mainly to the replacement of Sinc interpolation with polynomial-like interpolation.

Function interpolation by one of the above-stated methods frequently has one of the following features:

(E–E) The modulus of the error of approximation near one *endpoint* of an interval differs considerably from the modulus of the error near the other *endpoint*.

(E–M) The modulus of the error near the two endpoints of the interval is roughly equal but differs appreciably from the modulus of the error in the *mid-range* of the interval.

In this chapter, we extend the methods of [16] by introducing, respectively, "P_{E-E}" and "P_{E-M}"—polynomial-like interpolation at the points x_j. These polynomials are obtained, respectively, via use of transformations on the independent variable. The resulting new variables are rational functions of integer or fractional powers over finite, semi-infinite, infinite intervals, or even over arcs in the complex plane. The derivative of these newly constructed polynomials also enables more accurate approximations of the derivative than the derivative of the original method of interpolation.

As in [16], we initially study the errors in Sinc spaces ([16, Section 1.5.2]), inasmuch as these spaces contain the usual spaces of functions that are analytic on an interval containing Γ and inasmuch as these spaces also house functions that have singularities at endpoints of Γ.

Sinc methods enable uniform approximation of functions on Γ, in spite of possible singularities at endpoints of Γ. We thus also present a short review of basic Sinc notation, and we then present the usual and well-known Lagrange polynomial interpolation in Sect. 2.2. In Sect. 2.3, we introduce the polynomial P_{E-E} referred to above, for approximating a function on $[-1,1]$, as well as for approximating the derivative of f at x_j by differentiating this polynomial. In Sect. 2.4 we extend the results of Sect. 2.3 to the approximation over a more general arc Γ. In Sect. 2.5 we present the transformation P_{E-M}, initially for approximation f and its derivative on $[-1,1]$, and we then extend these results to approximation over a more general arc Γ, via use of the methods of Sect. 2.4. In Sect. 2.6 we present several examples of applications of our results.

2.2 Sinc Notation and Interpolation Formulas

In this section we recall some Sinc notation, and we also discuss polynomial interpolation. Our polynomial interpolation presentation differs somewhat from that of [16], where the interval $[0,1]$ was the starting point of derivation. Here it is more convenient to derive the initial polynomial approximation for the interval $[-1,1]$, for purposes of obtaining simpler formulas for other intervals.

2.2.1 Some Sinc Concepts

Let us first establish some mathematical notation which we shall require. Let $\mathbb{Z}$ denote the set of all integers, $\mathbb{R}$ the real line, and $\mathbb{C}$ the complex plane $\{a+ib : a \in \mathbb{R}, b \in \mathbb{R}\}$.

For a positive number $h \in \mathbb{R}$ and for $x \in \mathbb{C}$, the Sinc function $\mathrm{sinc}(x)$ is defined by ([14, 22])

$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}.$$

This function has value 1 at $x = 0$, and vanishes at all other integer values of x. It is more convenient for application to use the notation $S(k,h)(x)$ for the function

$$S(k,h)(x) = \mathrm{sinc}\left(\frac{x}{h} - k\right).$$

Given a number $d > 0$, we define the strip $\mathscr{D}_d$ as

$$\mathscr{D}_d = \{z \in \mathbb{C} : |\Im z| < d\}.$$

Let $\mathscr{D}$ be a simply connected domain having a boundary $\partial\mathscr{D}$, and let a and b denote two distinct points of $\partial\mathscr{D}$. Let φ denote a conformal map of $\mathscr{D}$ onto $\mathscr{D}_d$ such

that $\varphi(a) = -\infty$ and φ (b) = ∞, and let us define the inverse conformal map by $\psi = \varphi^{-1}$. Let us also define ρ by $\rho(z) = \mathrm{e}^{\varphi(z)}$. In addition, let Γ be an arc defined by

$$\Gamma = \{z \in \mathbb{C} : z = \psi(u) : u \in \mathbb{R}\}.$$

Letting α, β, and d denote arbitrary positive numbers, we denote by $\mathbb{L}_{\alpha,\beta}(\varphi)$ the family of all functions that are analytic in $\mathscr{D}$ such that for all $z \in \mathscr{D}$, we have

$$|f(z)| \le C \frac{|\rho(z)|^{\alpha}}{[1+|\rho(z)|]^{\alpha+\beta}}.$$

We next restrict the above numbers α, β, and d such that $0 < \alpha \le 1$, $0 < \beta \le 1$, and $0 < d < \pi$, to define another class of functions $\mathbb{M}_{\alpha,\beta}(\varphi)$ to be the set of all functions g defined on $\mathscr{D}$ that have finite limits $g(a) = \lim_{z\to a} g(z)$ and $g(b) = \lim_{z\to b} g(z)$ where the limits are taken from within $\mathscr{D}$, and such that $f \in \mathbb{L}_{\alpha,\beta}(\varphi)$, where

$$f = g - \frac{g(a) + \rho\, g(b)}{1+\rho}.$$

A one-dimensional approximation for a function f defined on an arc Γ can be obtained by applying the following Sinc interpolation formula:

$$f(x) \approx f_{M,N}(x) = \sum_{k=-M}^{N} f(x_k)\, \omega_k(x), \quad x \in \Gamma, \tag{2.2}$$

where $x_k = \psi(kh)$ are *Sinc points* on Γ and $\omega_k(x)$ are Sinc basis functions. These Sinc basis functions are defined in the above notation, for arbitrary Γ, as follows. Note that $\rho(x_k) = \mathrm{e}^{kh}$.

$$\begin{aligned}
\delta_j(x) &= S(j,h) \circ \varphi(x), \quad j = -M, \ldots, N \\
\omega_j(x) &= \delta_j(x), \quad j = -M+1, \ldots, N-1 \\
\omega_{-M}(x) &= \frac{1}{1+\rho(x)} - \sum_{j=-M+1}^{N} \frac{\delta_j(x)}{1+\mathrm{e}^{jh}} \\
\omega_N(x) &= \frac{\rho(x)}{1+\rho(x)} - \sum_{j=-M}^{N-1} \frac{\mathrm{e}^{jh}\,\delta_j(x)}{1+\mathrm{e}^{jh}},
\end{aligned} \tag{2.3}$$

cf. [14]. Using the interpolation (2.2) in calculations generates an accurate approximation with an exponentially decaying error rate which for $M = [\beta N/\alpha]$ is given for $h = (\pi d/(\beta N))^{1/2}$, by

$$|E_N| = \|f - f_N\| \le K N^{\frac{1}{2}} \mathrm{e}^{-\sqrt{\pi d \alpha N}}, \tag{2.4}$$

where $[\cdot]$ denotes the greatest integer function, where $\|\cdot\|$ denotes the sup norm on Γ, and where K is a constant independent of N.

It seems natural to obtain an approximation to $f^{(n)}$, the nth derivative of f, by differentiating equation (2.2) with respect to x, which yields the approximation

$$f^{(n)}(x) \approx f_{M,N}^{(n)}(x) = \sum_{k=-M}^{N} f(x_k)\, \omega_k^{(n)}(x), \quad x \in \Gamma, \tag{2.5}$$

where x_k are Sinc points on Γ and $\omega_k(x)$ are Sinc basis function defined in (2.3). *If* $M = \lfloor \beta N/\alpha \rfloor$, if h is selected as above, and if φ' is uniformly bounded on Γ, then this approximation has an error bounded by [8]

$$\| \{f - f_{M,N}\}^{(n)} \| \leq K_n \sqrt{N^{n+1}} \mathrm{e}^{-\sqrt{\pi d \alpha N}}, \tag{2.6}$$

with constant K_n depending only on f and n, but independent of N. Otherwise the approximation (2.5), while converging to $f^{(n)}$ on compact subsets of Γ, is unbounded at endpoints of Γ.

Unfortunately, the formula (2.5) is useful only for some transformations φ of $\mathbb{R}$ onto $\mathbb{R}$. The formula (2.5) always yields unbounded results in neighborhoods of finite endpoints of Γ (see e.g., [6]).

2.2.2 *Sinc and Lagrange Polynomial Approximation*

Our polynomial interpolation presentation differs somewhat from that of [16], where the interval $[0, 1]$ was the starting point of derivation. Here it is more convenient for purposes of simplicity of expression of the novel formulas which we shall derive to consider polynomial approximation on the interval $[-1, 1]$. The polynomial methods of [16] and of this chapter, in fact, polynomials of a rational function of $\rho = \mathrm{e}^{\varphi}$.

Consider, for example, the case when each of the basis functions ω_j of (2.2) are just the functions δ_j of the equation following (2.2). In this case, the Sinc approximation (2.2) may be written in the form ([14, 15])

$$f_{M,N}(x) = \frac{h}{\pi} \sin\left(\frac{\pi}{h}\varphi(x)\right) \sum_{k=-M}^{N} (-1)^k \frac{f(x_k)}{\varphi(x) - k h}, \quad x \in \Gamma. \tag{2.7}$$

This equation shows that the Sinc approximation is itself a product of $\sin(\pi\varphi(x)/h)$ and a rational function of $\varphi(x)$. Indeed, as was demonstrated in [1–3], via the introduction of a novel method of barycentric interpolation, the factor $\sin(\pi\varphi(x)/h)$ is not even necessary for purposes of evaluation of $f_{M,N}(x)$. Rational function methods of approximation of f were also introduced in ([13], [14, Thm. 5.2.5]); those also were rationals in $\rho = \mathrm{e}^{\varphi}$ (but different ones from the "polynomial-like" rationals of this chapter), which interpolated at the same Sinc points, and which had the same order of error as the Sinc approximation (2.7).

2.2.3 *Lagrange Polynomial Interpolation on* $[-1, 1]$

Generally Lagrange polynomial approximation over the interval $[-1,1]$ is defined in the following way.

Given a set of $m = M+N+1$ distinct points $\{x_j\}_{j=-M}^{N}$ on the interval $[-1,1]$ and function values, $\{f(x_j)\}_{j=-M}^{N}$, at these points, there exists a unique polynomial $p(x)$ of degree at most $m-1$ satisfying,

$$p(x_j) = f(x_j), \quad j = -M, -M+1, \ldots, N. \tag{2.8}$$

Here $p(x)$ can be expressed as follows:

$$p(x) = \sum_{j=-M}^{N} b_j(x) f(x_j), \tag{2.9}$$

where

$$b_j(x) = \frac{g(x)}{(x-x_j)g'(x)}, \tag{2.10}$$

where

$$g(x) = \prod_{l=-M}^{N} (x-x_l). \tag{2.11}$$

2.2.4 *The Derivative of* $p(x)$

We get the derivative of p by differentiating Eqs. (2.9) and (2.10) with respect to x. To this end, and for our later purposes, it is convenient to define an $m \times m$ matrix $\mathbf{A} = [a_{j,k}]$, $j,k = -M, \ldots, N$ such that

$$f'(x_j) \approx p'(x_j) = \sum_{k=-M}^{N} a_{j,k} f(x_k). \tag{2.12}$$

Here we have

$$a_{j,k} = b_k'(x_j) = \begin{cases} \dfrac{g'(x_j)}{(x_j - x_k)g'(x_k)} & \text{if } k \neq j. \\ \displaystyle\sum_{l=-M, l\neq j}^{N} \frac{1}{x_j - x_l} & \text{if } k = j. \end{cases} \tag{2.13}$$

2.2.5 Error Estimates for Sinc Data

We now derive bounds on the errors of the approximation for the case of Sinc data for the interval $[-1,1]$. These estimates are similar to the ones introduced in [16]. The following theorem will give an error estimate for the approximation of the function $f(x)$ by $p(x)$ as well as an error bound for the approximation of $f'(x_j)$ by $p'(x_j)$ for the case when the x_j are Sinc points, and when $f \in \mathbf{M}_{\alpha,\beta}(\varphi)$, with $\varphi(x) = \log((1+x)/(1-x))$, this being a conformal map of the region.

$D_2 = D \cup_{y\in(-1,1)} B(y,r)$ where $r > 0$, and $B(y,r) = \{z \in \mathbb{C} : |z-y| < r\}$.

We shall assume here for sake of simplicity of derivation of our results that $M = N$.

Theorem 1. *Let $M = N$, $h = \frac{\eta}{\sqrt{N}}$ with a constant $\eta > 0$ and independent of N, and let $\{x_j\}_{j=-N}^{N}$ denote the Sinc points as defined above. Let f be in $\mathbb{M}_{\alpha,\beta}(\varphi)$, analytic and bounded in D_2, and let p and $p'(x_j)$ given by (2.9) and (2.12). Then there exist two constants A and B, independent of N, such that*

$$\|f(x) - p(x)\| \le A\frac{\sqrt{N}}{r^{2N}} \exp\left(\frac{-\pi^2 N^{\frac{1}{2}}}{2\eta}\right) \tag{2.14}$$

and

$$\max_{j=-N,\ldots,N} |f'(x_j) - p'(x_j)| \le B\frac{N}{r^{2N}} \exp\left(\frac{-\pi^2 N^{\frac{1}{2}}}{2\eta}\right). \tag{2.15}$$

The proof of the theorem requires three lemmas, the proofs of which can be found in [16, 17].

Lemma 2. *Let $h = \frac{\eta}{\sqrt{N}}$, with η a positive constant and with N a positive integer. If $z = \mathrm{e}^{\frac{h}{2}}$ then*

$$\frac{z^N}{(1+z)^{2N}} \le 2^{-2N}, \tag{2.16}$$

and there exists a constant A_1 independent of N such that

$$\sum_{j=1}^{N} \log(1 - z\mathrm{e}^{-jh}) \le \frac{-\pi^2 N^{\frac{1}{2}}}{6c\eta} + \log(N^{\frac{1}{4}}) + A_1\,. \tag{2.17}$$

Lemma 3. *Let $h = \frac{\eta}{\sqrt{N}}$, with η a positive constant and with N a positive integer. Then there exist two constants A_2 and A_3 independent of N such that*

$$\sum_{j=1}^{N} \log(1 - \mathrm{e}^{-jh}) \le \frac{-\pi^2 N^{\frac{1}{2}}}{6\eta} + \log(N^{\frac{1}{2}}) + A_2, \tag{2.18}$$

and

$$\sum_{j=1}^{N} \log(1+\mathrm{e}^{-jh}) \geq \frac{-\pi^2 N^{\frac{1}{2}}}{12\,\eta} + \log(N^{\frac{1}{2}}) - A_3\,. \tag{2.19}$$

Lemma 4. *Let $g(x)$ be defined as (2.11). Then*

$$|g'(x_0)| = \max_{j=-N,\ldots,N} |g'(x_j)|. \tag{2.20}$$

We now prove Theorem 1. This will be carried out via a sequence of sub-proofs.

Proof of Theorem 1: Proof of (2.14)

The error in approximating a function $f(x)$ by $p(x)$ using m interpolation points is given by the following contour integral, for the error $E_m(f,x) = f(x) - p(x)$,

$$E_m(f,x) = \frac{g(x)}{2\pi i} \int\limits_{\partial D_2} \frac{f(z)}{(z-x)g(z)}\,\mathrm{d}z. \tag{2.21}$$

To find the upper bound of this integral, we will use the definition of D_2 which gives $|z-x| \geq r$ and $|z-x_j| \geq r$ for all x and x_j in $[-1,1]$ and $z \in \partial D_2$. Since we assumed $f(x)$ to be bounded in D_2, we take $|f(x)| \leq \mathscr{B}(f)$ in D_2. Then

$$|E_m(f,x)| = |f(x) - p(x)| \leq \frac{\mathscr{B}(f)}{r^{m+1}} \max_{x\in[-1,1]} |g(x)| \frac{L(\partial D_2)}{2\pi}, \tag{2.22}$$

where $L(\partial D_2) \leq 4 + 2\pi r$ is the length of ∂D_2.

Since $\rho(x) = \frac{1+x}{1-x}$, we have $x = \frac{\rho-1}{\rho+1}$; this transformation maps $\rho \in (0,\infty)$ to $x \in (-1,1)$. We estimate that the maximum value of g takes place approximately at $\rho = \exp(\frac{h}{2})$ (see [16]). Under these assumptions, we have

$$\begin{aligned}
\left| g\left(\frac{\rho-1}{\rho+1}\right)\right| &= \prod_{j=-N}^{N} \left|\frac{\rho-1}{\rho+1} - \frac{\mathrm{e}^{jh}-1}{\mathrm{e}^{jh}+1}\right| \\
&= \prod_{j=-N}^{N} \left|\frac{2\rho - 2\mathrm{e}^{jh}}{(\rho+1)(1+\mathrm{e}^{jh})}\right| \\
&= \frac{2^{2N}\rho^{N+1}}{(1+\rho)^{2N+1}} \prod_{j=1}^{N} \left(\frac{1-\rho\mathrm{e}^{-jh}}{1+\mathrm{e}^{-jh}}\right)^2 |1-\mathrm{e}^{-(N+1)h}|.
\end{aligned}$$

Now using the fact that $|1-\mathrm{e}^{-(N+1)h}| \leq 1$ and that $\log\prod_{j=1}^{N} a_j = \sum_{j=1}^{N} \log a_j$, we apply Lemmas 2 and 3 to get

$$\left| g\left(\frac{\rho-1}{\rho+1}\right)\right| \leq C_1\sqrt{N}\mathrm{e}^{\frac{-\pi^2 N^{\frac{1}{2}}}{2\eta}},$$

where C_1 is a constant independent of N and h. Using this bound in (2.22) yields the right-hand side of (2.14).

Proof of (2.15):

We can similarly bound the approximation error for the derivative, $f'(x_j)$, by $p'(x_j)$:

$$E_m(f') = |f'(x_j) - p'(x_j)| \leq \frac{B(f)}{r^{m+1}} \max_{j\in(-N,\ldots,N)} |g'(x_j)| \frac{L(\partial D_2)}{2\pi}. \tag{2.23}$$

Using Lemma 4 we have

$$\max_{j\in(-N,\ldots,N)} |g'(x_j)| = |g'(x_0)| = \prod_{k=1}^{N} |x_k \cdot x_{-k}|.$$

Since $M = N$, we have $x_{-k} = -x_k$ on $(-1,1)$, and therefore

$$|g'(x_0)| = \prod_{k=1}^{N} (x_k)^2 = \prod_{k=1}^{N} \left(\frac{e^{kh}-1}{e^{kh}+1}\right)^2.$$

Now we only need to use Lemma 3 to get the right-hand side of (2.15).

2.3 Approximating with Transformed Polynomials

All of the "polynomials" which we shall construct are for the interval $[-1,1]$. Hence given the data X of (2.1) on the arc Γ, let $\xi = \xi(x)$ denote a one-to-one transformation of the arc Γ to the interval $[-1,1]$. This transformation transforms the distinct points X of (2.1) to the distinct points

$$X' = \{\xi_j\}_{j=-M}^{N}, \quad \xi_j = \xi(x_j). \tag{2.24}$$

Our "polynomial" takes the form

$$P(\xi) \equiv \sum_{j=-M}^{N} \frac{G(\xi) f_j}{(\xi - \xi_j) G'(\xi_j)}, \tag{2.25}$$

in which

$$G(\xi) = \prod_{j=-M}^{N} (\xi - \xi_j). \tag{2.26}$$

The evaluation of $P_{E-E}(x)$ for approximating $f(x)$ using (2.25) is straightforward by taking $\xi = \xi(x)$ in (2.25). On the other hand, we find that

$$\frac{dP(\xi(x))}{dx}|_{x=x_j} = \sum_{k=-M}^{N} a_{j,k} \frac{d\xi(x)}{dx}|_{x=x_j} f_k, \tag{2.27}$$

where, from (2.13),

$$a_{j,k} = \begin{cases} \dfrac{G'(\xi_j)}{(\xi_j - \xi_k)G'(\xi_k)} & \text{if } k \neq j \\ \displaystyle\sum_{l=-M, l\neq j}^{N} \frac{1}{(\xi_j - \xi_k)} & \text{if } k = j \end{cases} \tag{2.28}$$

and where explicit expressions for the evaluation of $\frac{d\xi(x)}{dx}$ will depend on explicit expressions for $\xi(x)$ and will be given below.

2.4 Approximating with the Polynomials P_{E-E}

The purpose of the polynomials P_{E-E} is to make more equal the magnitudes of the errors of interpolation in neighborhoods of the endpoints of Γ.

2.4.1 The Polynomials P_{E-E} on $[-1, 1]$

We consider here the approximation of f and f' on $[-1, 1]$, but with polynomials P_{E-E}.

Recall, for the case of the arc $(-1, 1)$, we have $\varphi = \varphi_2$, with $\varphi_2(x) = \log((1 + x)/(1 - x))$, $\rho(x) = \rho_2(x) = \exp(\varphi_2(x))$, and $x = (\rho_2(x) - 1)/(\rho_2(x) + 1)$.

We assume that we are given a (polynomial, trigonometric polynomial, rational function, spline, wavelet, or Sinc) method of interpolating the data X, of (2.1), i.e., of the given data

$$X = \{(x_j, f_j)\}_{j=-M}^{N},$$

where the x_j are distinct points on $[-1, 1]$ (i.e., not necessarily Sinc points) listed in increasing order. Suppose furthermore that the magnitudes of the errors of approximation differ appreciably near the endpoints ± 1. In this case, consider replacing x with ξ, where

$$\xi = \frac{x - c}{1 - cx} \tag{2.29}$$

and where c is a point of $(-1,1)$. This transformation ξ is a one-to-one transformation of the interval $[-1,1]$ to itself. Our aim is to obtain a polynomial approximation with polynomial in the variable ξ, such that the errors of our polynomial near ± 1 are approximately equal, by suitably selecting c.

Our "polynomial" takes the form (2.25) and (2.26), in which $\xi_j = (x_j - c)/(1 - cx_j)$.

If $c < 0$, then more of the given Sinc points are shifted towards the right endpoint, which decreases the error at this endpoint and increases it at the left endpoint, while for $c > 0$, we achieve the opposite effect.

The evaluation of $P_{E-E}(x)$ for approximating $f(x)$ using (2.25) is straightforward by taking $\xi = \xi(x)$ in (2.25), with $\xi(x)$ given in (2.29).
Similarly, we can approximate the derivative of f at x_j using (2.27), with

$$\frac{\mathrm{d}\xi(x)}{\mathrm{d}x} = \frac{1-c^2}{(1-cx)^2}. \tag{2.30}$$

2.4.2 P_{E-E} for Other Intervals

Let φ be a conformal map defined as in Sect. 2.2.1 above, which transforms Γ to $\mathbb{R}$. Recall [8] that if $F \in M_{\alpha,\beta}(\varphi)$, then $f = F \circ \varphi^{-1} \circ \varphi_2 \in M_{\alpha,\beta}(\varphi_2)$. Furthermore for the case of $\Gamma_2 = (-1,1)$, we have $x = (\rho_2 - 1)/(\rho_2 + 1)$, where according to our above definition, $\rho_2 = \exp(\varphi_2)$. Inasmuch as the above polynomial $P(x)$ of Sect. 2.2.4 is a suitable polynomial for approximation on $\Gamma_2 = (-1,1)$, it follows from Sinc theory [17] that a polynomial $P_{E-E} = P_{E-E}(\xi)$ in the variable $\xi = (\rho(x) - 1)/(\rho(x) + 1)$ with $\rho(x) = \exp(\varphi(x))$ will then be a good approximation on $\Gamma = \varphi^{-1}((R))$.

Hence, if we take $\rho(x) = \exp(\phi(x))$, then (with $c \in (-1,1)$ as above), the new shifted polynomial P_{E-E} with shift from $\psi(0)$ to $\psi(\log((1+c)/(1-c)) = \psi(\log(q))$ is a polynomial $P_{E-E}(\xi)$ in the variable $\xi = (\rho(x) - q)/(\rho(x) + q)$, with $q = (1+c)/(1-c)$.

Now suppose that we are given the above set of data points X, with the x_j distinct points of Γ, and that we wish to interpolate this data with this "polynomial" $P_{E-E}(\xi)$. The interpolation points ξ_j then become $\xi_j = (\rho(x_j) - q)/(\rho(x_j) + q)$. The "polynomial" P_{E-E} thus takes the form of (2.25). This polynomial P_{E-E} interpolates the data X, i.e., we have $P_{E-E}(\xi_j) = f_j$, and it can be readily evaluated using (2.25) and (2.26) to get a method of approximating $f(x) \approx P_{E-E}(\xi(x))$ on Γ. To approximate $f'(x_j)$, on Γ, we use the derivative $\frac{\mathrm{d}P_{E-E}(\xi(x))}{\mathrm{d}x}$ which can be evaluated via use of (2.27) and (2.28), with

$$\frac{\mathrm{d}\xi(x)}{\mathrm{d}x} = \frac{2q\rho(x)\varphi'(x)}{(\rho(x)+q)^2}. \tag{2.31}$$

2.5 Approximating with the Polynomials P_{E-M}

In this case, we assume that the errors of the given method of approximation in neighborhoods of the endpoints of Γ are approximately equal in magnitude, but differ considerably in magnitude from the errors in the mid-range of Γ, and we now wish to have similar magnitudes of the errors at the endpoints and mid-range of Γ.

2.5.1 P_{E-M} for $[-1,1]$

Recall once again, that for the case of $\Gamma = \Gamma_2 = (-1,1)$, we gave $x = (\rho_2(x) - 1)/(\rho_2(x)+1)$ where $\rho_2(x) = (1+x)/(1-x)$. In order to equalize the errors at the endpoints of Γ_2 with those of the mid-range of γ_2, we take

$$\begin{aligned} \xi = \xi(x) &= \frac{\rho_2(x)^\alpha - 1}{\rho_2(x)^\alpha + 1} \\ &= \frac{(1+x)^\alpha - (1-x)^\alpha}{(1+x)^\alpha + (1-x)^\alpha}. \end{aligned} \tag{2.32}$$

The polynomial $P_{E-M}(\xi(x))$ can then be evaluated via use of (2.25), taking $\xi_j = \xi(x_j)$, and we can use this polynomial to get a new approximation to the function $f(x)$ on $[-1,1]$. Evidently, these new points of interpolation ξ_j have properties similar to those of the Sinc points of Γ_2. Hence, if the original method of interpolation is based on classical polynomial (e.g., Chebyshev, or Newton–Cotes) methods, then this new method of approximation should yield greater accuracy when f has singularities at the endpoints of Γ_2.

Similarly, we can use (2.27) to get a new approximation to $f'(x_j)$, where we now require use of the expression

$$\xi'(x_j) = \frac{4\,\alpha\,(1-x_j)^{2\,\alpha-2}}{((1+x_j)^\alpha + (1-x_j)^\alpha)^2}. \tag{2.33}$$

2.5.2 P_{E-M} for an Arbitrary Arc Γ

We now define ξ by the equation

$$\xi = \xi(x) = \frac{\rho(x)^\alpha - 1}{\rho(x)^\alpha + 1}, \tag{2.34}$$

where α is a positive number.

Towards evaluation of the resulting polynomial $P_{E-M}(\xi(x))$, we use (2.25) and (2.26) in which we take $\xi = \xi(x)$. To approximate $f'(x_j)$, we simply differentiate $P_{E-M}(\xi(x))$ with respect to x, and for this purpose, we require

$$\frac{\mathrm{d}\xi(x)}{\mathrm{d}x} = \frac{2\,\alpha\,\rho(x)^{\alpha}\,\varphi'(x)}{(\rho(x)^{\alpha}+1)^2}\,. \tag{2.35}$$

2.5.3 A P_{E-M} for Equi-Spaced Interpolation

Equi-spaced interpolation corresponds to Sinc interpolation on the real line, in which case, we have $\varphi(x) = x$, so that $\rho(x) = \exp(x)$. Thus (2.34) yields the transformation

$$\xi = \xi(x) = \frac{\rho(x)^{\alpha}-1}{\rho(x)^{\alpha}+1} = \frac{\mathrm{e}^{\alpha x}-1}{\mathrm{e}^{\alpha x}+1}\,. \tag{2.36}$$

In this case, we have

$$\frac{\mathrm{d}\xi(x)}{\mathrm{d}x} = \frac{2\,\alpha\,\mathrm{e}^{\alpha x}}{(\mathrm{e}^{\alpha x}+1)^2}\,. \tag{2.37}$$

2.5.4 A More General P_{E-M} for $[-1,1]$

We can postulate a further generalization of the transformation (2.32) above, for polynomial approximation on $[-1,1]$. These transformations are given by

$$\xi = \xi(x) = \frac{(1+x)^{\alpha}-(1-x)^{\beta}}{(1+x)^{\alpha}+(1-x)^{\beta}}, \tag{2.38}$$

or more generally, by

$$\xi = \xi(x) = \frac{(1+x)^{\alpha}-q\,(1-x)^{\beta}}{(1+x)^{\alpha}+q\,(1-x)^{\beta}}\,. \tag{2.39}$$

In (2.39), q, α, and β are arbitrary positive numbers. These transformations may enable polynomial approximations that are more efficient than those made possible using (2.32) in the case when f has different Lipschitz behavior at the two endpoints ± 1 of I_2. The transformation (2.39) is somewhat more general than (2.38), in that it also enables a shift of the origin, in the spirit of Sect. 2.4.1.

We may note that for the case of (2.39), we have

$$\frac{\mathrm{d}\xi(x)}{\mathrm{d}x} = \frac{2q\,(1-x)^{\alpha-1}(1+x)^{\beta-1}\,(\beta\,(1+x)+\alpha\,(1-x))}{((1+x)^{\alpha}+q\,(1-x)^{\beta})^2}, \tag{2.40}$$

which is positive on $(-1,1)$ for all positive values of α, β, and q and which shows that the transformations (2.38) and (2.39) are one-to-one transformations of the interval $(-1,1)$ to itself.

In this chapter, we neither consider further study of these transformations, nor their extension to other arcs.

2.5.5 *Some Explicit Transformations Based on Sinc*

At this point, we express some polynomial approximations in polynomials in the variable ξ as a function of x as follows (see [17, Section 1.5.3] for more detail and for other such transformations).

1. If $\varphi(x) = x$, then $\Gamma = \mathbb{R}$, the Sinc points are $x_j = jh$, and $\xi = (\rho(x)-1)/(\rho(x)+1) = (e^x-1)/(e^x+1)$.
2. If $\varphi(x) = \phi_2(x) = \log((1+x)/(1-x)$, then $\Gamma = \Gamma_2 = (-1,1)$, the Sinc points are $x_j = (e^{jh}-1)/(e^{jh}+1)$, and $(\rho_2-1)/(\rho_2+1) = x$.
3. If $\varphi(x) = \log(x)$, then $\Gamma = (0,\infty)$, the Sinc points are $x_j = e^{jh}$, and $\xi = (\rho(x)-1)/(\rho(x)+1) = (x-1)/(x+1)$.
4. If $\varphi(x) = \log(\sinh(x))$, then $\Gamma = (0,\infty)$, the Sinc points are $x_j = \log(e^{jh} + \sqrt{1+e^{2jh}}$, and $\xi = (\rho(x)-1)/(\rho(x)+1) = (\sinh(x)-1)/(\sinh(x)+1)$.
5. If $\varphi(x) = \log(x+\sqrt{1+x^2})$, then $\Gamma = \mathbb{R}$, the Sinc points are $x_j = \sinh(jh)$, and $\xi = (\rho(x)-1)/(\rho(x)+1) = (\sqrt{x^2+1}+x-1)/(\sqrt{x^2+1}+x+1)$.

Other interesting examples are possible via use of the excellent double exponential transformations of [9, 10, 18].

2.6 Numerical Examples

In this section we give some examples of improving a given approximation of a function as well as examples of improving approximations of the derivative of the given approximation. All of our plots are obtained by evaluation both functions and approximations at the points $x = -1 + (j-1/2)/1000$, for $j = 1,\ldots,1000$.

2.6.1 *Improving Approximations to Runge's Function*

In [11] Runge studied the approximation of the function

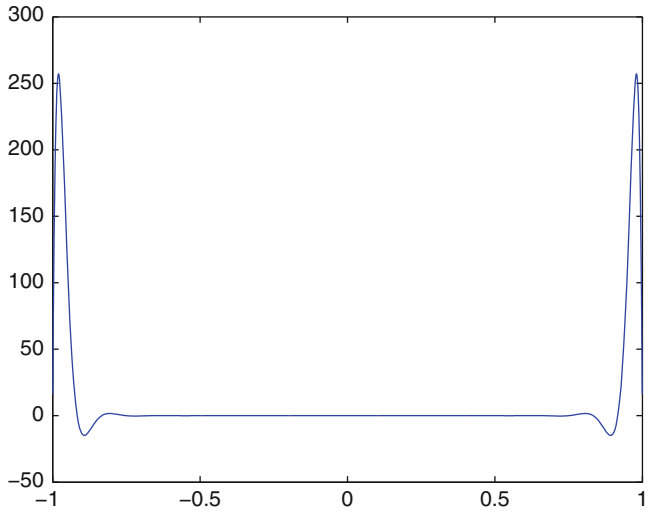

Fig. 2.1 Newton–Cotes polynomial error

$$\begin{aligned} f(x) &= \frac{1}{1+x^2} \Longleftrightarrow \quad -5 \le x \le 5 \\ &= \frac{1}{1+25x^2} \Longleftrightarrow -1 \le x \le 1 \end{aligned} \tag{2.41}$$

using equi-spaced Newton–Cotes polynomials. Although Runge's problem is over a century old, it is only in recent times that exponentially convergent Runge defeating methods have been developed [4, 5]. In this example we illustrate the errors of approximation of this function.

Example 1. In Fig. 2.1 we have plotted the difference between f and the Newton–Cotes po lynomial P_{2N+1} that interpolates f at the points $x_j = -1 + (j-1)/N$, $j = 1,\ 2,\ \ldots,\ 2N+1$, with $N = 12$.
In Fig. 2.2, we plotted the difference $f\prime - P'_{2N+1}$, with P_{2N+1} as in Fig. 2.1.

Example 2. In Figs. 2.3 and 2.4, we have plotted the errors of Sinc approximation of f, and f', namely,

$$\begin{aligned} f(x) &- \sum_{j=-N}^{N} f(x_j)\, S(j,h)(x) \\ f'(x) &- \sum_{j=-N}^{N} f(x_j)\, (S(j,h)(x))', \end{aligned} \tag{2.42}$$

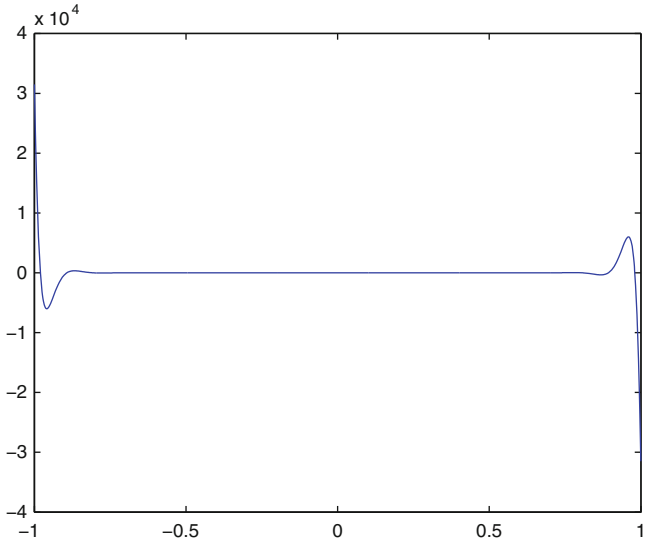

Fig. 2.2 Derivative of Newton–Cotes polynomial error

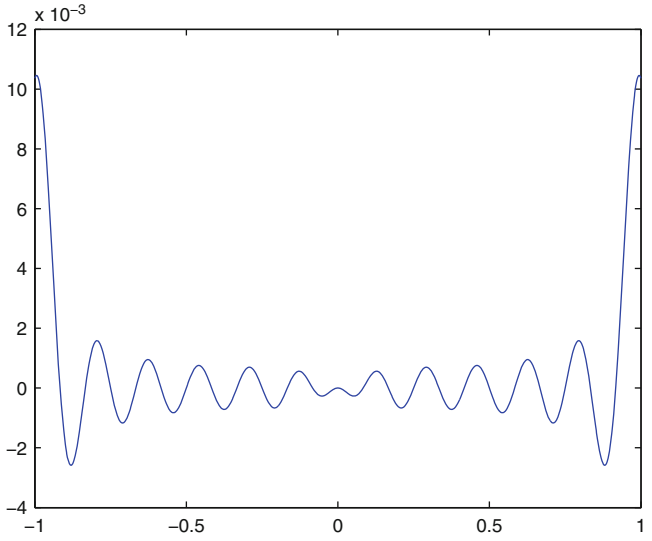

Fig. 2.3 Sinc approximation error

with $h = 1/N$, and with $N = 12$. The approximation is reasonably accurate, but not impressively so, since, while "Sinc points" $\{jh\}$ are used ([17, Section 1.5.1]), the function f does not belong to the space of [17, Example 1.5.4].

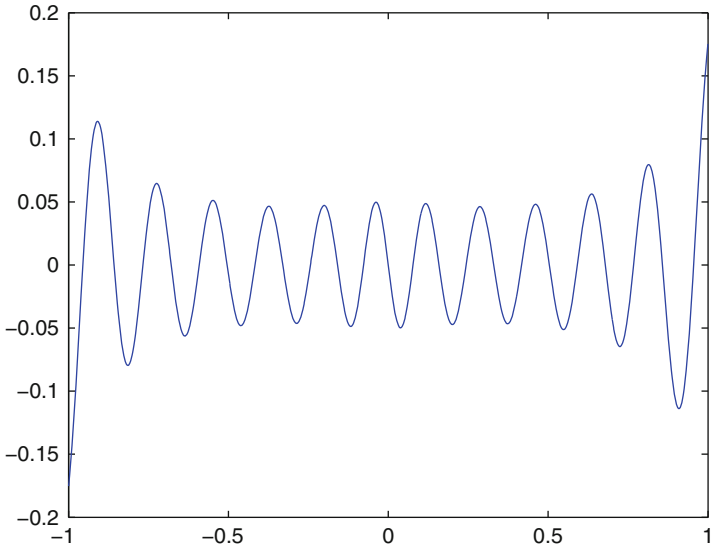

Fig. 2.4 Derivative of Sinc approximation error

Example 3. Here we use the polynomial $P_{E-M}(\xi(x))$ based on the transformation $\xi = (e^{\alpha x} - 1)/(e^{\alpha x} + 1)$, a transformation motivated by item #1 of Sect. 2.5.5, since the interpolation points $\{jh\}$ are Sinc points of the real line $\mathbb{R}$. Figures 2.5 and 2.6 illustrate plots of the differences

$$\begin{aligned} & f(x) - P_{E-M}(\xi(x)), \\ & \frac{\mathrm{d}f(x)}{\mathrm{d}x} - \frac{\mathrm{d}P_{E-M}(\xi(x))}{\mathrm{d}x}. \end{aligned} \tag{2.43}$$

Example 4. Here we use the polynomial $P_{E-M}(\xi(x))$ based on the transformation

$$\xi = \xi(x) = \frac{(1+x)^{\alpha} - (1-x)^{\alpha})}{1+x)^{\alpha} + (1-x)^{\alpha}}, \tag{2.44}$$

the transformation given in (2.32). We get the following figures that illustrate the approximations as given in (2.42), but with P_{E-M} now depending on the ξ of (2.43) (Figs. 2.3 and 2.4).

Example 5. Here we illustrate use of the transformation of Sect. 2.4.1, which we may, upon observing that the f is an even function, and that the interpolation points x_j are evenly distributed on $[-1,1]$. We first extend the domain of approximation from $[0,1]$ to $[-1,1]$ by means of the transformation $y = 2x^2 - 1$, and we then approximate over $[-1,1]$ with the variable $z = (y-c)/(1+yc)$. We have only

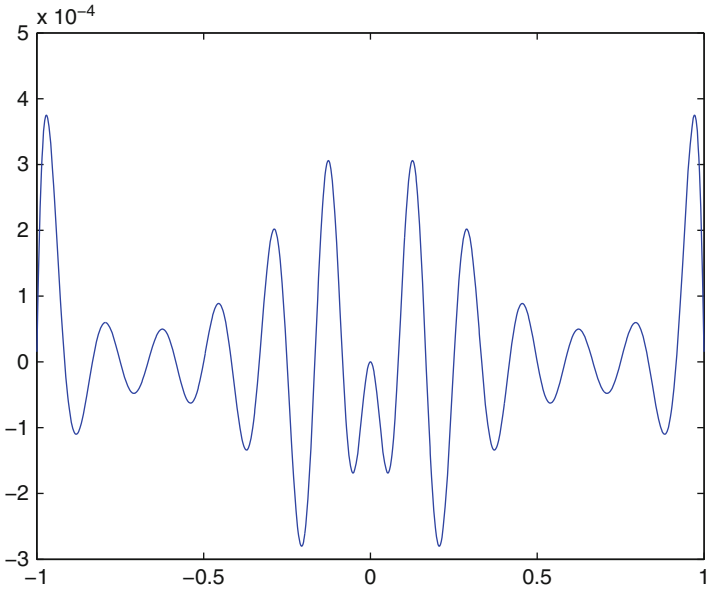

Fig. 2.5 Exponential polynomial error

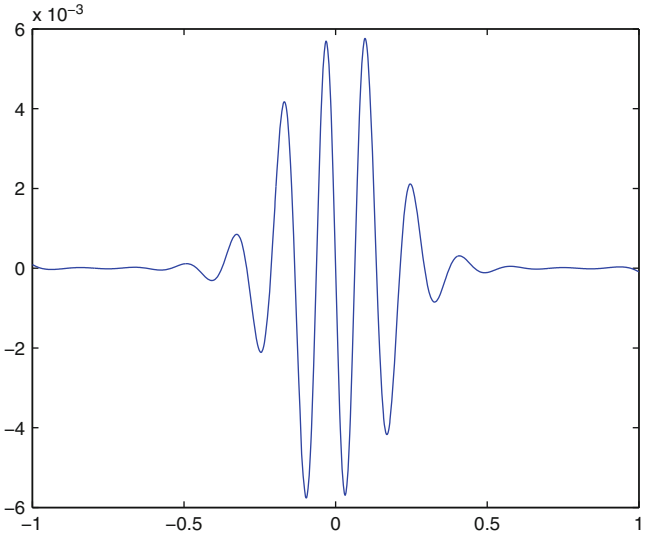

Fig. 2.6 Derivative of exponential polynomial error

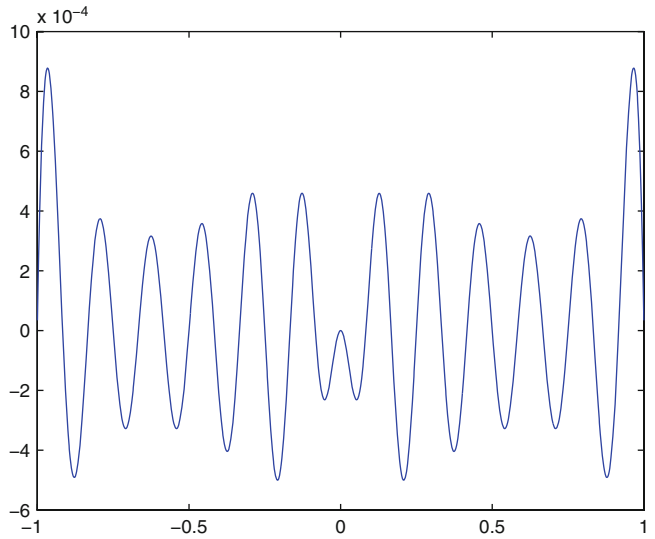

Fig. 2.7 Trans algebraic polynomial error

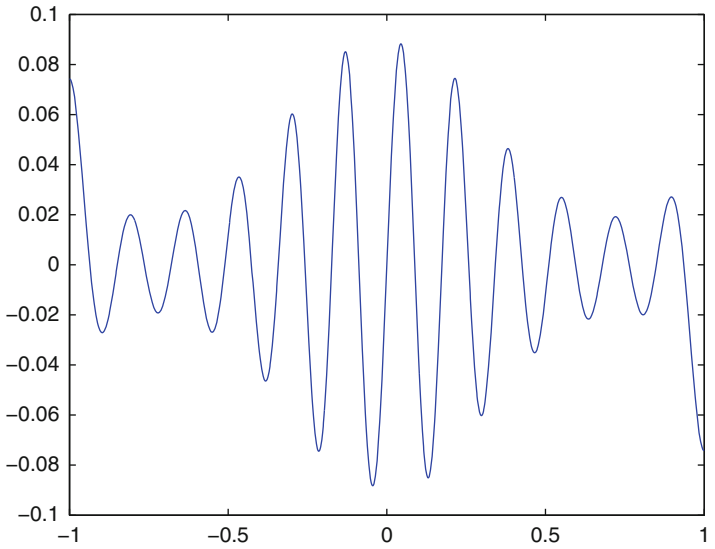

Fig. 2.8 f-prime minus trans algebraic polynomial prime

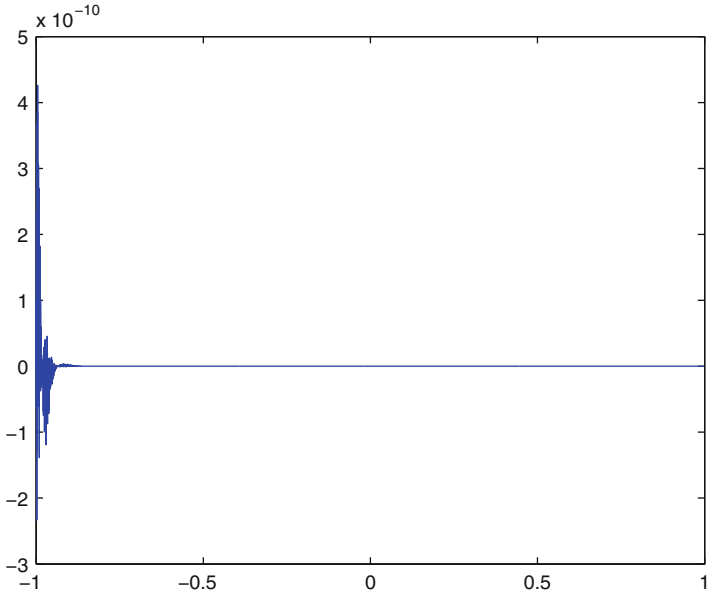

Fig. 2.9 f minus trans circ poly approx error

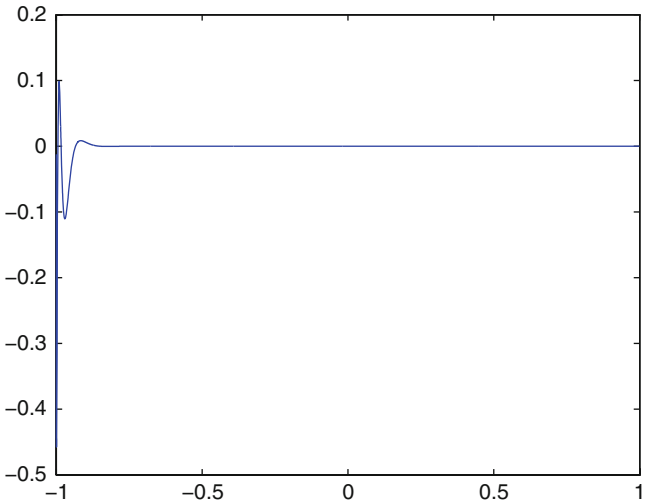

Fig. 2.10 Trans circ poly prime error

plotted for $0 \leq x \leq 1$; the plot for $x \to -x$ is similar, due to symmetry. Note the incredibly accurate approximation thus obtainable (Figs. 2.9 and 2.10).

2.6.2 More Uniform Chebyshev Polynomial Approximation

Use of Chebyshev polynomial approximation has become an accurate way to solve many problems in applications. The results are usually very accurate, except in the presence of singularities at endpoints of the interval of approximation. To this end we have carried out several tests of polynomial approximations to functions that have singularities at endpoints of the interval, and we have reached the following conclusions:

1. The approximation of $f(x) = (1-x)^{5/4}$ and its derivative by means of the transformation (2.28). Only about 10 % improvement was possible, mainly because this transformation is unable to remove the singularity.
2. The approximation of $(1-x^2)^{5/4}$ using the transformation (2.32). Here, too, only about 10 % improvement was possible.
3. The approximation of the function $(1-x)^{1/4}$ using the transformation (2.38). Fantastic improvement was possible, with proper choice of α and β, as is illustrated in Figs. 2.11 and 2.12 below. Best results were obtained for $N = 12$, with $\alpha = 0.86$, and $\beta = 0.387$. Evidently, this transformation has possibilities. We were not able to determine a priori the parameters α and β .
4. The transformation (2.39) had almost no advantages over (2.38), i.e., whenever we tried it, best approximation always occurred with q very close to 1 . These parameters were determined via a "trial and error" Bellman-style selection process, i.e., we first fixed q and we then determined α and β . A different order of selection may determine different values of these parameters.

2.6.3 More Uniform Wavelet Approximation

Here, too, we did several examples to study wavelet approximations. Methods for improving approximations in cases when f has on singularities at endpoints of the interval are well known (see [17, Section 3.10]). Hence we turned our studies only to cases of when f has singularities at endpoints of the interval. To this end, our conclusions were similar to those in the previous section, i.e., the transformations (2.29) and (2.36) produced relatively little improvement, whereas (2.37) produced excellent results.

We include here the function $f(x) = x(\pi^2 - x^2)^{1/4}$, for which we have assumed a given approximation at the points $j\pi/(N+1)$ of the interval $(0, \pi)$. We assumed that this odd period function was approximated via use of the wavelet approximation ([17, (1.4.15)])

$$W_N(x) = \frac{\sin(Nx)}{N} \sum_{k=1}^{N-1} \frac{(-1)^k \sin\left(\frac{k\pi}{N}\right) f\left(\frac{k\pi}{N}\right)}{\cos(x) - \cos\left(\frac{k\pi}{N}\right)}. \tag{2.45}$$

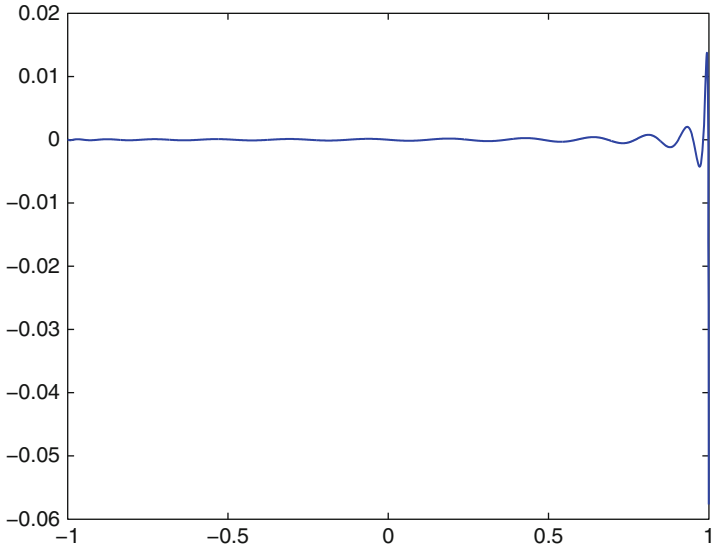

Fig. 2.11 Error of Chebyshev polynomial approximation

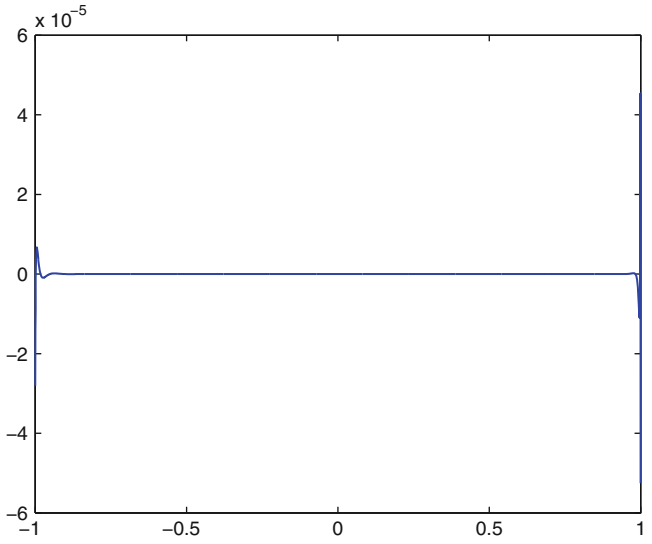

Fig. 2.12 Error of transformed Chebyshev polynomial approximation

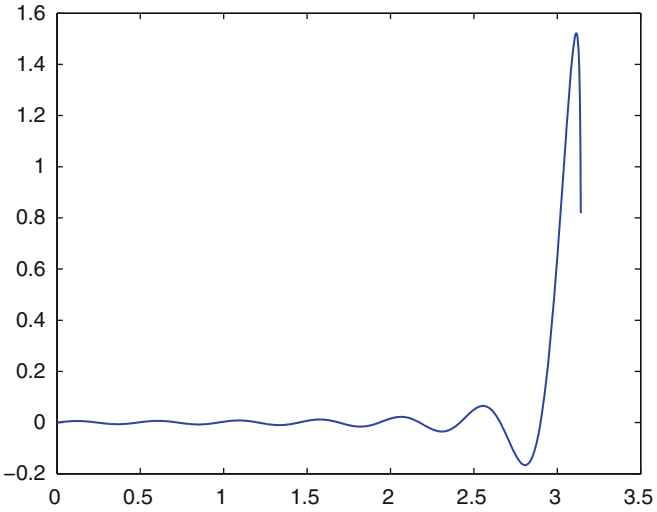

Fig. 2.13 Error of wavelet approximation

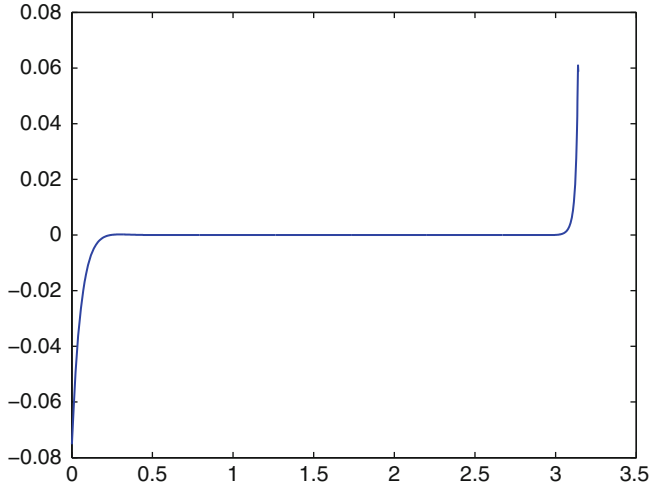

Fig. 2.14 Error of transformed wavelet approximation

We observe from Figs. 2.13 and 2.14 that in taking $N = 12$, using the transformation (2.39) and making the selections $\alpha = 0.94$, $\beta = 0.457$, and $c = 0.9999$ reduce the error of approximation by about a factor of 25 .

2.7 Conclusion

In this chapter, we introduced polynomial-like procedures for improving the approximation of a function using given interpolating data, although independently of the method of approximation. To this end, up to three parameters $c \leftrightarrow q$ and α and β were used in our approximations. These parameters were determined by trial and error. They depend not only on the function f that is being approximated but also on the order of approximation. At this time we do not have a method for determining these parameters a priori.

References

1. Berrut J-P (2007) A formula for the error of finite sinc-interpolation over a finite interval. Numer Algor 45:369–374
2. Berrut J-P (2011) A formula for the error of finite sinc interpolation with an even number of nodes. Numer Algor 56:143–157
3. Berrut J-P, Baltensperger R, Mittelmann H (2006) Recent developments in barycentric rational interpolation, international series of numerical mathematics. Trends Appl Constr Approx 151:27–51
4. Boyd JP (2001) Chebyshev and Fourier spectral methods, 2nd edn. Dover, Mineola, New York, p 665
5. Boyd JP, Ong J (2009) Exponentially–convergent strategies for defeating the runge-phenomenon for the approximation of non-periodic functions, part I: single-interval schemes. Commun Comput Phys 5(2–4):484–497
6. Hohn MH (2006) Solution of singular elliptic PDEs on Union of rectangles using sinc methods. ETNA 23:88–104
7. Hagmann MJ, Kumar G, Pandey S, Nahata A (2011) Analysis and simulation of generating ter-ahertz surface waves on a tapered field emission tip. Vac Sci Tech B 29(2):2B113-1–2B113-6
8. Kowalski M, Sikorski K, Stenger F (1995) Selected topics of approximation and computation. Oxford University Press
9. Mori M (2005) Discovery of the double exponential transformation and its developments. Publ Res Inst Math Sci Kyoto Univ 41:897–935
10. Okayama T, Matsuo T, Sugihara M (2010) Sinc collocation for weakly singular Fredholm integral equations of the second kind. J Comput Appl Math (234) 1211–1227
11. Runge C (1901) Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten. Zeitschrift fr Mathematik und Physik 46:224–243
12. Shen XA, Walter GG (2002) Meyer wavelet regularization. Numer Funct Anal Optim 23(1–2):195–215
13. Stenger F (1986) Explicit, nearly optimal, linear rational approximations with preassigned poles. Math Comput 47:225–252
14. Stenger F (1993) Numerical methods based on sinc and analytic functions. Springer, New York
15. Stenger F (2000) Summary of Sinc numerical methods. J Comput Appl Math 121:379–420
16. Stenger F (2009) Polynomial function and derivative approximation of Sinc data. J Complex 25:292–302
17. Stenger F (2010) Handbook of sinc numerical methods. CRC Press, Boca Raton, FL, USA
18. Sugihara M, Matsuo T (2004) Recent developments of the sinc numerical methods, proceedings of the 10th international congress on computational and applied mathematics (ICCAM-2002). J Comput Appl Math 164/165:673–689

19. Walter GG, Shen X (2000) Wavelets and other orthogonal systems, 2nd edn. CRC Press, Boca Raton, FL, USA
20. Walter GG, Shen X (2003) Sampling with prolate spheroidal wave functions. J Sampling Theor Signal Image Process 2(1):25–52
21. Walter GG, Shen X (2004) Wavelets based on spheroidal wave functions. J Fourier Anal Appl 10:1–25
22. Zayed A (1993) Advances in shannon's sampling theory. CRC Press, Boca Raton FL

Chapter 3
Generalized Sampling in $L^2(\mathbb{R}^d)$ Shift-Invariant Subspaces with Multiple Stable Generators

H.R. Fernández-Morales, A. G. García, and G. Pérez-Villalón

Abstract In order to avoid most of the problems associated with classical Shannon's sampling theory, nowadays, signals are assumed to belong to some shift-invariant subspace. In this work we consider a general shift-invariant space $V^2_{\boldsymbol{\Phi}}$ of $L^2(\mathbb{R}^d)$ with a set $\boldsymbol{\Phi}$ of r stable generators. Besides, in many common situations, the available data of a signal are samples of some filtered versions of the signal itself taken at a sub-lattice of $\mathbb{R}^d$. This leads to the problem of generalized sampling in shift-invariant spaces. Assuming that the ℓ^2-norm of the generalized samples of any $f \in V^2_{\boldsymbol{\Phi}}$ is stable with respect to the $L^2(\mathbb{R}^d)$-norm of the signal f, we derive frame expansions in the shift-invariant subspace allowing the recovery of the signals in $V^2_{\boldsymbol{\Phi}}$ from the available data. The mathematical technique used here mimics the Fourier duality technique which works for classical Paley–Wiener spaces.

3.1 By Way of Introduction

The classical Whittaker–Shannon–Kotel'nikov sampling theorem (WSK sampling theorem) [23, 52] states that any function f band-limited to $[-1/2, 1/2]$, i.e., $f(t) = \int_{-1/2}^{1/2} \widehat{f}(w) \mathrm{e}^{2\pi i t w} \mathrm{d}w$ for each $t \in \mathbb{R}$, may be reconstructed from the sequence of samples $\{f(n)\}_{n\in\mathbb{Z}}$ as

$$f(t) = \sum_{n=-\infty}^{\infty} f(n) \frac{\sin \pi(t-n)}{\pi(t-n)}, \quad t \in \mathbb{R}.$$

H.R. Fernández-Morales • A.G. García (✉)
Departamento de Matemáticas, Universidad Carlos III de Madrid, 28911 Leganés (Madrid), Spain
e-mail: hfernand@math.uc3m.es; agarcia@math.uc3m.es

G. Pérez-Villalón
Departamento de Matemática Aplicada, E.U.I.T.T., U.P.M., 28031 Madrid, Spain
e-mail: gperez@euitt.upm.es

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_3,

Thus, the Paley–Wiener space $PW_{1/2}$ of band-limited functions to $[-1/2,1/2]$ is generated by the integer shifts of the cardinal sine function, $\mathrm{sinc}(t) := \sin \pi t/\pi t$. A simple proof of this result is given by using the Fourier duality technique which uses that the Fourier transform

$$\begin{aligned} \mathscr{F} : PW_{1/2} &\longrightarrow L^2[-1/2,1/2] \\ f &\longmapsto \widehat{f} \end{aligned}$$

is a unitary operator from the Paley–Wiener space $PW_{1/2}$ of band-limited functions to $[-1/2,1/2]$ onto $L^2[-1/2,1/2]$. Thus, applying the inverse Fourier transform $\mathscr{F}^{-1}$ to the Fourier series $\widehat{f} = \sum_{n=-\infty}^{\infty} f(n)\,\mathrm{e}^{-2\pi i n w}$ of $\widehat{f}$ in $L^2[-1/2,1/2]$ one gets

$$\begin{aligned} f(t) &= \sum_{n=-\infty}^{\infty} f(n)\mathscr{F}^{-1}\left[\mathrm{e}^{-2\pi i n w}\chi_{[-1/2,1/2]}(w)\right](t) \\ &= \sum_{n=-\infty}^{\infty} f(n)\,\frac{\sin \pi(t-n)}{\pi(t-n)} \ \text{ in } \ L^2(\mathbb{R})\,. \end{aligned}$$

The pointwise convergence comes from the fact that $PW_{1/2}$ is a reproducing kernel Hilbert space (written shortly as RKHS) where convergence in norm implies pointwise convergence (which is, in this case, uniform on $\mathbb{R}$); this comes out from the inequality $|f(t)| \le \|f\|$ for each $t \in \mathbb{R}$ and $f \in PW_{1/2}$ (for the RKHS's theory and applications, see, for instance, [37]).

The WSK theorem has its d-dimensional counterpart. Any function f band-limited to the d-dimensional cube $[-1/2,1/2]^d$, i.e., $f(t) = \int_{[-1/2,1/2]^d} \widehat{f}(x)\mathrm{e}^{2\pi i x^\top t}dx$ for each $t \in \mathbb{R}^d$ (here we are using the notation $x^\top t := x_1 t_1 + \cdots + x_d t_d$ identifying elements in $\mathbb{R}^d$ with column vectors), may be reconstructed from the sequence of samples $\{f(\alpha)\}_{\alpha\in\mathbb{Z}^d}$ as

$$f(t) = \sum_{\alpha\in\mathbb{Z}^d} f(\alpha)\frac{\sin \pi(t_1-\alpha_1)}{\pi(t_1-\alpha_1)}\cdots\frac{\sin \pi(t_d-\alpha_d)}{\pi(t_d-\alpha_d)}, \quad t=(t_1,\ldots,t_d)\in\mathbb{R}^d,$$

where $\alpha = (\alpha_1,\ldots,\alpha_d)$. Although Shannon's sampling theory has had an enormous impact, it has a number of problems, as pointed out by Unser in [44,45]: It relies on the use of ideal filters; the band-limited hypothesis is in contradiction with the idea of a finite duration signal; the band-limiting operation generates Gibbs oscillations; and finally, the sinc function has a very slow decay at infinity which makes computation in the signal domain very inefficient. Besides, in several dimensions, it is also inefficient to assume that a multidimensional signal is band-limited to a d-dimensional interval. Moreover, many applied problems impose different a priori constraints on the type of signals. For this reason, sampling and reconstruction problems have been investigated in spline spaces, wavelet spaces, and general shift-invariant spaces; signals are assumed to belong to some shift-invariant space of the

form $V_\varphi^2 := \overline{\text{span}}_{L^2}\{\varphi(t-\alpha) : \alpha \in \mathbb{Z}^d\}$ where the function φ in $L^2(\mathbb{R}^d)$ is called the generator of V_φ^2. See, for instance, [1,3,4,6,7,10,24,45,47,49–51,53] and references therein.

In this new context, the analogous of the WSK sampling theorem in a shift-invariant space V_φ^2 was first time proved by Walter in [47].

3.1.1 Walter's Sampling Theorem in Shift-Invariant Spaces

Let $\varphi \in L^2(\mathbb{R})$ be a stable generator for the shift-invariant space V_φ^2 which means that the sequence $\{\varphi(\cdot - n)\}_{n\in\mathbb{Z}}$ is a Riesz basis for V_φ^2. A Riesz basis in a separable Hilbert space is the image of an orthonormal basis by means of a bounded invertible operator. Any Riesz basis $\{x_n\}_{n=1}^\infty$ has a unique biorthogonal (dual) Riesz basis $\{y_n\}_{n=1}^\infty$, i.e., $\langle x_n, y_m\rangle_{\mathscr{H}} = \delta_{n,m}$, such that the expansions

$$x = \sum_{n=1}^{\infty} \langle x, y_n\rangle_{\mathscr{H}}\, x_n = \sum_{n=1}^{\infty} \langle x, x_n\rangle_{\mathscr{H}}\, y_n$$

hold for every $x \in \mathscr{H}$ (see [11] for more details and proofs). Recall that the sequence $\{\varphi(\cdot - n)\}_{n\in\mathbb{Z}}$ is a Riesz sequence, i.e., a Riesz basis for V_φ^2 (see, for instance, [11, p 143]) if and only if there exist two positive constants $0 < A \le B$ such that

$$A \le \sum_{k\in\mathbb{Z}} |\widehat{\varphi}(w+k)|^2 \le B, \quad \text{a.e. } w \in [0,1].$$

Thus, we have that $V_\varphi^2 = \{\sum_{n\in\mathbb{Z}} a_n\, \varphi(\cdot - n) : \{a_n\} \in \ell^2(\mathbb{Z})\} \subset L^2(\mathbb{R})$.

We assume that the functions in the shift-invariant space V_φ^2 are continuous on $\mathbb{R}$. This is equivalent to say that the generator φ is continuous on $\mathbb{R}$ and the function $\sum_{n\in\mathbb{Z}} |\varphi(t-n)|^2$ is uniformly bounded on $\mathbb{R}$ (see [42]). Thus, any $f \in V_\varphi^2$ is defined on $\mathbb{R}$ as the pointwise sum $f(t) = \sum_{n\in\mathbb{Z}} a_n \varphi(t-n)$ for each $t \in \mathbb{R}$.

On the other hand, the space V_φ^2 is the image of the Hilbert space $L^2[0,1]$ by means of the isomorphism

$$\begin{aligned} \mathscr{T}_\varphi : L^2[0,1] &\longrightarrow V_\varphi^2 \\ \{\mathrm{e}^{-2\pi i n x}\}_{n\in\mathbb{Z}} &\longmapsto \{\varphi(t-n)\}_{n\in\mathbb{Z}}, \end{aligned}$$

which maps the orthonormal basis $\{\mathrm{e}^{-2\pi i n w}\}_{n\in\mathbb{Z}}$ for $L^2[0,1]$ onto the Riesz basis $\{\varphi(t-n)\}_{n\in\mathbb{Z}}$ for V_φ^2. For any $F \in L^2[0,1]$ we have

$$\begin{aligned} \mathscr{T}_\varphi F(t) &= \sum_{n\in\mathbb{Z}} \left\langle F, \mathrm{e}^{-2\pi i n x}\right\rangle \varphi(t-n) \\ &= \left\langle F, \sum_{n\in\mathbb{Z}} \overline{\varphi(t-n)}\mathrm{e}^{-2\pi i n x}\right\rangle = \langle F, K_t\rangle_{L^2[0,1]}, \quad t \in \mathbb{R}, \end{aligned}$$

where, for each $t \in \mathbb{R}$, the function $K_t \in L^2[0,1]$ is given by

$$K_t(x) = \sum_{n\in\mathbb{Z}} \overline{\varphi(t-n)}\mathrm{e}^{-2\pi inx} = \overline{\sum_{n\in\mathbb{Z}} \varphi(t+n)\mathrm{e}^{-2\pi inx}} = \overline{Z\varphi(t,x)}\,.$$

Here, $Z\varphi(t,x) := \sum_{n\in\mathbb{Z}} \varphi(t+n)\mathrm{e}^{-2\pi inx}$ denotes the Zak transform of the function φ. See [11, 22] for properties and uses of the Zak transform.

As a consequence, the samples in $\{f(a+m)\}_{m\in\mathbb{Z}}$ of $f \in V^2_\varphi$, where $a \in [0,1)$ is fixed, can be expressed as

$$f(a+m) = \langle F, K_{a+m}\rangle = \langle F, \mathrm{e}^{-2\pi imx}K_a\rangle\,, \quad m \in \mathbb{Z} \text{ where } F = \mathscr{T}_\varphi^{-1} f\,.$$

Thus, the stable recovery of $f \in V^2_\varphi$ from the sequence of its samples $\{f(a+m)\}_{m\in\mathbb{Z}}$ reduces to the study of the sequence $\left\{\mathrm{e}^{-2\pi imx}K_a(x)\right\}_{m\in\mathbb{Z}}$ in $L^2[0,1]$. Recall that the operator $m_F : L^2[0,1] \to L^2[0,1]$ given as the product $m_F(f) = Ff$ is well defined if and only if $F \in L^\infty[0,1]$, and then, it is bounded with norm $\|m_F\| = \|F\|_\infty$. As a consequence, the following result comes out:

Theorem 1. *The sequence of functions* $\left\{\mathrm{e}^{-2\pi imx}K_a(x)\right\}_{m\in\mathbb{Z}}$ *is a Riesz basis for* $L^2[0,1]$ *if and only if the inequalities* $0 < \|K_a\|_0 \le \|K_a\|_\infty < \infty$ *hold, where* $\|K_a\|_0 := \operatorname{ess\,inf}_{x\in[0,1]} |K_a(x)|$ *and* $\|K_a\|_\infty := \operatorname{ess\,sup}_{x\in[0,1]} |K_a(x)|$. *Moreover, its biorthogonal Riesz basis is* $\left\{\mathrm{e}^{-2\pi imx}/\overline{K_a(x)}\right\}_{m\in\mathbb{Z}}$.

In particular, the sequence $\left\{\mathrm{e}^{-2\pi imx}K_a(x)\right\}_{m\in\mathbb{Z}}$ is an orthonormal basis in $L^2[0,1]$ if and only if $|K_a(x)| = 1$ a.e. in $[0,1]$.

Let a be a real number in $[0,1)$ such that $0 < \|K_a\|_0 \le \|K_a\|_\infty < \infty$; next, we prove Walter's sampling theorem for V^2_φ in [47]. Given $f \in V^2_\varphi$, we expand the function $F = \mathscr{T}_\varphi^{-1} f \in L^2[0,1]$ with respect to the Riesz basis $\left\{\mathrm{e}^{-2\pi inx}/\overline{K_a(x)}\right\}_{n\in\mathbb{Z}}$. Thus, we get

$$F = \sum_{n\in\mathbb{Z}} \langle F, K_{a+n}\rangle \frac{\mathrm{e}^{-2\pi inx}}{\overline{K_a(x)}} = \sum_{n\in\mathbb{Z}} f(a+n)\frac{\mathrm{e}^{-2\pi inx}}{\overline{K_a(x)}} \text{ in } L^2[0,1]\,.$$

Applying the operator $\mathscr{T}_\varphi$ to the above expansion we obtain

$$f = \sum_{n\in\mathbb{Z}} f(a+n)\mathscr{T}_\varphi\left(\mathrm{e}^{-2\pi inx}/\overline{K_a(x)}\right) = \sum_{n\in\mathbb{Z}} f(a+n)S_a(\cdot - n) \text{ in } L^2(\mathbb{R})\,,$$

where we have used the shifting property $\mathscr{T}_\varphi(\mathrm{e}^{-2\pi inx}F)(t) = (\mathscr{T}_\varphi F)(t-n)$, $t \in \mathbb{R}$, and $n \in \mathbb{Z}$, satisfied by the isomorphism $\mathscr{T}_\varphi$ for the particular function $S_a := \mathscr{T}_\varphi(1/\overline{K_a}) \in V^2_\varphi$. As in the Paley–Wiener case, the shift-invariant space V^2_φ is a RKHS. Indeed, for each $t \in \mathbb{R}$, the evaluation functional at t is bounded:

$$|f(t)| \le \|F\|\|K_t\| \le \|\mathscr{T}_\varphi^{-1}\|\|K_t\|\|f\| = \|\mathscr{T}_\varphi^{-1}\|\left(\sum_{n\in\mathbb{Z}} |\varphi(t-n)|^2\right)^{1/2}\|f\|, \quad f \in V^2_\varphi.$$

Therefore, the L^2-convergence implies pointwise convergence which here is uniform on $\mathbb{R}$. The convergence is also absolute due to the unconditional convergence of a Riesz expansion. Thus, for each $f \in V^2_\varphi$, we get the sampling formula

$$f(t) = \sum_{n=-\infty}^{\infty} f(a+n) S_a(t-n), \quad t \in \mathbb{R}. \tag{3.1}$$

This mathematical technique, which mimics the Fourier duality technique for Paley–Wiener spaces [23], has been successfully used in deriving sampling formulas in other sampling settings [14, 16, 17, 19, 21, 25, 31, 32]. In this work, it will be used for obtaining generalized sampling formulas in $L^2(\mathbb{R}^d)$ shift-invariant subspaces with multiple stable generators.

3.1.2 Statement of the General Problem

Assume that our functions (signals) belong to some shift-invariant space of the form

$$V^2_{\boldsymbol{\Phi}} := \overline{\operatorname{span}}_{L^2(\mathbb{R}^d)} \left\{ \varphi_k(t-\alpha) : k = 1, 2, \ldots, r \text{ and } \alpha \in \mathbb{Z}^d \right\},$$

where the functions in $\boldsymbol{\Phi} := \{\varphi_1, \ldots, \varphi_r\}$ in $L^2(\mathbb{R}^d)$ are called a set of generators for $V^2_{\boldsymbol{\Phi}}$. Assuming that the sequence $\{\varphi_k(t-\alpha)\}_{\alpha \in \mathbb{Z}^d,\ k=1,2,\ldots,r}$ is a Riesz basis for $V^2_{\boldsymbol{\Phi}}$, the shift-invariant space $V^2_{\boldsymbol{\Phi}}$ can be described as

$$V^2_{\boldsymbol{\Phi}} = \left\{ \sum_{\alpha \in \mathbb{Z}^d} \sum_{k=1}^{r} d_k(\alpha)\ \varphi_k(t-\alpha) : d_k \in \ell^2(\mathbb{Z}^d), k = 1, 2, \ldots, r \right\}. \tag{3.2}$$

See [8, 9, 36] for the general theory of shift-invariant spaces and their applications. These spaces and the scaling functions $\boldsymbol{\Phi} = \{\varphi_1, \ldots, \varphi_r\}$ appear in the multiwavelet setting. Multiwavelets lead to multiresolution analyses and fast algorithms just as scalar wavelets, but they have some advantages: they can have short support coupled with high smoothness and high approximation order, and they can be both symmetric and orthogonal (see, for instance, [29]). Classical sampling in multiwavelet subspaces has been studied in [38, 43].

On the other hand, in many common situations, the available data are samples of some filtered versions $f * \mathsf{h}_j$ of the signal f itself, where the average function h_j reflects the characteristics of the acquisition device. This leads to generalized sampling (also called average sampling) in $V^2_{\boldsymbol{\Phi}}$ (see, among others, [1, 5, 14, 16, 17, 30, 34, 35, 40, 41, 43]).

Suppose that s convolution systems (linear time-invariant systems or filters in engineering jargon) $\mathscr{L}_j$, $j = 1, 2, \ldots, s$, are defined on the shift-invariant subspace $V^2_{\boldsymbol{\Phi}}$ of $L^2(\mathbb{R}^d)$. Assume also that the sequence of samples $\{(\mathscr{L}_j f)(M\alpha)\}_{\alpha \in \mathbb{Z}^d,\ j=1,2,\ldots,s}$ for f in $V^2_{\boldsymbol{\Phi}}$ is available, where the samples are taken at the sub-lattice $M\mathbb{Z}^d$ of $\mathbb{Z}^d$, where M denotes a matrix of integer entries with positive determinant. If we

sample any function $f \in V^2_{\Phi}$ on $M\mathbb{Z}^d$, we are using the sampling rate $1/r(\det M)$ and, roughly speaking, we will need, for the recovery of $f \in V^2_{\Phi}$, the sequence of generalized samples $\{(\mathscr{L}_j f)(M\alpha)\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ coming from $s \geq r(\det M)$ convolution systems $\mathscr{L}_j$.

Assume that the sequences of generalized samples satisfy the following stability condition: There exist two positive constants $0 < A \leq B$ such that

$$A\|f\|^2 \leq \sum_{j=1}^{s} \sum_{\alpha\in\mathbb{Z}^d} |\mathscr{L}_j f(M\alpha)|^2 \leq B\|f\|^2 \quad \text{for all } f \in V^2_{\Phi}.$$

In [5] the set of systems $\{\mathscr{L}_1, \mathscr{L}_2, \ldots, \mathscr{L}_s\}$ is said to be an M-stable filtering sampler for V^2_{Φ}. The aim of this work is to obtain sampling formulas in V^2_{Φ} having the form

$$f(t) = (\det M) \sum_{j=1}^{s} \sum_{\alpha\in\mathbb{Z}^d} (\mathscr{L}_j f)(M\alpha) S_j(t - M\alpha), \quad t \in \mathbb{R}^d, \tag{3.3}$$

such that the sequence of reconstruction functions $\{S_j(\cdot - M\alpha)\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a frame for the shift-invariant space V^2_{Φ}. This will be done in the light of the frame theory for separable Hilbert spaces, by using a similar mathematical technique as in the above section.

Recall that a sequence $\{x_n\}$ is a frame for a separable Hilbert space $\mathscr{H}$ if there exist two constants $A, B > 0$ (frame bounds) such that

$$A\|x\|^2 \leq \sum_n |\langle x, x_n\rangle|^2 \leq B\|x\|^2 \quad \text{for all } x \in \mathscr{H}.$$

Given a frame $\{x_n\}$ for $\mathscr{H}$ the representation property of any vector $x \in \mathscr{H}$ as a series $x = \sum_n c_n x_n$ is retained, but, unlike the case of Riesz bases, the uniqueness of this representation (for overcomplete frames) is sacrificed. Suitable frame coefficients c_n, depending linearly and continuously on x, are obtained by using the dual frames $\{y_n\}$ of $\{x_n\}$, i.e., the sequence $\{y_n\}$ is another frame for $\mathscr{H}$ such that, for each $x \in \mathscr{H}$, the expansions $x = \sum_n \langle x, y_n\rangle x_n = \sum_n \langle x, x_n\rangle y_n$ hold. For more details on the frame theory see the superb monograph [11] and the references therein.

3.2 Preliminaries on $L^2(\mathbb{R}^d)$ Shift-Invariant Subspaces

Let $\Phi := \{\varphi_1, \varphi_2, \ldots, \varphi_r\}$ be a set of functions, where $\varphi_k \in L^2(\mathbb{R}^d)$ $k = 1, 2, \ldots, r$, such that the sequence $\{\varphi_k(t - \alpha)\}_{\alpha\in\mathbb{Z}^d,\ k=1,2\ldots,r}$ is a Riesz basis for the shift-invariant space

$$V^2_{\Phi} := \left\{ \sum_{\alpha\in\mathbb{Z}^d} \sum_{k=1}^{r} d_k(\alpha)\ \varphi_k(t - \alpha) : d_k \in \ell^2(\mathbb{Z}^d), k = 1, 2 \ldots, r \right\} \subset L^2(\mathbb{R}^d).$$

There exists a necessary and sufficient condition involving the Gramian matrix function

$$G_{\boldsymbol{\Phi}}(w) := \sum_{\alpha\in\mathbb{Z}^d} \widehat{\boldsymbol{\Phi}}(w+\alpha)\overline{\widehat{\boldsymbol{\Phi}}(w+\alpha)}^{\top}, \text{ where } \widehat{\boldsymbol{\Phi}} := (\widehat{\varphi}_1, \widehat{\varphi}_2, \ldots, \widehat{\varphi}_r)^{\top},$$

which assures that the sequence $\{\varphi_k(\cdot-\alpha)\}_{\alpha\in\mathbb{Z}^d,\ k=1,2\ldots,r}$ is a Riesz basis for $V^2_{\boldsymbol{\Phi}}$, namely (see, for instance, [5]): There exist two positive constants c and C such that

$$c\,\mathbb{I}_r \le G_{\boldsymbol{\Phi}}(w) \le C\,\mathbb{I}_r \quad \text{a.e. } w\in[0,1)^d. \tag{3.4}$$

We assume throughout this chapter that the functions in the shift-invariant space $V^2_{\boldsymbol{\Phi}}$ are continuous on $\mathbb{R}^d$. As in the case of one generator, this is equivalent to the generators $\boldsymbol{\Phi}$ being continuous on $\mathbb{R}^d$ with $\sum_{\alpha\in\mathbb{Z}^d}|\boldsymbol{\Phi}(t-\alpha)|^2$ uniformly bounded on $\mathbb{R}^d$. Thus, any $f\in V^2_{\boldsymbol{\Phi}}$ is defined on $\mathbb{R}^d$ as the pointwise sum

$$f(t) = \sum_{k=1}^{r}\sum_{\alpha\in\mathbb{Z}^d} d_k(\alpha)\ \varphi_k(t-\alpha), \quad t\in\mathbb{R}^d. \tag{3.5}$$

Besides, the space $V^2_{\boldsymbol{\Phi}}$ is an RKHS since the evaluation functionals, $E_t f := f(t)$, are bounded on $V^2_{\boldsymbol{\Phi}}$. Indeed, for each fixed $t\in\mathbb{R}^d$, we have

$$\begin{aligned}
|f(t)|^2 &= \left|\sum_{\alpha\in\mathbb{Z}^d}\sum_{k=1}^{r} d_k(\alpha)\ \varphi_k(t-\alpha)\right|^2\\
&\le \left(\sum_{\alpha\in\mathbb{Z}^d}\sum_{k=1}^{r}|d_k(\alpha)|^2\right)\left(\sum_{\alpha\in\mathbb{Z}^d}\sum_{k=1}^{r}|\varphi_k(t-\alpha)|^2\right)\\
&= \left(\sum_{\alpha\in\mathbb{Z}^d}\sum_{k=1}^{r}|d_k(\alpha)|^2\right)\left(\sum_{\alpha\in\mathbb{Z}^d}|\boldsymbol{\Phi}(t-\alpha)|^2\right)\\
&\le \frac{\|f\|^2}{c}\sum_{\alpha\in\mathbb{Z}^d}|\boldsymbol{\Phi}(t-\alpha)|^2, \quad f\in V^2_{\boldsymbol{\Phi}},
\end{aligned}$$

where we have used Cauchy–Schwarz's inequality in (3.5), and the inequality satisfied for any lower Riesz bound c of the Riesz basis $\{\varphi_k(\cdot-\alpha)\}_{\alpha\in\mathbb{Z}^d,\ k=1,2\ldots,r}$ for $V^2_{\boldsymbol{\Phi}}$, i.e., $c\sum_{\alpha\in\mathbb{Z}^d}\sum_{k=1}^{r}|d_k(\alpha)|^2 \le \|f\|^2$.

Thus, the convergence in $V^2_{\boldsymbol{\Phi}}$ in the $L^2(\mathbb{R}^d)$ sense implies pointwise convergence which is uniform on $\mathbb{R}^d$.

The product space

$$L^2_r[0,1)^d := \left\{\mathbf{F} = (F_1, F_2, \ldots, F_r)^{\top} : F_k\in L^2[0,1)^d,\ k=1,2,\ldots,r\right\}$$

with its usual inner product

$$\langle \mathbf{F}, \mathbf{H} \rangle_{L_r^2[0,1)^d} := \sum_{k=1}^{r} \langle F_k, H_k \rangle_{L^2[0,1)^d} = \int_{[0,1)^d} \mathbf{H}^*(w)\mathbf{F}(w)\mathrm{d}w$$

becomes a Hilbert space. Similarly, we introduce the product Banach space $L_r^\infty[0,1)^d$.

The system $\left\{\mathrm{e}^{-2\pi i\alpha^\top w}\mathbf{e}_k\right\}_{\alpha\in\mathbb{Z}^d,\ k=1,2,\ldots,r}$, where $\mathbf{e}_k$ denotes the vector of $\mathbb{R}^r$ with all the components null except the kth component which is equal to one, is an orthonormal basis for $L_r^2[0,1)^d$.

The shift-invariant space $V_{\boldsymbol{\Phi}}^2$ is the image of $L_r^2[0,1)^d$ by means of the isomorphism

$$\mathscr{T}_{\boldsymbol{\Phi}} : L_r^2[0,1)^d \longrightarrow V_{\boldsymbol{\Phi}}^2$$

$$\left\{\mathrm{e}^{-2\pi i\alpha^\top w}\mathbf{e}_k\right\}_{\alpha\in\mathbb{Z}^d,\ k=1,2,\ldots,r} \longmapsto \left\{\varphi_k(t-\alpha)\right\}_{\alpha\in\mathbb{Z}^d,\ k=1,2,\ldots,r},$$

which maps the orthonormal basis $\left\{\mathrm{e}^{-2\pi i\alpha^\top w}\mathbf{e}_k\right\}_{\alpha\in\mathbb{Z}^d,\ k=1,2,\ldots,r}$ for $L_r^2[0,1)^d$ onto the Riesz basis $\left\{\varphi_k(t-\alpha)\right\}_{\alpha\in\mathbb{Z}^d,\ k=1,2,\ldots,r}$ for $V_{\boldsymbol{\Phi}}^2$. For each $\mathbf{F} = (F_1,\ldots,F_r)^\top \in L_r^2[0,1)^d$ we have

$$\mathscr{T}_{\boldsymbol{\Phi}}\mathbf{F}(t) := \sum_{\alpha\in\mathbb{Z}^d}\sum_{k=1}^{r} \left\langle F_k, \mathrm{e}^{-2\pi i\alpha^\top\cdot}\right\rangle_{L^2[0,1)^d} \varphi_k(t-\alpha), \quad t\in\mathbb{R}^d. \tag{3.6}$$

The isomorphism $\mathscr{T}_{\boldsymbol{\Phi}}$ can also be expressed by

$$f(t) = \mathscr{T}_{\boldsymbol{\Phi}}\mathbf{F}(t) = \langle \mathbf{F}, \mathbf{K}_t\rangle_{L_r^2[0,1)^d}, \quad t\in\mathbb{R}^d,$$

where the kernel transform $\mathbb{R}^d \ni t \mapsto \mathbf{K}_t \in L_r^2[0,1)^d$ is defined as $\mathbf{K}_t(x) := \overline{\mathbf{Z}\boldsymbol{\Phi}}(t,x)$, and $\mathbf{Z}\boldsymbol{\Phi}$ denotes the Zak transform of $\boldsymbol{\Phi}$, i.e.,

$$(\mathbf{Z}\boldsymbol{\Phi})(t,w) := \sum_{\alpha\in\mathbb{Z}^d} \boldsymbol{\Phi}(t+\alpha)\mathrm{e}^{-2\pi i\alpha^\top w}.$$

Note that $(\mathbf{Z}\boldsymbol{\Phi}) = (Z\varphi_1,\ldots,Z\varphi_r)^\top$ where Z denotes the usual Zak transform.

The following shifting property of $\mathscr{T}_{\boldsymbol{\Phi}}$ will be used later: For $\mathbf{F}\in L_r^2[0,1)^d$ and $\alpha\in\mathbb{Z}^d$, we have

$$\mathscr{T}_{\boldsymbol{\Phi}}\left[\mathbf{F}(\cdot)\mathrm{e}^{-2\pi i\alpha^\top\cdot}\right](t) = \mathscr{T}_{\boldsymbol{\Phi}}\mathbf{F}(t-\alpha), \quad t\in\mathbb{R}^d. \tag{3.7}$$

3.2.1 The Convolution Systems $\mathscr{L}_j$ on $V_{\boldsymbol{\Phi}}^2$

We consider s convolution systems $\mathscr{L}_j f = f * \mathsf{h}_j$, $j = 1,2,\ldots,s$, defined for $f\in V_{\boldsymbol{\Phi}}^2$ where each impulse response h_j belongs to one of the following three types:

(a) The impulse response h_j is a linear combination of partial derivatives of shifted delta functionals, i.e.,

$$\big(\mathscr{L}_j f\big)(t) := \sum_{|\beta|\le N_j} c_{j,\beta} D^\beta f(t+d_{j,\beta}), \quad t\in\mathbb{R}^d.$$

If there is a system of this type, we also assume that $\sum_{\alpha\in\mathbb{Z}^d}|D^\beta\varphi(t-\alpha)|^2$ is uniformly bounded on $\mathbb{R}^d$ for $|\beta|\le N_j$.

(b) The impulse response h_j of $\mathscr{L}_j$ belongs to $L^2(\mathbb{R}^d)$. Thus, for any $f\in V^2_\varphi$, we have

$$\big(\mathscr{L}_j f\big)(t) := [f*\mathsf{h}_j](t) = \int_{\mathbb{R}^d} f(x)\mathsf{h}_j(t-x)\mathrm{d}x, \quad t\in\mathbb{R}^d.$$

(c) The function $\widehat{\mathsf{h}}_j\in L^\infty(\mathbb{R}^d)$ whenever $H_{\varphi_k}(x) := \sum_{\alpha\in\mathbb{Z}^d}|\widehat{\varphi}_k(x+\alpha)|\in L^2[0,1)^d$ for all $k=1,2,\dots,r$.

Lemma 1. *Let $\mathscr{L}$ be a convolution system of the type (b) or (c). Then, for each fixed $t\in\mathbb{R}^d$ the sequence $\{(\mathscr{L}\varphi_k)(t+\alpha)\}_{\alpha\in\mathbb{Z}^d}$ belongs to $\ell^2(\mathbb{Z}^d)$ for each $k=1,\dots,r$.*

Proof. First assume that $\mathsf{h}\in L^2(\mathbb{R}^d)$; then, we have

$$\sum_{\alpha\in\mathbb{Z}^d}|\mathscr{L}\varphi_k(t+\alpha)|^2 = \left\|\sum_{\alpha\in\mathbb{Z}^d}\mathscr{L}\varphi_k(t+\alpha)\mathrm{e}^{-2\pi i\alpha^\top x}\right\|^2_{L^2[0,1)^d} = \|Z\mathscr{L}\varphi_k(t,x)\|^2_{L^2[0,1)^d}$$
$$= \left\|\sum_{\alpha\in\mathbb{Z}^d}\big(\widehat{\mathscr{L}\varphi_k}\big)(x+\alpha)\mathrm{e}^{2\pi i(x+\alpha)^\top t}\right\|^2_{L^2[0,1)^d},$$

where, in the last equality, we have used a version of the Poisson summation formula [20, Lemma 2.1]. Notice that $\widehat{\varphi}_k,\widehat{\mathsf{h}}\in L^2(\mathbb{R}^d)$ implies, by Cauchy–Schwarz's inequality, that $\widehat{\varphi}_k\widehat{\mathsf{h}} = \widehat{\mathscr{L}\varphi_k}\in L^1(\mathbb{R}^d)$. Now,

$$\left\|\sum_{\alpha\in\mathbb{Z}^d}\left(\widehat{\mathscr{L}\varphi_k}\right)(x+\alpha)\mathrm{e}^{2\pi i(x+\alpha)^\top t}\right\|^2_{L^2[0,1)^d}$$
$$= \left\|\sum_{\alpha\in\mathbb{Z}^d}\widehat{\varphi}_k(x+\alpha)\widehat{\mathsf{h}}(x+\alpha)\mathrm{e}^{2\pi i(x+\alpha)^\top t}\right\|^2_{L^2[0,1)^d}$$
$$\le \left\|\left(\sum_{\alpha\in\mathbb{Z}^d}|\widehat{\varphi}_k(x+\alpha)|^2\right)^{1/2}\left(\sum_{\alpha\in\mathbb{Z}^d}|\widehat{\mathsf{h}}(x+\alpha)|^2\right)^{1/2}\right\|^2_{L^2[0,1)^d} \le C^{1/2}\|\mathsf{h}\|^2_{L^2[0,1)^d},$$

where we have used (3.4) and the fact that $\|\mathsf{h}\|^2_{L^2(\mathbb{R}^d)} = \|\sum_{\alpha\in\mathbb{Z}^d} |\widehat{\mathsf{h}}(x+\alpha)|^2\|_{L^1[0,1)^d}$. Finally, assume that $H_{\varphi_k} \in L^2[0,1)^d$; since $\widehat{\varphi}_k \in L^1(\mathbb{R}^d)\cap L^2(\mathbb{R}^d)$, we obtain that $\widehat{\mathscr{L}\varphi_k} = \widehat{\varphi}_k\widehat{\mathsf{h}} \in L^1(\mathbb{R}^d)\cap L^2(\mathbb{R}^d)$. Since $\sum_{\alpha\in\mathbb{Z}^d} |\widehat{\mathscr{L}\varphi_k}(x+\alpha)| \le \|\widehat{\mathsf{h}}\|_{L^\infty(\mathbb{R}^d)} H_{\varphi_k}(x)$, using again [20, Lemma 2.1], we get

$$\sum_{\alpha\in\mathbb{Z}^d} |\mathscr{L}\varphi_k(t+\alpha)|^2 = \left\| \sum_{\alpha\in\mathbb{Z}^d} \left(\widehat{\mathscr{L}\varphi_k}\right)(x+\alpha) \mathrm{e}^{2\pi i(x+\alpha)^\top t} \right\|^2_{L^2[0,1)^d}$$

$$\le \left\| \sum_{\alpha\in\mathbb{Z}^d} |\widehat{\mathscr{L}\varphi_k}(x+\alpha)| \right\|^2_{L^2[0,1)^d} \le \left\|\widehat{\mathsf{h}}\right\|^2_{L^\infty(\mathbb{R}^d)} \left\|H_{\varphi_k}\right\|^2_{L^2[0,1)^d}.$$

□

Lemma 2. *Let $\mathscr{L}$ be a convolution system of the type (a), (b), or (c). Then, for each $f \in V^2_{\boldsymbol{\Phi}}$, we have*

$$(\mathscr{L}f)(t) = \left\langle \mathbf{F}, \left(\overline{\mathbf{Z}\mathscr{L}\boldsymbol{\Phi}}\right)(t,\cdot)\right\rangle_{L^2_r[0,1)^d}, \quad \textit{where} \quad \mathbf{F} = \mathscr{T}^{-1}_{\boldsymbol{\Phi}} f.$$

Proof. Assume that $\mathscr{L}$ is a convolution system of type (a). Under our hypothesis on $\mathscr{L}$, for $m = 0,1,2\ldots,N$, we have that

$$f^{(m)}(t) = \sum_{\alpha\in\mathbb{Z}^d} \sum_{k=1}^r \left\langle F_k, \mathrm{e}^{-2\pi i\alpha^\top \cdot}\right\rangle \varphi_k^{(m)}(t-\alpha).$$

Having in mind we have assumed that $\sum_{\alpha\in\mathbb{Z}^d} |\boldsymbol{\Phi}^{(m)}(t-\alpha)|^2$ is uniformly bounded on $\mathbb{R}^d$, we obtain that

$$(\mathscr{L}f)(t) = \sum_{m=0}^N c_m f^{(m)}(t+d_m) = \sum_{m=0}^N c_m \sum_{\alpha\in\mathbb{Z}^d} \sum_{k=1}^r \left\langle F_k, \mathrm{e}^{-2\pi i\alpha^\top \cdot}\right\rangle \varphi_k^{(m)}(t+d_m-\alpha)$$

$$= \sum_{k=1}^r \left\langle F_k, \sum_{m=0}^N \overline{c_m} \sum_{\alpha\in\mathbb{Z}^d} \overline{\varphi_k}^{(m)}(t+d_m-\alpha)\mathrm{e}^{-2\pi i\alpha^\top \cdot}\right\rangle_{L^2[0,1)^d}$$

$$= \sum_{k=1}^r \left\langle F_k, \sum_{\alpha\in\mathbb{Z}^d} \overline{\mathscr{L}\varphi_k}(t-\alpha)\mathrm{e}^{-2\pi i\alpha^\top \cdot}\right\rangle = \sum_{k=1}^r \left\langle F_k, (\overline{Z\mathscr{L}\varphi_k})(t,\cdot)\right\rangle_{L^2[0,1)^d}.$$

Assume now that $\mathscr{L}$ is a convolution system of the type (b) or (c). For each $t\in\mathbb{R}^d$, considering the function $\psi(x) := \overline{\mathsf{h}(-x)}$, we have

$$(\mathscr{L}f)(t) = \langle f, \psi(\cdot - t)\rangle_{L^2(\mathbb{R}^d)} = \left\langle \sum_{\alpha\in\mathbb{Z}^d} \sum_{k=1}^r \left\langle F_k, \mathrm{e}^{-2\pi i\alpha^\top \cdot}\right\rangle \varphi_k(\cdot-\alpha), \psi(\cdot-t)\right\rangle_{L^2(\mathbb{R}^d)}$$

$$= \sum_{\alpha\in\mathbb{Z}^d}\sum_{k=1}^{r}\left\langle F_k, \mathrm{e}^{-2\pi i\alpha^\top\cdot}\right\rangle_{L^2[0,1)^d}\langle\varphi_k,\psi(\cdot-t+\alpha)\rangle_{L^2(\mathbb{R}^d)}$$

$$= \sum_{\alpha\in\mathbb{Z}^d}\sum_{k=1}^{r}\left\langle F_k, \mathrm{e}^{-2\pi i\alpha^\top\cdot}\right\rangle_{L^2[0,1)^d}\mathscr{L}\varphi_k(t-\alpha)\,.$$

Since the sequence $\{(\mathscr{L}\varphi_k)(t+\alpha)\}_{\alpha\in\mathbb{Z}^d}\in\ell^2(\mathbb{Z}^d)$, Parseval's equality gives

$$\begin{aligned}(\mathscr{L}f)(t) &= \sum_{k=1}^{r}\left\langle F_k, \sum_{\alpha\in\mathbb{Z}^d}\overline{\mathscr{L}\varphi_k}(t-\alpha)\,\mathrm{e}^{-2\pi i\alpha^\top\cdot}\right\rangle_{L^2[0,1)^d}\\ &= \left\langle\mathbf{F},(\overline{\mathbf{Z}\mathscr{L}\boldsymbol{\Phi}})(t,\cdot)\right\rangle_{L^2_r(0,1)}\,,\end{aligned}$$

which ends the proof. □

3.2.2 *Sampling at a Lattice of $\mathbb{Z}^d$: An Expression for the Samples*

Given a nonsingular matrix M with integer entries, we consider the lattice in $\mathbb{Z}^d$ generated by M, i.e.,

$$\Lambda_M := \{M\alpha : \alpha\in\mathbb{Z}^d\}\subset\mathbb{Z}^d\,.$$

Without loss of generality, we can assume that $\det M>0$; otherwise, we can consider $M'=ME$ where E is some $d\times d$ integer matrix satisfying $\det E=-1$. Trivially, $\Lambda_M=\Lambda_M'$. We denote by $M^\top$ and $M^{-\top}$ the transpose matrices of M and M^{-1}, respectively. The following useful generalized orthogonal relationship holds (see [46]):

$$\sum_{p\in\mathscr{N}(M^\top)}\mathrm{e}^{-2\pi i\alpha^\top M^{-T}p} = \begin{cases}\det M, & \alpha\in\Lambda_M\\ 0 & \alpha\in\mathbb{Z}^d\setminus\Lambda_M\end{cases} \tag{3.8}$$

where

$$\mathscr{N}(M^\top) := \mathbb{Z}^d\cap\{M^\top x : x\in[0,1)^d\}. \tag{3.9}$$

The set $\mathscr{N}(M^\top)$ has $\det M$ elements (see [46] or [48]). One of these elements is zero, say $i_1=0$; we denote the rest of elements by $i_2,\ldots,i_{\det M}$ ordered in any form; from now on, $\mathscr{N}(M^\top)=\{i_1=0,i_2,\ldots,i_{\det M}\}\subset\mathbb{Z}^d$.

Note that the sets, defined as $Q_l := M^{-\top}i_l + M^{-\top}[0,1)^d$, $l=1,2,\ldots,\det M$, satisfy (see [48, p 110])

$$Q_l\cap Q_{l'}=\emptyset \text{ if } l\neq l' \quad\text{and}\quad \mathrm{Vol}\left(\bigcup_{l=1}^{\det M}Q_l\right)=1\,.$$

Thus, $\int_{[0,1)^d} F(x)\mathrm{d}x = \sum_{l=1}^{\det M} \int_{Q_l} F(x)\mathrm{d}x$, for any function F integrable in $[0,1)^d$ and $\mathbb{Z}^d$-periodic. See also [39] and references therein for an abstract version of sampling in lattice invariant subspaces.

Now, assume that we sample the filtered versions $\mathscr{L}_j f$ of $f \in V^2_{\boldsymbol{\Phi}}$, $j = 1,2,\ldots,s$, at a lattice Λ_M. Having in mind Lemma 2, for $j = 1,2,\ldots,s$ and $\alpha \in \mathbb{Z}^d$, we obtain that

$$\big(\mathscr{L}_j f\big)(M\alpha) = \langle \mathbf{F}, \overline{\mathbf{Z}\mathscr{L}_j\boldsymbol{\Phi}}(M\alpha,\cdot)\rangle = \Big\langle \mathbf{F}, \overline{\mathbf{Z}\mathscr{L}_j\boldsymbol{\Phi}}(0,\cdot)\mathrm{e}^{-2\pi i\alpha^\top M^\top \cdot}\Big\rangle_{L^2_r[0,1)^d}, \quad (3.10)$$

where $\mathbf{F} = \mathscr{T}_{\boldsymbol{\Phi}}^{-1} f \in L^2_r[0,1)^d$. Denote

$$\mathbf{g}_j(x) := \mathbf{Z}\mathscr{L}_j\boldsymbol{\Phi}(0,x), \quad j = 1,2,\ldots,s; \quad (3.11)$$

in other words, $\mathbf{g}_j^\top(x) := \big(g_{j,1}(x), g_{j,2}(x), \ldots, g_{j,r}(x)\big)$, where $g_{j,k}(x) = Z\mathscr{L}_j\varphi_k(0,x)$ for $1 \le j \le s$ and $1 \le k \le r$.

As a consequence of expression (3.10) for generalized samples, a challenging problem is to study the completeness, Bessel, frame, or Riesz basis properties of any sequence $\big\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\big\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ in $L^2_r[0,1)^d$. To this end we introduce the $s \times r(\det M)$ matrix of functions

$$\mathbb{G}(x) := \begin{bmatrix} \mathbf{g}_1^\top(x) & \mathbf{g}_1^\top(x+M^{-\top}i_2) & \cdots & \mathbf{g}_1^\top(x+M^{-\top}i_{\det M}) \\ \mathbf{g}_2^\top(x) & \mathbf{g}_2^\top(x+M^{-\top}i_2) & \cdots & \mathbf{g}_2^\top(x+M^{-\top}i_{\det M}) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{g}_s^\top(x) & \mathbf{g}_s^\top(x+M^{-\top}i_2) & \cdots & \mathbf{g}_s^\top(x+M^{-\top}i_{\det M}) \end{bmatrix} \quad (3.12)$$

and its related constants

$$A_{\mathbb{G}} := \operatorname*{ess\,inf}_{x\in[0,1)^d} \lambda_{\min}[\mathbb{G}^*(x)\mathbb{G}(x)], \quad B_{\mathbb{G}} := \operatorname*{ess\,sup}_{x\in[0,1)^d} \lambda_{\max}[\mathbb{G}^*(x)\mathbb{G}(x)],$$

where $\mathbb{G}^*(x)$ denotes the transpose conjugate of the matrix $\mathbb{G}(x)$ and $\lambda_{\min}$ (respectively $\lambda_{\max}$), the smallest (respectively the largest) eigenvalue of the positive semidefinite matrix $\mathbb{G}^*(x)\mathbb{G}(x)$. Observe that $0 \le A_{\mathbb{G}} \le B_{\mathbb{G}} \le \infty$. Note that in the definition of the matrix $\mathbb{G}(x)$ we are considering the $\mathbb{Z}^d$-periodic extension of the involved functions $\mathbf{g}_j$, $j = 1,2,\ldots,s$.

We now present a general result valid for functions $\mathbf{g}_j$ in $L^2_r[0,1)^d$, $j = 1,2,\ldots,s$, even if they are not given by (3.11).

Lemma 3. *Let $\mathbf{g}_j$ be in $L^2_r[0,1)^d$ for $j = 1,2,\ldots,s$ and let $\mathbb{G}(x)$ be its associated matrix as in (3.12). Then:*

(a) The sequence $\big\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\big\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a complete system for $L^2_r[0,1)^d$ if and only if the rank of the matrix $\mathbb{G}(x)$ is $r(\det M)$ a.e. in $[0,1)^d$.

(b) The sequence $\left\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ *is a Bessel sequence for* $L^2_r[0,1)^d$ *if and only if* $\mathbf{g}_j\in L^\infty_r[0,1)^d$ *(or equivalently* $B_{\mathbb{G}}<\infty$*). In this case, the optimal Bessel bound is* $B_{\mathbb{G}}/(\det M)$.

(c) The sequence $\left\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ *is a frame for* $L^2_r[0,1)^d$ *if and only if* $0<A_{\mathbb{G}}\le B_{\mathbb{G}}<\infty$. *In this case, the optimal frame bounds are* $A_{\mathbb{G}}/(\det M)$ *and* $B_{\mathbb{G}}/(\det M)$.

(d) The sequence $\left\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ *is a Riesz basis for* $L^2_r[0,1)^d$ *if and only if it is a frame and* $s=r(\det M)$.

Proof. For any $\mathbf{F}\in L^2_r[0,1)^d$, we have

$$\begin{aligned}
&\left\langle \mathbf{F}(x),\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\rangle_{L^2_r[0,1)^d}\\
&\quad=\int_{[0,1)^d}\sum_{k=1}^{r}F_k(x)g_{j,k}(x)\mathrm{e}^{2\pi i\alpha^\top M^\top x}\,\mathrm{d}x\\
&\quad=\sum_{k=1}^{r}\sum_{l=1}^{\det M}\int_{Q_l}F_k(x)g_{j,k}(x)\mathrm{e}^{2\pi i\alpha^\top M^\top x}\,\mathrm{d}x\\
&\quad=\sum_{k=1}^{r}\int_{M^{-\top}[0,1)^d}\sum_{l=1}^{\det M}F_k\left(x+M^{-\top}i_l\right)g_{j,k}\left(x+M^{-\top}i_l\right)\,\mathrm{e}^{2\pi i\alpha^\top M^\top x}\,\mathrm{d}x\\
&\quad=\int_{M^{-\top}[0,1)^d}\sum_{k=1}^{r}\sum_{l=1}^{\det M}F_k\left(x+M^{-\top}i_l\right)g_{j,k}\left(x+M^{-\top}i_l\right)\,\mathrm{e}^{2\pi i\alpha^\top M^\top x}\,\mathrm{d}x\\
&\quad=\int_{M^{-\top}[0,1)^d}\sum_{l=1}^{\det M}\mathbf{g}_j^\top\left(x+M^{-\top}i_l\right)\mathbf{F}\left(x+M^{-\top}i_l\right)\,\mathrm{e}^{2\pi i\alpha^\top M^\top x}\,\mathrm{d}x,
\end{aligned}\tag{3.13}$$

where we have considered the $\mathbb{Z}^d$-periodic extension of $\mathbf{F}$. By using that the sequence $\{\mathrm{e}^{2\pi i\alpha^\top M^\top x}\}_{\alpha\in\mathbb{Z}^d}$ is an orthogonal basis for $L^2(M^{-\top}[0,1)^d)$ we obtain

$$\begin{aligned}
&\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}\left|\left\langle\mathbf{F}(x),\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\rangle_{L^2_r[0,1)^d}\right|^2\\
&\qquad=\frac{1}{\det M}\sum_{j=1}^{s}\left\|\sum_{l=1}^{\det M}\mathbf{g}_j^\top(x+M^{-\top}i_l)\mathbf{F}(x+M^{-\top}i_l)\right\|^2_{L^2_r(M^{-\top}[0,1)^d)}.
\end{aligned}$$

Denoting $\mathbb{F}(x):=\left[\mathbf{F}^\top(x),\mathbf{F}^\top(x+M^{-\top}i_2),\ldots,\mathbf{F}^\top(x+M^{-\top}i_{\det M})\right]^\top$, the equality above reads

$$\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}\left|\left\langle \mathbf{F}(x),\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\rangle_{L^2_r[0,1)^d}\right|^2=\frac{1}{\det M}\left\|\mathbb{G}(x)\mathbb{F}(x)\right\|^2_{L^2_s(M^{-\top}[0,1)^d)}. \tag{3.14}$$

On the other hand, using that the function $\mathbf{g}_j$ is $\mathbb{Z}^d$-periodic, we obtain that the set $\left\{\mathbf{g}_j(x+M^{-\top}i_l+M^{-\top}i_1),\mathbf{g}_j(x+M^{-\top}i_l+M^{-\top}i_2),\ldots,\mathbf{g}_j(x+M^{-\top}i_l+M^{-\top}i_{\det M})\right\}$ has the same elements as $\left\{\mathbf{g}_j(x+M^{-\top}i_1),\mathbf{g}_j(x+M^{-\top}i_2),\ldots,\mathbf{g}_j(x+M^{-\top}i_{\det M})\right\}$. Thus, the matrix $\mathbb{G}(x+M^{-\top}i_l)$ has the same columns of $\mathbb{G}(x)$, possibly in a different order. Hence, $\operatorname{rank}\mathbb{G}(x)=r(\det M)$ a.e. in $[0,1)^d$ if and only if $\operatorname{rank}\mathbb{G}(x)=r(\det M)$ a.e. in $M^{-\top}[0,1)^d$. Moreover,

$$A_{\mathbb{G}}=\operatorname*{ess\,inf}_{x\in M^{-\top}[0,1)^d}\lambda_{\min}[\mathbb{G}^*(x)\mathbb{G}(x)],\quad B_{\mathbb{G}}=\operatorname*{ess\,sup}_{x\in M^{-\top}[0,1)^d}\lambda_{\max}[\mathbb{G}^*(x)\mathbb{G}(x)]. \tag{3.15}$$

To prove (a), assume that there exists a set $\Omega\subseteq M^{-\top}[0,1)^d$ with positive measure such that $\operatorname{rank}\mathbb{G}(x)<r(\det M)$ for each $x\in\Omega$. Then, there exists a measurable function $v(x)$, $x\in\Omega$, such that $\mathbb{G}(x)v(x)=0$ and $\|v(x)\|_{L^2_{r(\det M)}(M^{-\top}[0,1)^d)}=1$ in Ω. This function can be constructed as in [28, Lemma 2.4]. Define $\mathbf{F}\in L^2_r[0,1)^d$ such that $\mathbb{F}(x)=v(x)$ if $x\in\Omega$ and $\mathbb{F}(x)=0$ if $x\in M^{-\top}[0,1)^d\setminus\Omega$. Hence, from (3.14), we obtain that the system is not complete. Conversely, if the system is not complete, by using (3.14), we obtain an $\mathbb{F}(x)$ different from 0 in a set with positive measure such that $\mathbb{G}(x)\mathbb{F}(x)=0$. Thus, $\operatorname{rank}\mathbb{G}(x)<r(\det M)$ on a set with positive measure. To prove (b) notice that

$$\begin{aligned}\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}\left|\left\langle \mathbf{F}(x),\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\rangle_{L^2_r[0,1)^d}\right|^2&=\frac{1}{\det M}\left\|\mathbb{G}(x)\mathbb{F}(x)\right\|^2_{L^2_s(M^{-\top}[0,1)^d)}\\&=\frac{1}{\det M}\int_{M^{-\top}[0,1)^d}\mathbb{F}^*(x)\mathbb{G}^*(x)\mathbb{G}(x)\mathbb{F}(x)\mathrm{d}x.\end{aligned} \tag{3.16}$$

If $B_{\mathbb{G}}<\infty$, then, for each $\mathbb{F}$, we have

$$\begin{aligned}\frac{1}{\det M}\int_{M^{-\top}[0,1)^d}\mathbb{F}^*(x)\mathbb{G}^*(x)\mathbb{G}(x)\mathbb{F}(x)\mathrm{d}x&\le\frac{B_{\mathbb{G}}}{\det M}\left\|\mathbb{F}\right\|^2_{L^2_{r(\det M)}(M^{-\top}[0,1)^d)}\\&=\frac{B_{\mathbb{G}}}{\det M}\left\|\mathbf{F}\right\|^2_{L^2_r[0,1)^d},\end{aligned} \tag{3.17}$$

from which the sequence $\left\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a Bessel sequence and its optimal Bessel bound is less than or equal to $B_{\mathbb{G}}/(\det M)$.

Let $K<B_{\mathbf{G}}$; there exists a set $\Omega_K\subset M^{-\top}[0,1)^d$ with positive measure such that $\lambda_{\max_{x\in\Omega_K}}[\mathbb{G}^*(x)\mathbb{G}(x)]\ge K$. Let $\mathbf{F}\in L^2_r[0,1)^d$ such that its associated vector function

$\mathbb{F}$ is 0 if $x \in M^{-\top}[0,1)^d \setminus \Omega_K$ and $\mathbb{F}$ is an eigenvector of norm 1 associated with the largest eigenvalue of $\mathbb{G}^*(x)\mathbb{G}(x)$ if $x \in \Omega_K$. Using (3.16), we obtain

$$\sum_{j=1}^{s} \sum_{\alpha \in \mathbb{Z}^d} \left| \left\langle \mathbf{F}(x), \overline{\mathbf{g}_j(x)} \mathrm{e}^{-2\pi i \alpha^\top M^\top x} \right\rangle_{L^2_r[0,1)^d} \right|^2 \geq \frac{K}{\det M} \|\mathbf{F}\|^2_{L^2_r[0,1)^d}.$$

Therefore, if $B_{\mathbf{G}} = \infty$, the sequence $\left\{\overline{\mathbf{g}_j(x)} \mathrm{e}^{-2\pi i \alpha^\top M^\top x}\right\}_{\alpha \in \mathbb{Z}^d,\ j=1,2,\ldots,s}$ is not a Bessel sequence, and the optimal Bessel bound is $B_{\mathbb{G}}/(\det M)$.

To prove (c) assume first that $0 < A_{\mathbb{G}} \leq B_{\mathbb{G}} < \infty$. By using part (b), the sequence $\left\{\overline{\mathbf{g}_j(x)} \mathrm{e}^{-2\pi i \alpha^\top M^\top x}\right\}_{\alpha \in \mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a Bessel sequence in $L^2_r[0,1)^d$. Moreover, using (3.16) and the Rayleigh–Ritz theorem (see [26, p 176]), for each $\mathbf{F} \in L^2_r[0,1)^d$, we obtain

$$\begin{aligned} \sum_{j=1}^{s} \sum_{\alpha \in \mathbb{Z}^d} \left| \left\langle \mathbf{F}(x), \overline{\mathbf{g}_j(x)} \mathrm{e}^{-2\pi i \alpha^\top M^\top x} \right\rangle_{L^2_r[0,1)^d} \right|^2 &\geq \frac{A_{\mathbb{G}}}{\det M} \|\mathbb{F}\|^2_{L^2_{r(\det M)}(M^{-\top}[0,1)^d)} \\ &= \frac{A_{\mathbb{G}}}{\det M} \|\mathbf{F}\|^2_{L^2_r[0,1)^d}. \end{aligned} \tag{3.18}$$

Hence, the sequence $\left\{\overline{\mathbf{g}_j(x)} \mathrm{e}^{-2\pi i \alpha^\top M^\top x}\right\}_{\alpha \in \mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a frame with optimal lower bound larger than or equal to $A_{\mathbb{G}}/(\det M)$.

Conversely, if $\left\{\overline{\mathbf{g}_j(x)} \mathrm{e}^{-2\pi i \alpha^\top M^\top x}\right\}_{\alpha \in \mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a frame for $L^2_r[0,1)^d$, we know by part (b) that $B_{\mathbb{G}} < \infty$. In order to prove that $A_{\mathbb{G}} > 0$, consider any constant $K > A_{\mathbb{G}}$. Then, there exists a set $\Omega_K \subset M^{-\top}[0,1)^d$ with positive measure such that $\lambda_{\min_{x \in \Omega_K}}[\mathbb{G}^*(x)\mathbb{G}(x)] \leq K$. Let $\mathbf{F} \in L^2_r[0,1)^d$ such that its associated $\mathbb{F}(x)$ is 0 if $x \in M^{-\top}[0,1)^d \setminus \Omega_K$ and $\mathbb{F}(x)$ is an eigenvector of norm 1 associated with the smallest eigenvalue of $\mathbb{G}^*(x)\mathbb{G}(x)$ if $x \in \Omega_K$. Since $\mathbb{F}$ is bounded, we have that $\mathbb{G}(x)\mathbb{F}(x) \in L^2_s(M^{-\top}[0,1)^d)$. From (3.16) we get

$$\begin{aligned} \sum_{j=1}^{s} \sum_{\alpha \in \mathbb{Z}^d} \left| \left\langle \mathbf{F}(x), \overline{\mathbf{g}_j(x)} \mathrm{e}^{-2\pi i \alpha^\top M^\top x} \right\rangle_{L^2_r[0,1)^d} \right|^2 &\leq \frac{K}{\det M} \|\mathbb{F}\|^2_{L^2_{r(\det M)}(M^{-\top}[0,1)^d)} \\ &= \frac{K}{\det M} \|\mathbf{F}\|^2_{L^2_r[0,1)^d}. \end{aligned} \tag{3.19}$$

Denoting by A the optimal lower frame bound of $\left\{\overline{\mathbf{g}_j(x)} \mathrm{e}^{-2\pi i \alpha^\top M^\top x}\right\}_{\alpha \in \mathbb{Z}^d,\ j=1,2,\ldots,s}$, we have obtained that $K/(\det M) \geq A$ for each $K > A_{\mathbb{G}}$; thus, $A_{\mathbb{G}}/(\det M) \geq A$ and, consequently, $A_{\mathbb{G}} > 0$. Moreover, under the hypotheses of part (c), we deduce that $A_{\mathbb{G}}/(\det M)$ and $B_{\mathbb{G}}/(\det M)$ are the optimal frame bounds.

The proof of (d) is based on the following result ([11, Theorem 6.1.1]): A frame is a Riesz basis if and only if it has a biorthogonal sequence. Assume that the sequence $\left\{\overline{\mathbf{g}_j(x)} \mathrm{e}^{-2\pi i \alpha^\top M^\top x}\right\}_{\alpha \in \mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a Riesz basis for $L^2_r[0,1)^d$ being the sequence $\{\mathbf{h}_{j,\alpha}\}_{\alpha \in \mathbb{Z}^d,\ j=1,2,\ldots,s}$ its biorthogonal sequence. Using (3.13) we get

$$\int_{M^{-\top}[0,1)^d} \sum_{l=1}^{\det M} \mathbf{g}_j^{\top}\left(x+M^{-\top}i_l\right)\mathbf{h}_{j',0}\left(x+M^{-\top}i_l\right)\,\mathrm{e}^{2\pi i\alpha^{\top}M^{\top}x}\,\mathrm{d}x$$
$$= \left\langle \mathbf{h}_{j',0}(\cdot), \overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}\cdot}\right\rangle = \delta_{j,j'}\delta_{\alpha,0}\,.$$

Therefore,

$$\sum_{l=1}^{\det M} \mathbf{g}_j^{\top}\left(x+M^{-\top}i_l\right)\mathbf{h}_{j',0}\left(x+M^{-\top}i_l\right)\,\mathrm{e}^{2\pi i\alpha^{\top}M^{\top}x}$$
$$= (\det M)\delta_{j,j'} \quad \text{a.e. in } M^{-\top}[0,1)^d\,.$$

Thus, the matrix $\mathbb{G}(x)$ has a right inverse a.e. in $M^{-\top}[0,1)^d$ and, in particular, $s \le r(\det M)$. On the other hand, $A_{\mathbb{G}} > 0$ implies that $\det[\mathbb{G}^*(x)\mathbb{G}(x)] > 0$, a.e. in $M^{-\top}[0,1)^d$, and there exists the matrix $[\mathbb{G}^*(x)\mathbb{G}(x)]^{-1}\mathbb{G}^*(x)$ a.e. in $M^{-\top}[0,1)^d$. This matrix is a left inverse of the matrix $\mathbb{G}(x)$ which implies $s \ge r(\det M)$. Thus, we obtain that $r(\det M) = s$.

Conversely, assume that $\left\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a frame for $L^2_r[0,1)^d$ and $r(\det M) = s$. In this case, $\mathbb{G}(x)$ is a square matrix, and $\det[\mathbb{G}(x)^*(x)\mathbf{G}(x)(x)] > 0$ a.e. in $M^{-\top}[0,1)^d$ implies that $\det\mathbb{G}(x) \neq 0$ a.e. in $M^{-\top}[0,1)^d$. Having in mind the structure of $\mathbb{G}(x)$ its inverse must be the $r(\det M)\times s$ matrix

$$\mathbb{G}^{-1}(x) = \begin{bmatrix} \mathbf{c}_1(x) & \ldots & \mathbf{c}_s(x) \\ \mathbf{c}_1(x+M^{-\top}i_2) & \ldots & \mathbf{c}_s(x+M^{-\top}i_2) \\ \vdots & & \vdots \\ \mathbf{c}_1(x+M^{-\top}i_{\det M}) & \ldots & \mathbf{c}_s(x+M^{-\top}i_{\det M}) \end{bmatrix},$$

where, for each $j = 1,2,\ldots,s$, the function $\mathbf{c}_j \in L^2_r[0,1)^d$.

It is easy to verify that the sequence $\left\{(\det M)\mathbf{c}_j(x)\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a biorthogonal sequence of $\left\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$, and therefore it is a Riesz basis for $L^2_r[0,1)^d$. □

3.3 Generalized Regular Sampling in V^2_{Φ}

In this section we prove that expression (3.10) allows us to obtain $\mathbf{F} = \mathscr{T}_{\Phi}^{-1}f$ from the generalized samples $\{\mathscr{L}_j f(M\alpha)\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$; as a consequence, applying the isomorphism $\mathscr{T}_{\Phi}$, we recover the function f in V^2_{Φ}.

Assume that the functions $\mathbf{g}_j$ given in (3.11) belong to $\in L^\infty_r[0,1)^d$ for $j = 1,2,\ldots,s$; thus, $\mathbf{g}_j^\top(x)\mathbf{F}(x) \in L^2[0,1)^d$. Having in mind (3.8) and the expression (3.10) for the generalized samples, we have that

$$
\begin{aligned}
&(\det M)\sum_{\alpha\in\mathbb{Z}^d}\big(\mathscr{L}_j f\big)(M\alpha)\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\\
&\quad=\sum_{\alpha\in\mathbb{Z}^d}\big(\mathscr{L}_j f\big)(\alpha)\mathrm{e}^{-2\pi i\alpha^\top x}\sum_{p\in\mathscr{N}(M^\top)}\mathrm{e}^{-2\pi i\alpha^\top M^{-\top}p}\\
&\quad=\sum_{p\in\mathscr{N}(M^\top)}\sum_{\alpha\in\mathbb{Z}^d}\big(\mathscr{L}_j f\big)(\alpha)\mathrm{e}^{-2\pi i\alpha^\top(x+M^{-\top}p)}\\
&\quad=\sum_{p\in\mathscr{N}(M^\top)}\sum_{\alpha\in\mathbb{Z}^d}\left\langle\mathbf{F},\overline{\mathbf{g}_j(\cdot)}\mathrm{e}^{-2\pi i\alpha^\top M^\top\cdot}\right\rangle_{L^2_r[0,1)^d}\mathrm{e}^{-2\pi i\alpha^\top(x+M^{-\top}p)}\\
&\quad=\sum_{p\in\mathscr{N}(M^\top)}\sum_{\alpha\in\mathbb{Z}^d}\left(\int_{[0,1)^d}\sum_{k=1}^{r}F_k(y)g_{j,k}(y)\mathrm{e}^{-2\pi i\alpha^\top M^\top y}\mathrm{d}y\right)\mathrm{e}^{-2\pi i\alpha^\top(x+M^{-\top}p)}\\
&\quad=\sum_{p\in\mathscr{N}(M^\top)}\sum_{k=1}^{r}F_k\left(x+M^{-\top}p\right)g_{j,k}(x+M^{-\top}p)\\
&\quad=\sum_{p\in\mathscr{N}(M^\top)}\mathbf{g}_j^\top\left(x+M^{-\top}p\right)\mathbf{F}(x+M^{-\top}p)\,.
\end{aligned}
$$

Defining $\mathbb{F}(x) := \left[\mathbf{F}^\top(x),\mathbf{F}^\top(x+M^{-\top}i_2),\ldots,\mathbf{F}^\top(x+M^{-\top}i_{\det M})\right]^\top$, the above equality allows us to write, in matrix form, that $\mathbb{G}(x)\,\mathbb{F}(x)$ equals to

$$
(\det M)\left[\sum_{\alpha\in\mathbb{Z}^d}\big(\mathscr{L}_1 f\big)(M\alpha)\mathrm{e}^{-2\pi i\alpha^\top M^\top x},\ldots,\sum_{\alpha\in\mathbb{Z}^d}\big(\mathscr{L}_s f\big)(M\alpha)\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right]^\top.
$$

In order to recover the function $\mathbf{F} = \mathscr{T}_{\boldsymbol{\Phi}}^{-1}f$, assume the existence of an $r\times s$ matrix $\mathbf{a}(x) := [\mathbf{a}_1(x),\ldots,\mathbf{a}_s(x)]$, with entries in $L^\infty[0,1)^d$, such that

$$
[\mathbf{a}_1(x),\ldots,\mathbf{a}_s(x)]\;\mathbb{G}(x) = [\mathbb{I}_r,\mathbb{O}_{(\det M-1)r\times r}]\quad\text{a.e. in }[0,1)^d\,.
$$

If we left multiply $\mathbb{G}(x)\mathbb{F}(x)$ by $\mathbf{a}(x)$, we get

$$
\mathbf{F}(x) = (\det M)\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}\left(\mathscr{L}_j f\right)(M\alpha)\,\mathbf{a}_j(x)\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\quad\text{in }L^2_r[0,1)^d\,. \tag{3.20}
$$

Finally, the isomorphism $\mathscr{T}_{\boldsymbol{\Phi}}$ gives

$$
f(t) = (\det M)\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}\big(\mathscr{L}_j f\big)(M\alpha)(\mathscr{T}_{\boldsymbol{\Phi}}\mathbf{a}_j)(t-M\alpha)\,,\quad t\in\mathbb{R}^d\,,
$$

where we have used the shifting property (3.7) and that the space V_{Φ}^2 is an RKHS. Much more can be said about the above sampling result. In fact, the following theorem holds:

Theorem 2. *Assume that the functions* $\mathbf{g}_j$ *given in* (3.11) *belong to* $L_r^{\infty}[0,1)^d$ *for each* $j = 1,2,\ldots,s$. *Let* $\mathbb{G}(x)$ *be the associated matrix defined in* $[0,1)^d$ *as in* (3.12). *The following statements are equivalents:*

(a) $A_{\mathbb{G}} > 0$.

(b) *There exists an* $r \times s$ *matrix* $\mathbf{a}(x) := \big[\mathbf{a}_1(x),\ldots,\mathbf{a}_s(x)\big]$ *with columns* $\mathbf{a}_j \in L_r^{\infty}[0,1)^d$ *and satisfying*

$$\big[\mathbf{a}_1(x),\ldots,\mathbf{a}_s(x)\big]\mathbb{G}(x) = [\mathbb{I}_r, \mathbb{O}_{(\det M-1)r\times r}] \quad \text{a.e. in } [0,1)^d. \tag{3.21}$$

(c) *There exists a frame for* V_{Φ}^2 *having the form* $\{S_{j,\mathbf{a}}(\cdot - M\alpha)\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ *such that for any* $f \in V_{\Phi}^2$

$$f = (\det M)\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}(\mathscr{L}_j f)(M\alpha)\,S_{j,\mathbf{a}}(\cdot - M\alpha) \quad \text{in } L^2(\mathbb{R}^d). \tag{3.22}$$

(d) *There exists a frame* $\{S_{j,\alpha}(\cdot)\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ *for* V_{Φ}^2 *such that for any* $f \in V_{\Phi}^2$

$$f = (\det M)\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}(\mathscr{L}_j f)(M\alpha)\,S_{j,\alpha} \quad \text{in } L^2(\mathbb{R}^d). \tag{3.23}$$

Proof. First we prove that (a) implies (b). As the determinant of the positive semidefinite matrix $\mathbb{G}^*(x)\mathbb{G}(x)$ is equal to the product of its eigenvalues, condition (a) implies that $\operatorname{ess\,inf}_{x\in\mathbb{R}^d}\det[\mathbb{G}^*(x)\mathbb{G}(x)] > 0$. Hence, there exists the left pseudo-inverse matrix $\mathbb{G}^{\dagger}(x) := [\mathbb{G}^*(x)\mathbb{G}(x)]^{-1}\mathbb{G}^*(x)$, a.e. in $[0,1)^d$, and it satisfies $\mathbb{G}^{\dagger}(x)\mathbb{G}(x) = \mathbb{I}_{r(\det M)}$. The first r rows of $\mathbb{G}^{\dagger}(x)$ form an $r \times s$ matrix $[\mathbf{a}_1(x),\ldots,\mathbf{a}_s(x)]$ which satisfies (3.21). Moreover, the functions $\mathbf{a}_j(x)$, $j = 1,2,\ldots,s$, are essentially bounded since the condition $\operatorname{ess\,inf}_{x\in[0,1)^d}\det[\mathbb{G}^*(x)\mathbb{G}(x)] > 0$ holds.

Next, we prove that (b) implies (c). For $j = 1,2,\ldots,s$, let $\mathbf{a}_j(x)$ be a function in $L_r^{\infty}[0,1)^d$ and satisfying $[\mathbf{a}_1(x),\ldots,\mathbf{a}_s(x)]\mathbb{G}(x) = [\mathbb{I}_r, \mathbb{O}_{(\det M-1)r\times r}]$. In (3.20) we have proved that, for each $\mathbf{F} = \mathscr{T}_{\Phi}^{-1}(f) \in L_r^2[0,1)^d$, we have the expansion

$$\mathbf{F}(x) = (\det M)\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}\big(\mathscr{L}_j f\big)(M\alpha)\,\mathbf{a}_j(x)\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}x} \quad \text{in } L_r^2[0,1)^d,$$

from which

$$f = (\det M)\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}\big(\mathscr{L}_j f\big)(M\alpha)\,S_{j,\mathbf{a}}(\cdot - M\alpha) \quad \text{in } L^2(\mathbb{R}^d),$$

where $S_{j,\mathbf{a}} := \mathscr{T}_{\Phi}\mathbf{a}_j$ for $j = 1,2,\dots,s$. Since we have assumed that $\mathbf{g}_j \in L^{\infty}_r[0,1)^d$ for each $j = 1,2,\dots,s$, the sequence $\left\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\dots,s}$ is a Bessel sequence in $L^2_r[0,1)^d$ by using part (b) in Lemma 3. The same argument proves that the sequence $\left\{(\det M)\mathbf{a}_j(x)\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\dots,s}$ is also a Bessel sequence in $L^2_r[0,1)^d$. These two Bessel sequences satisfy for each $\mathbf{F}\in L^2_r[0,1)^d$

$$\mathbf{F}(x) = (\det M)\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}\left\langle \mathbf{F}, \overline{\mathbf{g}_j}\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}\cdot}\right\rangle \mathbf{a}_j(x)\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}x} \quad \text{in } L^2_r[0,1)^d.$$

Hence, they are a pair of dual frames for $L^2_r[0,1)^d$ (see [11, Lemma 5.6.2]). Since $\mathscr{T}_{\Phi}$ is an isomorphism, the sequence $\left\{S_{j,\mathbf{a}}(t-M\alpha)\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\dots,s}$ is a frame for V^2_{Φ}; hence, (b) implies (c). Statement (c) implies (d) trivially.

Assume condition (d), applying the isomorphism $\mathscr{T}_{\Phi}^{-1}$ to the expansion (3.23) we get

$$\mathbf{F}(x) = (\det M)\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}\left\langle \mathbf{F}, \overline{\mathbf{g}_j}\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}\cdot}\right\rangle \mathscr{T}_{\Phi}^{-1}(S_{j,\alpha})(x) \quad \text{in } L^2_r[0,1)^d, \tag{3.24}$$

where $\left\{\mathscr{T}_{\Phi}^{-1}S_{j,\alpha}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\dots,s}$ is a frame for $L^2_r[0,1)^d$. By using Lemma 3, the sequence $\left\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^{\top}M^{\top}x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\dots,s}$ is a Bessel sequence; expansion (3.24) implies that is also a frame (see [11, Lemma 5.6.2]). Hence, by using again Lemma 3, condition (a) holds. □

In the case that the functions $\mathbf{g}_j$, $j = 1,2,\dots,s$, are continuous on $\mathbb{R}^d$ (for instance, if the sequences of generalized samples $\left\{\mathscr{L}_j\varphi_k(\alpha)\right\}_{\alpha\in\mathbb{Z}^d}$ belongs to $\ell^1(\mathbb{Z}^d)$ for $1\le j\le s$ and $1\le k\le r$), the following corollary holds:

Corollary 1. *Assume that the functions $\mathbf{g}_j$, $j = 1,2,\dots,s$, in (3.11) are continuous on $\mathbb{R}^d$. Then, the following assertions are equivalents:*

(a) *rank $\mathbb{G}(x) = r(\det M)$ for all $x\in\mathbb{R}^d$.*
(b) *There exists a frame $\{S_{j,\mathbf{a}}(\cdot - rn)\}_{n\in\mathbb{Z},\ j=1,2,\dots,s}$ for V^2_{Φ} satisfying the sampling formula (3.22).*

Proof. Whenever the functions $\mathbf{g}_j$, $j = 1,2,\dots,s$, are continuous on $\mathbb{R}^d$, condition $A_{\mathbb{G}} > 0$ is equivalent to that $\det\left[\mathbb{G}^*(x)\mathbb{G}(x)\right] \neq 0$ for all $x\in\mathbb{R}^d$. Indeed, if $\det\mathbb{G}^*(x)\mathbb{G}(x) > 0$, then the r first rows of the matrix $\mathbb{G}^{\dagger}(x) := [\mathbb{G}^*(x)\mathbb{G}(x)]^{-1}\mathbb{G}^*(x)$ give an $r\times s$ matrix $\mathbf{a}(x) = [\mathbf{a}_1(x),\mathbf{a}_2(x),\dots,\mathbf{a}_s(x)]$ satisfying statement (b) in Theorem 2, and therefore $A_{\mathbb{G}} > 0$.

The reciprocal follows from the fact that $\det\left[\mathbb{G}^*(x)\mathbb{G}(x)\right] \ge A_{\mathbb{G}}^{r(\det M)}$ for all $x\in\mathbb{R}^d$. Since $\det\left[\mathbb{G}^*(x)\mathbb{G}(x)\right]\neq 0$ is equivalent to rank $\mathbb{G}(x) = r(\det M)$ for all $x\in\mathbb{R}^d$, the result is a consequence of Theorem 2. □

The reconstruction functions $S_{j,\mathbf{a}}$, $j=1,2,\dots,s$, are determined from the Fourier coefficients of the components of $\mathbf{a}_j(x) := [a_{1,j}(x), a_{2,j}(x), \dots, a_{r,j}]^\top$, $j=1,2,\dots,s$. More specifically, if $\widehat{a}_{k,j}(\alpha) := \int_{[0,1)^d} a_{k,j}(x)\mathrm{e}^{2\pi i\alpha^\top x}\mathrm{d}x$, we get (see (3.6))

$$S_{j,\mathbf{a}}(t) = \sum_{\alpha\in\mathbb{Z}^d}\sum_{k=1}^{r} \widehat{a}_{k,j}(\alpha)\varphi_k(t-\alpha)\,, \quad t\in\mathbb{R}^d\,. \tag{3.25}$$

The Fourier transform in (3.25) gives $\widehat{S}_{j,\mathbf{a}}(x) = \sum_{k=1}^{r} a_{k,j}(x)\widehat{\varphi_k}(x)$.

Assume that the $r\times s$ matrix $\mathbf{a}(x) = \big[\mathbf{a}_1(x), \mathbf{a}_2(x), \dots, \mathbf{a}_s(x)\big]$ satisfies (3.21). We consider the periodic extension of $a_{k,j}$, i.e., $a_{k,j}(x+\alpha) = a_{k,j}(x)$, $\alpha\in\mathbb{Z}^d$. For all $x\in[0,1)^d$, the $r(\det M)\times s$ matrix

$$\mathbb{A}^\top(x) := \begin{bmatrix} \mathbf{a}_1(x) & \mathbf{a}_2(x) & \cdots & \mathbf{a}_s(x) \\ \mathbf{a}_1(x+M^{-\top}i_2) & \mathbf{a}_2(x+M^{-\top}i_2) & \cdots & \mathbf{a}_s(x+M^{-\top}i_2) \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_1(x+M^{-\top}i_{\det M}) & \mathbf{a}_2(x+M^{-\top}i_{\det M}) & \cdots & \mathbf{a}_s(x+M^{-\top}i_{\det M}) \end{bmatrix} \tag{3.26}$$

is a left inverse matrix of $\mathbb{G}(x)$, i.e., $\mathbb{A}^\top(x)\mathbb{G}(x) = \mathbb{I}_{r(\det M)}$.

Provided that condition (3.21) is satisfied, it can be easily checked that all matrices $\mathbf{a}(x)$ with entries in $L^\infty[0,1)^d$ and satisfying (3.21) correspond to the first r rows of the matrices of the form

$$\mathbb{A}^\top(x) = \mathbb{G}^\dagger(x) + \mathbb{U}(x)\big[\mathbb{I}_s - \mathbb{G}(x)\mathbb{G}^\dagger(x)\big]\,, \tag{3.27}$$

where $\mathbb{U}(x)$ is any $r(\det M)\times s$ matrix with entries in $L^\infty[0,1)^d$, and $\mathbb{G}^\dagger$ denotes the left pseudo-inverse $\mathbb{G}^\dagger(x) := [\mathbb{G}^*(x)\mathbb{G}(x)]^{-1}\mathbb{G}^*(x)$.

Notice that if $s = r(\det M)$, there exists a unique matrix $\mathbf{a}(x)$, given by the first r rows of $\mathbb{G}^{-1}(x)$; if $s > r(\det M)$, there are infinitely many solutions according to (3.27).

Moreover, the sequence $\big\{(\det M)\mathbf{a}_j^\dagger(\cdot)\mathrm{e}^{-2\pi i\alpha^\top M^\top\cdot}\big\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\dots,s}$, associated with the $r\times s$ matrix $[\mathbf{a}_1^\dagger(x), \mathbf{a}_2^\dagger(x), \dots, \mathbf{a}_s^\dagger(x)]$ obtained from the r first rows of $\mathbb{G}^\dagger(x)$, gives precisely the canonical dual frame of the frame $\big\{\overline{\mathbf{g}_j(\cdot)}\mathrm{e}^{-2\pi i\alpha^\top M^\top\cdot}\big\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\dots,s}$. Indeed, the frame operator $\mathscr{S}$ associated to $\big\{\overline{\mathbf{g}_j(\cdot)}\mathrm{e}^{-2\pi i\alpha^\top M^\top\cdot}\big\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\dots,s}$ is given by

$$\mathscr{S}\mathbf{F}(x) = \frac{1}{\det M}\Big[\overline{\mathbf{g}_1(x)}, \overline{\mathbf{g}_2(x)}, \dots, \overline{\mathbf{g}_s(x)}\Big]\,\mathbb{G}(x)\mathbb{F}(x)\,, \quad \mathbf{F}\in L^2_r[0,1)^d\,,$$

from which one gets

$$\mathscr{S}\Big[(\det M)\mathbf{a}_j^\dagger(\cdot)\mathrm{e}^{-2\pi i\alpha^\top M^\top\cdot}\Big](x) = \overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\,, \quad j=1,2,\dots,s \text{ and } \alpha\in\mathbb{Z}^d\,.$$

Something more can be said in the case where $s = r(\det M)$:

Theorem 3. *Assume that the functions* $\mathbf{g}_j$, $j = 1, 2, \ldots, s$, *given in* (3.11) *belong to* $L^\infty_r[0,1)^d$ *and* $s = r(\det M)$. *The following statements are equivalent:*

(a) $A_{\mathbb{G}} > 0$.

(b) There exists a Riesz basis $\{S_{j,\alpha}\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ *for* V^2_{Φ} *such that for any* $f \in V^2_{\Phi}$, *the expansion*

$$f = (\det M) \sum_{\alpha\in\mathbb{Z}^d} \sum_{j=1}^{s} (\mathscr{L}_j f)(M\alpha)\, S_{j,\alpha} \tag{3.28}$$

holds in $L^2(\mathbb{R}^d)$.

In case the equivalent conditions are satisfied, necessarily $S_{j,\alpha}(t) = S_{j,\mathbf{a}}(t - M\alpha)$, $t \in \mathbb{R}^d$, *where* $S_{j,\mathbf{a}} = \mathscr{T}_{\Phi}(\mathbf{a}_j)$, $j = 1, 2, \ldots, s$, *and the* $r \times s$ *matrix* $\mathbf{a} := [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_s]$ *is formed with the* r *first rows of the inverse matrix* $\mathbb{G}^{-1}$. *The sampling functions* $S_{j,\mathbf{a}}$, $j = 1, 2, \ldots, s$, *satisfy the interpolation property* $(\mathscr{L}_{j'} S_{j,\mathbf{a}})(M\alpha) = \delta_{j,j'}\delta_{\alpha,0}$, *where* $j, j' = 1, 2, \ldots, s$ *and* $\alpha \in \mathbb{Z}^d$.

Proof. Assume that $A_{\mathbb{G}} > 0$; since $\mathbb{G}(x)$ is a square matrix, this implies that $\operatorname{ess\,inf}_{x\in\mathbb{R}^d} |\det \mathbb{G}(x)| > 0$. Therefore, the r first rows of $\mathbb{G}^{-1}(x)$ gives a solution of the equation $[\mathbf{a}_1(x), \ldots, \mathbf{a}_s(x)]\mathbb{G}(x) = [\mathbb{I}_r, \mathbb{O}_{(\det M - 1)r\times r}]$ with $\mathbf{a}_j \in L^\infty_r[0,1)^d$ for $j = 1, 2, \ldots, s$. According to Theorem 2, the sequence

$$\{S_{j,\alpha}\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s} := \{S_{j,\mathbf{a}}(t - M\alpha)\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s},$$

where $S_{j,\mathbf{a}} = \mathscr{T}_{\Phi}(\mathbf{a}_j)$, satisfies the sampling formula (3.28). Moreover, the sequence

$$\left\{(\det M)\mathbf{a}_j(x)\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s} = \left\{\mathscr{T}_{\Phi}^{-1} S_{j,\mathbf{a}}(\cdot - M\alpha)\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$$

is a frame for $L^2_r[0,1)^d$. Since $r(\det M) = s$, according to Lemma 3, it is a Riesz basis for $L^2_r[0,1)^d$. Hence, the sequence $\{S_{j,\mathbf{a}}(t - M\alpha)\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a Riesz basis for V^2_{Φ}, and condition (b) is proved.

Conversely, assume now that $\{S_{j,\alpha}\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a Riesz basis for V^2_{Φ} satisfying (3.28). From the uniqueness of the coefficients in a Riesz basis, we get that the interpolatory condition $(\mathscr{L}_{j'} S_{j,\alpha})(M\alpha') = \delta_{j,j'}\delta_{\alpha,\alpha'}$ holds for $j, j' = 1, 2, \ldots, s$ and $\alpha, \alpha' \in \mathbb{Z}^d$. Since $\mathscr{T}_{\Phi}^{-1}$ is an isomorphism, $\{\mathscr{T}_{\Phi}^{-1} S_{j,\alpha}\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ is a Riesz basis for $L^2_r[0,1)^d$. Expanding the function $\overline{\mathbf{g}_{j'}(x)}\mathrm{e}^{-2\pi i\alpha'^\top M^\top x}$ with respect to the dual basis of $\{\mathscr{T}_{\Phi}^{-1} S_{j,\alpha}\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$, denoted by $\{G_{j,\alpha}\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$, we obtain

$$\begin{aligned}\overline{\mathbf{g}_{j'}(x)}\mathrm{e}^{-2\pi i\alpha'^\top M^\top x} &= \sum_{\alpha\in\mathbb{Z}^d}\sum_{j=1}^{s} \left\langle \overline{\mathbf{g}_{j'}(\cdot)}\mathrm{e}^{-2\pi i\alpha'^\top M^\top \cdot}, \mathscr{T}_{\Phi}^{-1} S_{j,\alpha}\right\rangle_{L^2[0,1)^d} G_{j,\alpha}(x)\\ &= \sum_{\alpha\in\mathbb{Z}^d} \overline{\mathscr{L}_{j'} S_{j,\alpha}}(M\alpha') G_{j,\alpha}(x) = G_{j',\alpha'}(x)\,.\end{aligned}$$

Therefore, the sequence $\left\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ is the dual basis of the Riesz basis $\left\{\mathscr{T}_{\boldsymbol{\Phi}}^{-1}S_{j,\alpha}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$. In particular, it is a Riesz basis for $L^2_r[0,1)^d$, which implies, according to Lemma 3, that $A_{\mathbf{G}}>0$; this proves (a). Moreover, the sequence $\left\{\mathscr{T}_{\boldsymbol{\Phi}}^{-1}S_{j,\alpha}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ is necessarily the unique dual basis of the Riesz basis $\left\{\overline{\mathbf{g}_j(x)}\mathrm{e}^{-2\pi i\alpha^\top M^\top x}\right\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$. Therefore, this proves the uniqueness of the Riesz basis $\{S_{j,\alpha}\}_{\alpha\in\mathbb{Z}^d,\ j=1,2,\ldots,s}$ for $V^2_{\boldsymbol{\Phi}}$ satisfying (3.28). □

3.3.1 Reconstruction Functions with Prescribed Properties

A generalized sampling formula in the shift-invariant space $V^2_{\boldsymbol{\Phi}}$ as

$$f(t)=(\det M)\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}(\mathscr{L}_jf)(M\alpha)S_{j,\mathbf{a}}(t-M\alpha)\,,\quad t\in\mathbb{R}^d\,, \tag{3.29}$$

can be read as a filter bank. Indeed, introducing the expression for the sampling functions $S_{j,\mathbf{a}}(t)=\sum_{\beta\in\mathbb{Z}^d}\sum_{k=1}^{r}\widehat{a}_{k,j}(\beta)\varphi_k(t-\beta)\,,\ t\in\mathbb{R}^d$, the change $\gamma:=\beta+M\alpha$ in the summation's index gives

$$f(t)=(\det M)\sum_{k=1}^{r}\sum_{\gamma\in\mathbb{Z}^d}\left\{\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}(\mathscr{L}_jf)(M\alpha)\widehat{a}_{k,j}(\gamma-M\alpha)\right\}\varphi_k(t-\gamma)\,,\quad t\in\mathbb{R}^d\,.$$

Thus, the relevant data for the recovery of the signal $f\in V^2_{\boldsymbol{\Phi}}$,

$$d_k(\gamma):=\sum_{j=1}^{s}\sum_{\alpha\in\mathbb{Z}^d}(\mathscr{L}_jf)(M\alpha)\widehat{a}_{k,j}(\gamma-M\alpha)\,,\quad \gamma\in\mathbb{Z}^d\,,\quad 1\le k\le r\,,$$

is obtained by means of r filter banks whose impulse responses involve the Fourier coefficients of the entries of the $r\times s$ matrix $\mathbf{a}:=\left[\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_s\right]$ in (3.21), and the input is given by the sampling data.

Notice that reconstruction functions $S_{j,\mathbf{a}}$ with compact support in the above sampling formula implies low computational complexities and avoids truncation errors. This occurs whenever the generators φ_k have compact support and the sum in (3.25) is finite. These sums are finite if and only if the entries of the $r\times s$ matrix $\mathbf{a}$ are trigonometric polynomials. In this case, all the filter banks involved in the reconstruction process are finite impulse response (FIR) filters.

In order to give a necessary and sufficient condition assuring compactly supported reconstruction functions $S_{j,\mathbf{a}}$ in formula (3.29), we introduce first some complex notation, more convenient for this study. We denote $\mathbf{z}^\alpha:=z_1^{\alpha_1}z_2^{\alpha_2}\cdots z_d^{\alpha_d}$ for $\mathbf{z}=(z_1,\ldots,z_d)\in\mathbb{C}^d$, $\alpha=(\alpha_1,\ldots,\alpha_d)\in\mathbb{Z}^d$, and the d-torus by $\mathbb{T}^d:=\{\mathbf{z}\in\mathbb{C}^d:|z_1|=|z_2|=\cdots=|z_d|=1\}$. For $1\le j\le s$ and $1\le k\le r$, we define

$$\mathsf{g}_{j,k}(\mathbf{z}) := \sum_{\mu\in\mathbb{Z}^d} \mathscr{L}_j\varphi_k(\mu)\mathbf{z}^{-\mu}, \quad \mathsf{g}_j^\top(\mathbf{z}) := \big(\mathsf{g}_{j,1}(\mathbf{z}), \mathsf{g}_{j,2}(\mathbf{z}), \ldots, \mathsf{g}_{j,r}(\mathbf{z})\big)$$

and the $s \times r(\det M)$ matrix

$$\mathsf{G}(\mathbf{z}) := \Big[\mathsf{g}_j^\top\left(z_1 \mathrm{e}^{2\pi i m_1^\top i_l}, \ldots, z_d \mathrm{e}^{2\pi i m_d^\top i_l}\right)\Big]_{\substack{j=1,2,\ldots,s \\ k=1,2,\ldots,r;\ l=1,2,\ldots,\det M}} \tag{3.30}$$

where $m_1, \ldots, m_d$ denote the columns of the matrix M^{-1}. Recall that $i_1, i_2, \ldots, i_{\det M}$ in $\mathbb{Z}^d$ are the elements of $\mathscr{N}(M^\top)$ defined in (3.9). Note also that for the values $x = (x_1, \ldots, x_d) \in [0,1)^d$ and $\mathbf{z} = (\mathrm{e}^{2\pi i x_1}, \ldots, \mathrm{e}^{2\pi i x_d}) \in \mathbb{T}^d$, we have $\mathbb{G}(x) = \mathsf{G}(\mathbf{z})$.

Provided that the functions $\mathbf{g}_j$ are continuous on $\mathbb{R}^d$, Corollary 1 can be reformulated as follows: There exists an $r \times s$ matrix $\mathsf{a}(\mathbf{z}) = \big[\mathsf{a}_1(\mathbf{z}), \ldots, \mathsf{a}_s(\mathbf{z})\big]$ with entries essentially bounded in the torus $\mathbb{T}^d$ and satisfying

$$\mathsf{a}(\mathbf{z})\mathsf{G}(\mathbf{z}) = [\mathbb{I}_r, \mathbb{O}_{(\det M - 1)r\times r}] \quad \text{for all } \mathbf{z} \in \mathbb{T}^d \tag{3.31}$$

if and only if

$$\operatorname{rank} \mathsf{G}(\mathbf{z}) = r(\det M) \quad \text{for all } \mathbf{z} \in \mathbb{T}^d\,. \tag{3.32}$$

Denoting the columns of the matrix $\mathsf{a}(\mathbf{z})$ as $\mathsf{a}_j^\top(\mathbf{z}) = \big(\mathsf{a}_{1,j}(\mathbf{z}), \ldots, \mathsf{a}_{r,j}(\mathbf{z})\big)$, $j = 1, 2, \ldots, s$, the corresponding reconstruction functions $S_{j,\mathsf{a}}$ in sampling formula (3.29) are

$$S_{j,\mathsf{a}}(t) = \sum_{\alpha\in\mathbb{Z}^d} \sum_{k=1}^{r} \widehat{\mathsf{a}}_{k,j}(\alpha)\varphi_k(t-\alpha)\,, \quad t \in \mathbb{R}^d\,, \tag{3.33}$$

where $\widehat{\mathsf{a}}_{k,j}(\alpha)$, $\alpha \in \mathbb{Z}^d$, are the Laurent coefficients of the functions $\mathsf{a}_{k,j}(\mathbf{z})$, i.e.,

$$\mathsf{a}_{k,j}(\mathbf{z}) = \sum_{\alpha\in\mathbb{Z}^d} \widehat{\mathsf{a}}_{k,j}(\alpha)\mathbf{z}^{-\alpha}\,. \tag{3.34}$$

Note that, in order to obtain compactly supported reconstruction functions $S_{j,\mathsf{a}}$ in (3.29), we need an $r \times s$ matrix $\mathsf{a}(\mathbf{z})$ whose entries are Laurent polynomials, i.e., the sum in (3.34) is finite. The following result, which proof can be found in [16] under minor changes, holds:

Theorem 4. *Assume that the generators φ_k and the functions $\mathscr{L}_j\varphi_k$, $1 \le k \le r$ and $1 \le j \le s$, have compact support. Then, there exists an $r \times s$ matrix $\mathsf{a}(\mathbf{z})$ whose entries are Laurent polynomials and satisfying (3.31) if and only if*

$$\textit{rank } \mathsf{G}(\mathbf{z}) = r(\det M) \quad \textit{for all } \mathbf{z} \in (\mathbb{C}\setminus\{0\})^d\,.$$

The reconstruction functions $S_{j,\mathsf{a}}$, $j = 1, 2, \ldots, s$, obtained from such matrix $\mathsf{a}(\mathbf{z})$ through (3.33) have compact support.

From one of these $r \times s$ matrices, say $\widetilde{\mathsf{a}}(\mathbf{z}) = [\widetilde{\mathsf{a}}_1(\mathbf{z}), \ldots, \widetilde{\mathsf{a}}_s(\mathbf{z})]$, we can get all of them. Indeed, it is easy to check that they are given by the r first rows of the $r(\det M) \times s$ matrices of the form

$$\mathsf{A}(\mathbf{z}) = \widetilde{\mathsf{A}}(\mathbf{z}) + \mathsf{U}(\mathbf{z}) \left[\mathbb{I}_s - \mathsf{G}(\mathbf{z}) \widetilde{\mathsf{A}}(\mathbf{z}) \right], \tag{3.35}$$

where

$$\widetilde{\mathsf{A}}(\mathbf{z}) := \left[\widetilde{\mathsf{a}}_j(z_1 \mathrm{e}^{2\pi i m_1^\top i_l}, \ldots, z_d \mathrm{e}^{2\pi i m_d^\top i_l}) \right]_{\substack{k=1,2,\ldots,r;\ l=1,2,\ldots,\det M \\ j=1,2,\ldots,s}},$$

and $\mathsf{U}(\mathbf{z})$ is any $r(\det M) \times s$ matrix with Laurent polynomial entries. Remember that $m_1, \ldots, m_d$ denote the columns of the matrix M^{-1} and $i_1, \ldots, i_{\det M}$ the elements of $\mathscr{N}(M^\top)$ defined in (3.9).

Next, we study the existence of reconstruction functions $S_{j,\mathsf{a}}$, $j = 1, 2, \ldots, s$, in (3.29) having exponential decay; it means that there exist constants $C > 0$ and $q \in (0,1)$ such that $|S_{j,\mathsf{a}}(t)| \leq C q^{|t|}$ for each $t \in \mathbb{R}^d$. In so doing, we introduce the algebra $\mathscr{H}(\mathbb{T}^d)$ of all holomorphic functions in a neighborhood of the d-torus $\mathbb{T}^d$. Note that the elements in $\mathscr{H}(\mathbb{T}^d)$ are characterized as admitting a Laurent series where the sequence of coefficients decays exponentially fast [27].

The following theorem, which proof can be found in [16] under minor changes, holds:

Theorem 5. *Assume that the generators φ_k and the functions $\mathscr{L}_j \varphi_k$, $j = 1, 2, \ldots, s$ and $k = 1, 2, \ldots, r$, have exponential decay. Then, there exists an $r \times s$ matrix $\mathsf{a}(\mathbf{z}) = [\mathsf{a}_1(\mathbf{z}), \ldots, \mathsf{a}_s(\mathbf{z})]$ with entries in $\mathscr{H}(\mathbb{T}^d)$ and satisfying (3.31) if and only if rank $\mathsf{G}(\mathbf{z}) = r(\det M)$ for all $\mathbf{z} \in \mathbb{T}^d$.*

In this case, all of such matrices $\mathsf{a}(\mathbf{z})$ are given as the first r rows of a $r(\det M) \times s$ matrix $\mathsf{A}(\mathbf{z})$ of the form

$$\mathsf{A}(\mathbf{z}) = \mathsf{G}^\dagger(\mathbf{z}) + \mathsf{U}(\mathbf{z}) \left[\mathbb{I}_s - \mathsf{G}(\mathbf{z}) \mathsf{G}^\dagger(\mathbf{z}) \right], \tag{3.36}$$

where $\mathsf{U}(\mathbf{z})$ denotes any $r(\det M) \times s$ matrix with entries in the algebra $\mathscr{H}(\mathbb{T}^d)$ and $\mathsf{G}^\dagger(\mathbf{z}) := \left[\mathsf{G}^*(\mathbf{z}) \mathsf{G}(\mathbf{z}) \right]^{-1} \mathsf{G}^*(\mathbf{z})$. The corresponding reconstruction functions $S_{j,\mathsf{a}}$, $j = 1, 2, \ldots, s$, given by (3.33) have exponential decay.

3.3.2 Some Illustrative Examples

We include here some examples illustrating Theorem 4, a particular case of Theorem 2, by taking B-splines as generators; they certainly are important for practical purposes [44].

First notice that if the generator φ has compact support, the only situation when the reconstruction function S_a in formula (3.1) has compact support as well is the special case when φ is the linear B-spline $N_2(t) := \chi_{[0,1)} * \chi_{[0,1)}(t)$, where $\chi_{[0,1)}$ denotes the characteristic function of the interval $[0,1)$. For any $f \in V^2_{N_2}$, the following sampling formula holds:

$$f(t) = \sum_{n=-\infty}^{\infty} f(n) N_2(t+1-n), \quad t \in \mathbb{R}.$$

In this special case where $d = 1$ and $r = s = 1$, we have $\mathsf{G}(z) = z$, and consequently, $\mathsf{a}(z) = z^{-1}$ in Theorem 4.

3.3.2.1 The Case $d = 1, r = 1, M = 2$, and $s = 3$

Let $N_3(t) := \chi_{[0,1)} * \chi_{[0,1)} * \chi_{[0,1)}(t)$ be the quadratic B-spline, and let $\mathscr{L}_j$, $j = 1,2,3$, be the systems

$$\mathscr{L}_1 f(t) = f(t); \quad \mathscr{L}_2 f(t) = f\left(t + \frac{2}{3}\right) \quad \text{and} \quad \mathscr{L}_3 f(t) = f\left(t + \frac{4}{3}\right).$$

Since the functions $\mathscr{L}_j N_3$, $j = 1,2,3$, have compact support, then the entries of the 3×2 matrix $\mathsf{G}(z)$ in (3.30) are Laurent polynomials, and we can try to search a vector $\mathbf{a}(z) := [\mathbf{a}_1(z), \mathbf{a}_2(z), \mathbf{a}_3(z)]$ satisfying (3.31) with Laurent polynomials entries also. This implies reconstruction functions $S_{j,\mathbf{a}}$, $j = 1,2,3$, with compact support. Proceeding as in [14], we obtain that any function $f \in V^2_{N_3}$ can be recovered through the sampling formula

$$f(t) = \sum_{n \in \mathbb{Z}} \sum_{j=1}^{3} \mathscr{L}_j f(2n) S_{j,\mathbf{a}}(t-2n), \quad t \in \mathbb{R},$$

where the reconstruction functions, according to (3.33), are given by

$$S_{1,\mathbf{a}}(t) = \frac{1}{16}\left[N_3(t+3) - 3N_3(t+2) - 3N_3(t+1) + N_3(t)\right],$$

$$S_{2,\mathbf{a}}(t) = \frac{1}{16}\left[27N_3(t+1) - 9N_3(t)\right],$$

$$S_{3,\mathbf{a}}(t) = \frac{1}{16}\left[-9N_3(t+1) + 27N_3(t)\right], \quad t \in \mathbb{R}.$$

3.3.2.2 The Case $d = 1, r = 2, M = 1$, and $s = 3$

Consider the Hermite cubic splines defined as

$$\varphi_1(t) = \begin{cases} (t+1)^2(1-2t), & t \in [-1,0] \\ (1-t)^2(1+2t), & t \in [0,1] \\ 0, & |t| > 1 \end{cases} \quad \text{and} \quad \varphi_2(t) = \begin{cases} (t+1)^2 t, & t \in [-1,0] \\ (1-t)^2 t, & t \in [0,1] \\ 0, & |t| > 1 \end{cases}.$$

They are stable generators for the space $V^2_{\varphi_1,\varphi_2}$ (see [12]). Consider the sampling period $M = 1$ and the systems $\mathscr{L}_j$, $j = 1,2,3$, defined by

$$\mathscr{L}_1 f(t) := \int_t^{t+1/3} f(u)\mathrm{d}u, \quad \mathscr{L}_2 f(t) := \mathscr{L}_1 f\Big(t + \frac{1}{3}\Big), \quad \mathscr{L}_3 f(t) := \mathscr{L}_1 f\Big(t + \frac{2}{3}\Big).$$

Since the functions $\mathscr{L}_j \varphi_k$, $j = 1,2,3$ and $k = 1,2$, have compact support, then the entries of the 3×2 matrix $\mathsf{G}(z)$ in (3.30) are Laurent polynomials, and we can try to search an 2×3 matrix $\mathbf{a}(z) := [\mathbf{a}_1(z), \mathbf{a}_2(z), \mathbf{a}_3(z)]$ satisfying (3.31) with Laurent polynomials entries also. This leads to reconstruction functions $S_{j,\mathbf{a}}$, $j = 1,2,3$, with compact support. Proceeding as in [17], we obtain in $V^2_{\varphi_1,\varphi_2}$ the following sampling formula:

$$f(t) = \sum_{n \in \mathbb{Z}} \sum_{j=1}^{3} \mathscr{L}_j f(n) S_{j,\mathbf{a}}(t-n), \quad t \in \mathbb{R},$$

where the sampling functions, according to (3.33), are

$$\begin{aligned} S_{1,\mathbf{a}}(t) &:= \frac{85}{44}\varphi_1(t) + \frac{1}{11}\varphi_1(t-1) + \frac{85}{4}\varphi_2(t) - \varphi_2(t-1), \\ S_{2,\mathbf{a}}(t) &:= \frac{-23}{44}\varphi_1(t) - \frac{23}{44}\varphi_1(t-1) - \frac{23}{4}\varphi_2(t) + \frac{23}{4}\varphi_2(t-1), \\ S_{3,\mathbf{a}}(t) &:= \frac{1}{11}\varphi_1(t) + \frac{85}{44}\varphi_1(t-1) + \varphi_2(t) - \frac{85}{4}\varphi_2(t-1), \quad t \in \mathbb{R}. \end{aligned}$$

3.3.3 L^2-Approximation Properties

Consider an $r \times s$ matrix $\mathbf{a}(x) := \big[\mathbf{a}_1(x), \mathbf{a}_2(x), \ldots, \mathbf{a}_s(x)\big]$ with entries $a_{k,j} \in L^\infty[0,1)^d$, $1 \le k \le r$, $1 \le j \le s$, and satisfying (3.21). Let $S_{j,\mathbf{a}}$ be the associated reconstruction functions, $j = 1,2,\ldots,s$, given in Theorem 2. The aim of this section is to show that if the set of generators Φ satisfies the Strang–Fix conditions of order ℓ, then the scaled version of the sampling operator

$$\Gamma_{\mathbf{a}} f(t) := \sum_{j=1}^{s} \sum_{\alpha \in \mathbb{Z}^d} (\mathscr{L}_j f)(M\alpha) S_{j,\mathbf{a}}(t - M\alpha), \quad t \in \mathbb{R}^d,$$

gives L^2-approximation order ℓ for any smooth function f (in a Sobolev space). In so doing, we take advantage of the good approximation properties of the scaled space $\sigma_{1/h} V_{\Phi}^2$, where for $h > 0$, we are using the notation $\sigma_h f(t) := f(ht)$, $t \in \mathbb{R}^d$.

The set of generators $\boldsymbol{\Phi} = \{\varphi_k\}_{k=1}^r$ is said to satisfy the Strang–Fix conditions of order ℓ if there exist r finitely supported sequences $b_k : \mathbb{Z}^d \to \mathbb{C}$ such that the function $\varphi(t) = \sum_{k=1}^r \sum_{\alpha \in \mathbb{Z}^d} b_k(\alpha) \varphi_k(t - \alpha)$ satisfies the Strang–Fix conditions of order ℓ, i.e.,

$$\widehat{\varphi}(0) \neq 0, \quad D^{\beta} \widehat{\varphi}(\alpha) = 0, \quad |\beta| < \ell, \quad \alpha \in \mathbb{Z}^d \setminus \{0\}. \tag{3.37}$$

We denote by $W_2^{\ell}(\mathbb{R}^d) := \{f : \|D^{\gamma} f\|_2 < \infty, |\gamma| \leq \ell\}$ the usual Sobolev space and by $|f|_{\ell,2} := \sum_{|\beta|=\ell} \|D^{\beta} f\|_2$ the corresponding seminorm of a function $f \in W_2^{\ell}(\mathbb{R}^d)$. When $2\ell > d$, we identify $f \in W_2^{\ell}(\mathbb{R}^d)$ with its continuous choice (see [2]).

It is well known that if $\boldsymbol{\Phi}$ satisfies the Strang–Fix conditions of order ℓ and the generators φ_k satisfy a suitable decay condition, the space V_{Φ}^2 provides L^2-approximation order ℓ for any function f regular enough. For instance, Lei et al. proved in [33, Theorem 5.2] the following result: If a set $\boldsymbol{\Phi} = \{\varphi_k\}_{k=1}^r$ of stable generators satisfies the Strang–Fix conditions of order ℓ and the decay condition $\varphi_k(t) = O\big([1 + |t|]^{-d-\ell-\varepsilon}\big)$ for each $k = 1, 2, \ldots, r$ and some $\varepsilon > 0$, then, for any $f \in W_2^{\ell}(\mathbb{R}^d)$, there exists a function $f_h \in \sigma_{1/h} V_{\Phi}^2$ such that

$$\|f - f_h\|_2 \leq C\, |f|_{\ell,2}\, h^{\ell}, \tag{3.38}$$

where the constant C does not depend on h and f.

In this section we assume that all the systems $\mathscr{L}_j$, $j = 1, 2, \ldots, s$, are of type (b), i.e., $\mathscr{L}_j f = f * \mathsf{h}_j$, belonging the impulse response h_j to the Hilbert space $\mathscr{L}^2(\mathbb{R}^d)$. Recall that a Lebesgue measurable function $\mathsf{h} : \mathbb{R}^d \longrightarrow \mathbb{C}$ belongs to the Hilbert space $\mathscr{L}^2(\mathbb{R}^d)$ if

$$|\mathsf{h}|_2 := \left(\int_{[0,1)^d} \Big(\sum_{\alpha \in \mathbb{Z}^d} |\mathsf{h}(t - \alpha)| \Big)^2 \mathrm{d}t \right)^{1/2} < \infty.$$

Notice that the space $\mathscr{L}^2(\mathbb{R}^d)$ coincides with the amalgam space $W(\ell^1, L^2)$ and that $\mathscr{L}^2(\mathbb{R}^d) \subset L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. For $f \in L^2(\mathbb{R}^d)$ and $\mathsf{h} \in \mathscr{L}^2(\mathbb{R}^d)$, the following inequality holds: $\big\| \{\mathsf{h} * f(\alpha)\}_{\alpha \in \mathbb{Z}^d} \big\|_2 \leq |\mathsf{h}|_2 \, \|f\|_2$ (see [27, Theorem 3.1]); thus, the sequence of generalized samples $\{(\mathscr{L}_j f)(M\alpha)\}_{\alpha \in \mathbb{Z}^d,\ j=1,2,\ldots,s}$ belongs to $\ell^2(\mathbb{Z}^d)$ for any $f \in L^2(\mathbb{R}^d)$.

First we note that the operator $\Gamma_{\mathbf{a}} : \big(L^2(\mathbb{R}^d), \|\cdot\|_2\big) \longrightarrow \big(V^2_{\Phi}, \|\cdot\|_2\big)$ given by

$$(\Gamma_{\mathbf{a}} f)(t) := (\det M) \sum_{j=1}^{s} \sum_{\alpha \in \mathbb{Z}^d} (\mathscr{L}_j f)(M\alpha) S_{j,\mathbf{a}}(t - M\alpha), \quad t \in \mathbb{R}^d,$$

is a well-defined bounded operator onto V^2_{Φ}. Besides, $\Gamma_{\mathbf{d}} f = f$ for all $f \in V^2_{\Phi}$.

Under appropriate hypotheses we prove that the scaled operator $\Gamma^h_{\mathbf{a}} := \sigma_{1/h} \Gamma_{\mathbf{a}} \sigma_h$ approximates, in the L^2-norm sense, any function f in the Sobolev space $W^{\ell}_2(\mathbb{R}^d)$ as $h \to 0^+$. Specifically we have the following:

Theorem 6. *Assume $2\ell > d$ and that all the systems $\mathscr{L}_j$ satisfy $\mathscr{L}_j f = f * \mathsf{h}_j$ with $\mathsf{h}_j \in \mathscr{L}^2(\mathbb{R}^d)$, $j = 1, \ldots, s$. Then,*

$$\|f - \Gamma^h_{\mathbf{a}} f\|_2 \leq (1 + \|\Gamma_{\mathbf{a}}\|) \inf_{g \in \sigma_{1/h} V^2_{\Phi}} \|f - g\|_2, \quad f \in W^{\ell}_2(\mathbb{R}^d),$$

where $\|\Gamma_{\mathbf{a}}\|$ denotes the norm of the sampling operator $\Gamma_{\mathbf{a}}$. If the set of generators $\Phi = \{\varphi_k\}_{k=1}^r$ satisfies the Strang–Fix conditions of order ℓ and, for each $k = 1, 2, \ldots, r$, the decay condition $\varphi_k(t) = O\big([1 + |t|]^{-d-\ell-\varepsilon}\big)$ for some $\varepsilon > 0$, then

$$\|f - \Gamma^h_{\mathbf{a}} f\|_2 \leq C\, |f|_{\ell,2}\, h^{\ell} \quad \text{for all } f \in W^{\ell}_2(\mathbb{R}^d),$$

where the constant C does not depend on h and f.

Proof. Using that $\Gamma^h_{\mathbf{a}} g = g$ for each $g \in \sigma_{1/h} V^2_{\Phi}$, then, for each $f \in L^2(\mathbb{R}^d)$ and $g \in \sigma_{1/h} V^2_{\Phi}$, Lebesgue's Lemma [13, p 30] gives

$$\|f - \Gamma^h_{\mathbf{a}} f\|_2 \leq \|f - g\|_2 + \|\Gamma^h_{\mathbf{a}} g - \Gamma^h_{\mathbf{a}} f\|_2 \leq (1 + \|\Gamma_{\mathbf{a}}\|) \inf_{g \in \sigma_{1/h} V^2_{\Phi}} \|f - g\|_2,$$

where we have used that $\|\Gamma^h_{\mathbf{a}}\| = \|\Gamma_{\mathbf{a}}\|$ for $h > 0$. Now, for each $f \in W^{\ell}_2(\mathbb{R}^d)$ and $h > 0$, there exists a function $f_h \in \sigma_{1/h} V^2_{\Phi}$ such that (3.38) holds, from which we obtain the desired result. □

More results on approximation by means of generalized sampling formulas can be found in [15, 18].

Acknowledgments The authors are very pleased to dedicate this work to Professor Gilbert. G. Walter on the occasion of his 80th birthday. Professor Walter's research and mentorship have, over the years, inspired and influenced many mathematicians throughout the world; we are fortunate to be three of these mathematicians.

This work has been supported by the grant MTM2009–08345 from the Spanish *Ministerio de Ciencia e Innovación* (MICINN).

References

1. Acosta-Reyes E, Aldroubi A, Krishtal I (2009) On stability of sampling-reconstruction models. Adv Comput Math 31:5–34
2. Adams RA, Fournier JJF (2003) Sobolev spaces. Academic Press, Amsterdam
3. Aldroubi A (2002) Non-uniform weighted average sampling and reconstruction in shift-invariant and wavelets spaces. Appl Comput Harmon Anal 13:151–161
4. Aldroubi A, Gröchenig K (2001) Non-uniform sampling and reconstruction in shift-invariant spaces. SIAM Rev 43:585–620
5. Aldroubi A, Sun Q, Tang W-S (2005) Convolution, average sampling, and a Calderon resolution of the identity for shift-invariant spaces. J Fourier Anal Appl 11:215–244
6. Aldroubi A, Unser M (1994) Sampling procedure in function spaces and asymptotic equivalence with Shannon's sampling theorem. Numer Funct Anal Optim 15:1–21
7. Aldroubi A, Unser M, Eden M (1992) Cardinal spline filters: stability and convergence to the ideal sinc interpolator. Signal Process 28:127–138
8. Boor C, DeVore R, Ron A (1994) Approximation from shift-invariant subspaces in $L^2(\mathbb{R}^d)$. Trans Am Math Soc 341:787–806
9. Boor C, DeVore R, Ron A (1994) The structure of finitely generated shift-invariant spaces in $L^2(\mathbb{R}^d)$. J Funct Anal 119:37–78
10. Chen W, Itoh S, Shiki J (2002) On sampling in shift invariant spaces. IEEE Trans Signal Process 48:2802–2810
11. Christensen O (2003) An introduction to frames and Riesz bases. Birkhäuser, Boston
12. Dahmen W, Han B, Jia RQ, Kunoth A (2000) Biorthogonal multiwavelets on the interval: cubic Hermite spline. Constr Approx 16:221–259
13. DeVore R, Lorentz G (1993) Constructive approximation. Springer, Berlin
14. García AG, Pérez-Villalón G (2006) Dual frames in $L^2(0,1)$ connected with generalized sampling in shift-invariant spaces. Appl Comput Harmon Anal 20:422–433
15. García AG, Pérez-Villalón G (2008) Approximation from shift-invariant spaces by generalized sampling formulas. Appl Comput Harmon Anal 24:58–69
16. García AG, Pérez-Villalón G (2009) Multivariate generalized sampling in shift-invariant spaces and its approximation properties. J Math Anal Appl 355:397–413
17. García AG, Hernández-Medina MA, Pérez-Villalón G (2008) Generalized sampling in shift-invariant spaces with multiple stable generators. J Math Anal Appl 337:69–84
18. García AG, Muñoz-Bouzo MJ, Pérez-Villalón G (2011) Regular multivariate sampling and approximation in L^p shift-invariant spaces. J Math Anal Appl 380:607–627
19. García AG, Pérez-Villalón G, Portal A (2005) Riesz bases in $L^2(0,1)$ related to sampling in shift invariant spaces. J Math Anal Appl 308:703–713
20. García AG, Kim JM, Kwon KH, Pérez-Villalón G (2008) Aliasing error of sampling series in wavelets subspaces. Numer Funct Anal Optim 29:126–144
21. García AG, Kim JM, Kwon KH, Yoon GJ (2012) Multi-channel sampling on shift-invariant spaces with frame generators. Int J Wavelets Multiresolut Inform Process 10(1):41–60
22. Gröchenig K (2001) Foundations of time-frequency analysis. Birkhäuser, Boston
23. Higgins JR (1996) Sampling theory in fourier and signal analysis: foundations. Oxford University Press, Oxford
24. Hogan JA, Lakey JD (2005) Sampling in principal shif-invariant spaces. In: Time-frequency and time-scale methods. Birkhäuser, Boston
25. Hong YM, Kim JM, Kwon KH, Lee EM (2007) Channeled sampling in shift-invariant spaces. Int J Wavelets Multiresolut Inform Process 5:753–767
26. Horn RA, Johnson CR (1999) Matrix analysis. Cambridge University Press, Cambridge
27. Jia RQ, Micchelli CA (1991) Using the refinement equations for the construction of pre-wavelets II: Powers of two. In: Laurent PJ, Le Méhauté A, Schumaker LL (eds) Curves and surfaces. Academic Press, New York, pp 209–246
28. Jia RQ, Shen Z (1994) Multiresolution and wavelets. Proc Edinbur Math Soc 37:271–300

29. Keinert F (2004) Wavelets and multiwavelets. Chapman & Hall/CRC, Boca Raton FL
30. Kang S, Kwon KH (2011) Generalized average sampling in shift-invariant spaces. J Math Anal Appl 377:70–78
31. Kang S, Kim JM, Kwon KH (2010) Asymmetric multi-channel sampling in shift-invariant spaces. J Math Anal Appl 367:20–28
32. Kim KH, Kwon KH (2008) Sampling expansions in shift-invariant spaces. Int J Wavelets Multiresolut Inform Process 6:223–248
33. Lei JJ, Jia RQ, Cheney EW (1997) Approximation from shift-invariant spaces by integral operators. SIAM J Math Anal 28:481–498
34. Nashed MZ, Sun Q, Tang W-S (2009) Average sampling in L^2. Compt Rend Acad Sci Paris, Ser I 347:1007–1010
35. Papoulis A (1977) Generalized sampling expansion. IEEE Trans Circuits Syst 24:652–654
36. Ron A, Shen Z (1995) Frames and stable bases for shift-invariant subspaces of $L_2(\mathbb{R}^d)$. Can J Math 47:1051–1094
37. Saitoh S (1997) Integral transforms, reproducing kernels and their applications. Longman, Essex
38. Selesnick IW (1999) Interpolating multiwavelet bases and the sampling theorem. IEEE Trans Signal Process 47:1615–1620
39. Šikić H, Wilson EN (2011) Lattice invariant subspaces and sampling. Appl Comput Harmon Anal 31:26–43
40. Sun Q (2006) Nonuniform average sampling and reconstruction of signals with finite rate of innovation. SIAM J Math Anal 38:1389–1422
41. Sun Q (2010) Local reconstruction for sampling in shift-invariant spaces. Adv Comput Math 32:335–352
42. Sun W (1999) On the stability of multivariate trigonometric systems. J Math Anal Appl 235:159–167
43. Sun W, Zhou X (1999) Sampling theorem for multiwavelet subspaces Chin Sci Bull 44:1283–1285
44. Unser M (1999) Splines: a perfect fit for signal and image processing. IEEE Signal Process Mag 16:22–38
45. Unser M (2000) Sampling 50 years after Shannon. Proc IEEE 88:569–587
46. Vaidyanathan PP (1993) Multirate systems and filter banks. Prentice-Hall, Englewood Hills
47. Walter GG (1992) A sampling theorem for wavelet subspaces. IEEE Trans Inform Theor 38:881–884
48. Wojtaszczyk P (1997) A mathematical introduction to wavelets. Cambridge University Press, Cambridge
49. Xian J, Li S (2007) Sampling set conditions in weighted multiply generated shift-invariant spaces and their applications. Appl Comput Harmon Anal 23:171–180
50. Xian J, Sun W (2010) Local sampling and reconstruction in shift-invariant spaces and their applications in spline subspaces. Numer Funct Anal Optim 31:366–386
51. Xian J, Luo SP, Lin W (2006) Weighted sampling and signal reconstruction in spline subspace. Signal Process 86:331–340
52. Zayed AI (1993) Advances in Shannon's sampling theory. CRC Press, Boca Raton
53. Zhou X, Sun W (1999) On the sampling theorem for wavelet subspaces. J Fourier Anal Appl 5:347–354

Chapter 4
Function Spaces for Sampling Expansions

M. Zuhair Nashed and Qiyu Sun

Abstract In this paper, we consider a variety of Hilbert and Banach spaces that admit sampling expansions $f(t) = \sum_{n=1}^{\infty} f(t_n)S_n(t)$, where $\{S_n\}_{n=1}^{\infty}$ is a family of functions that depend on the sampling points $\{t_n\}_{n=1}^{\infty}$ but not on the function f. Those function spaces, that arise in connection with sampling expansions, include reproducing kernel spaces, Sobolev spaces, Wiener amalgam space, shift-invariant spaces, translation-invariant spaces, and spaces modeling signals with finite rate of innovation. Representative sampling theorems are presented for signals in each of these spaces. The paper also includes recent results on nonlinear sampling of signals with finite rate of innovation, convolution sampling on Banach spaces, and certain foundational issues in sampling expansions.

4.1 Introduction

Series expansions and integral representations of functions and operators play a fundamental role in the analysis of *direct problems* of applied mathematics—witness the role of power series, Fourier series, Karhunen–Loève expansion, eigenfunction expansions of symmetric linear operators, and sampling expansions in signal processing; and the role of Fourier transform, spectral integral representations, and various integral representations in boundary value problems, potential theory, complex analysis, and other areas.

Expansion theorems also play a fundamental role in *inverse problems*. Two important problems discussed below are: (1) the recovery of a function from inner product with a given set of functions (i.e., the moment problem), and (2) the recovery of a function from its values on a subset of its domain (i.e., reconstruction of a function from its samples via sampling expansion).

M.Z. Nashed (✉) • Q. Sun
Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA
e-mail: zuhair.nashed@ucf.edu; qiyu.sun@ucf.edu

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_4,

One of the important problems in analysis is to expand a given function f in a separable Banach space by a series of the form

$$f(t) = \sum_{n=1}^{\infty} c_n f_n(t), \tag{4.1}$$

where $\{f_n\}_{n=1}^{\infty}$ is a suitable sequence of functions. This is not always possible. We consider two important cases of such an expansion:

1. Let H be a separable Hilbert space and $\{f_n\}_{n=1}^{\infty}$ be an orthonormal sequence in H. If $\{c_n\}_{n=1}^{\infty}$ is a sequence for which the right-hand side of the expansion (4.1) converges to f in H, then $c_n = \langle f, f_n\rangle$ for all $n \geq 1$, and hence $c_n, n \geq 1$, are the (generalized) Fourier coefficients. In this case, the series

$$f(t) = \sum_{n=1}^{\infty} \langle f, f_n\rangle f_n(t) \tag{4.2}$$

is called the *(generalized) Fourier expansion*. This series converges to f if and only if the orthonormal sequence $\{f_n\}$ is complete in the sense that the only function orthogonal to all the f_n's is the zero function. Equivalently, Parseval equality

$$\sum_{n=1}^{\infty} |\langle f, f_n\rangle|^2 = \|f\|^2$$

holds for every $f \in H$.

This expansion theorem is a *direct problem*: Given f, find its expansion. The associated inverse problem is the *moment problem*: determine f given moments $\langle f, f_n\rangle, n \in J$ (an index set). Given any sequence of real number $s_n, n \in J$, the existence of the moment problem is whether there exists a function f such that $s_n = \langle f, f_n\rangle, n \in J$, while the uniqueness is whether such a function f is uniquely determined by its moment sequence $s_n, n \in J$.

2. The second type of expansion that is the central theme of this chapter, is what is called a *sampling expansion*:

$$f(t) = \sum_{n=1}^{\infty} f(t_n) S_n(t), \tag{4.3}$$

where $\{S_n\}_{n=1}^{\infty}$ is called a *sampling sequence* and $\{t_n\}_{n=1}^{\infty}$ is the *sampling set*. The inverse problem for sampling is to determine f from given samples $f(t_n), n \geq 1$.

There have been many advances in sampling theory and its applications to signal and image processing. In the past three decades, many authors developed sampling theorems based on (1) the theory of regular and singular boundary value problems and (2) transforms other than the Fourier transform, including such transforms as the Sturm–Liouville transform, Jacobi transform, and Hankel transform, see [43]. Another main thrust has been in nonuniform sampling for non-bandlimited signals.

In the past 20 years, there have been major advances in sampling theory and its foundational aspects, where methods of functional analysis and harmonic analysis have played pivotal roles. In particular, new directions in sampling theory have been pursued using various function spaces that admit sampling expansions, such as reproducing kernel Hilbert (Banach) spaces, Sobolev spaces, Wiener amalgam space, shift-invariant spaces, translation-invariant spaces, and spaces modeling signals with finite rate of innovation. Another direction of research on sampling in the past decade involves average sampling, convolution sampling, nonlinear sampling, and other fundamental issues in sampling theory. The reader may refer to [1, 6, 7, 9–11, 15–24, 30, 33, 37, 39, 43] for various prospectives on these advances.

The purpose of this chapter is to consider a variety of function spaces mentioned above in which every function admits the sampling expansion (4.3). Representative sampling expansions are presented for signals in each of these spaces. This chapter also includes recent results on nonlinear sampling for signals with finite rate of innovation, convolution sampling on Banach spaces, and certain foundational issues in sampling expansions.

4.2 Fourier Series/Fourier Integral Approach to the Whittaker, Shannon, and Kotel'nikov Sampling Theorem

Let f be a bandlimited signal with finite energy; i.e.,

$$f(t) = \int_{-\Omega}^{\Omega} F(\omega) \mathrm{e}^{i\omega t} \mathrm{d}\omega,\ t \in (-\infty, \infty), \tag{4.4}$$

for some square-integrable function F on $[-\Omega, \Omega], \Omega > 0$. We extend $F(\omega)$ periodically to the real line and expand the extension in complex Fourier series:

$$F(\omega) = \sum_{n=-\infty}^{\infty} c_n \exp(in\pi\omega/\Omega),\ |\omega| < \Omega, \tag{4.5}$$

where

$$c_n = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\omega) \exp(-in\pi\omega/\Omega) \mathrm{d}\omega. \tag{4.6}$$

Comparing (4.4) and (4.6) leads to

$$c_n = \frac{1}{2\Omega} f(-n\pi/\Omega). \tag{4.7}$$

Substituting (4.7) in (4.5) gives

$$F(\omega) = \frac{1}{2\Omega} \sum_{n=-\infty}^{\infty} f(-n\pi/\Omega) \exp(in\pi\omega/\Omega)$$

$$= \frac{1}{2\Omega} \sum_{n=-\infty}^{\infty} f(n\pi/\Omega) \exp(-in\pi\omega/\Omega),\ |\omega| < \Omega.$$

Substituting this in (4.4) and interchanging the order of integration and summation leads to

$$f(t) = \frac{1}{2\Omega} \sum_{n=-\infty}^{\infty} f(n\pi/\Omega) \int_{-\Omega}^{\Omega} \exp(-in\pi\omega/\Omega) \exp(i\omega t) \mathrm{d}\omega,$$

which yields the celebrated classical expansion of a bandlimited signal f:

$$f(t) = \sum_{n=-\infty}^{\infty} f(n\pi/\Omega) \frac{\sin(\Omega t - n\pi)}{\Omega t - n\pi}. \tag{4.8}$$

This can be simplified to

$$f(t) = \sin(\Omega t) \sum_{n=-\infty}^{\infty} f(n\pi/\Omega) \frac{(-1)^n}{\Omega t - n\pi}.$$

This classical proof is rigorous. The interchange of integration and summation can be easily justified. But the proof is not very revealing: We perform this interchange and a theorem pops up. A theorem is born, but *we did not hear the heartbeat of its proof.*

4.3 Properties of the Sinc Function and the Paley–Wiener Space

Let B_π consist of all signals that are bandlimited to $[-\pi, \pi]$ (i.e., $\Omega = \pi$ in (4.4)) and have finite energy. The space B_π is the same as the Paley–Wiener space of restrictions to the real axis of entire functions of exponential type π. The Paley–Wiener space B_π has many interesting properties that are not exploited or used in the classical proof of the Whittaker–Shannon–Kotel'nikov sampling theorem. Some of these properties are stated in the following theorem.

Theorem 1. *(i) B_π is a reproducing kernel space with kernel*

$$\mathrm{sinc}(s-t) := \frac{\sin \pi(s-t)}{\pi(s-t)}.$$

(ii) *The sequence* $\{S_n\}_{n\in\mathbf{Z}}$, *where* $S_n(t) := \operatorname{sinc}(t-n)$, *is an orthonormal basis for* B_π.

(iii) *The sequence* $\{S_n\}_{n\in\mathbf{Z}}$ *has the discrete orthogonality property:*

$$S_n(m) = \delta_{mn} := \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases}$$

(iv) $f(\cdot - c) \in B_\pi$ *and* $\|f(\cdot - c)\|_2 = \|f\|_2$ *for all* $f \in B_\pi$ *and* $c \in \mathbf{R}$. *Hence* B_π *is unitarily translation-invariant subspace of* $L^2(\mathbf{R})$.

(v) B_π *is a shift-invariant subspace of* $L^2(\mathbf{R})$ *generated by the sinc function:*

$$B_\pi = \left\{ \sum_{n\in\mathbf{Z}} c(n)\, \operatorname{sinc}(t-n) : \sum_{n\in\mathbf{Z}} |c(n)|^2 < \infty \right\}.$$

Proof. (i) Take $f \in B_\pi$, and let F be the square-integrable function in (4.4). Then

$$f(t) = \int_{-\pi}^{\pi} \mathrm{e}^{it\omega} F(\omega)\mathrm{d}\omega,\ t \in (-\infty,\infty). \tag{4.9}$$

This implies that

$$|f(t)| \le \int_{-\pi}^{\pi} |F(\omega)| d\omega \le (2\pi)^{1/2} \|F\|_2 = \|f\|_2$$

and

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{it\omega} \left(\int_{-\infty}^{\infty} f(s)\mathrm{e}^{-is\omega}\mathrm{d}s \right) \mathrm{d}\omega \\ &= \int_{-\infty}^{\infty} f(s) \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{i(t-s)\omega}\mathrm{d}\omega \right) \mathrm{d}s \\ &= \int_{-\infty}^{\infty} f(s) \frac{\sin\pi(s-t)}{\pi(s-t)}\mathrm{d}s,\ t \in (-\infty,\infty). \end{aligned}$$

Hence B_π is a reproducing kernel Hilbert space with kernel $\frac{\sin\pi(s-t)}{\pi(s-t)}$.

(ii) By the reproducing property and symmetry of the sinc kernel $k(s,t) := \frac{\sin\pi(s-t)}{\pi(s-t)}$, we have that $\langle S_n, S_m \rangle = k(n,m)$, which takes value one if $m = n$ and zero if $m \neq n$. Hence $S_n, n \in \mathbf{Z}$, is an orthonormal set. The completeness of the orthonormal set $\{S_n\}_{n\in\mathbf{Z}}$ follows from (4.8).

(iii) The discrete orthogonality property is obvious.

(iv) Take $f \in B_\pi$, and let F be the square-integrable function in (4.4). Then it follows from (4.9) that for any $c \in (-\infty,\infty)$,

$$f(t-c) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{i(t-c)\omega} F(\omega)\mathrm{d}\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{it\omega} F_c(\omega)\mathrm{d}\omega,$$

where $F_c(\omega) = \mathrm{e}^{-ic\omega}F(\omega)$ is square-integrable. Hence $f(t-c) \in B_\pi$ for all $t \in (-\infty, \infty)$.

(v) This follows from the conclusion that $\{S_n\}_{n\in\mathbf{Z}}$ is an orthonormal basis for the Paley–Wiener space B_π. □

4.4 The Engineering Approach to Sampling and Its Mathematical Deficiencies

We now turn to the engineering clever approach to the sampling theorem, see, e.g., [13, 30, 32]. We paraphrase some of the description in [13]. Let us consider what happens when we sample $f(t)$ at uniformly spaced times. If the sampling frequency is f_s, then we can model this with a multiplication of $f(t)$ by a train of Dirac impulsed spaces $T_s := 1/f_s$ second apart:

$$f^*(t) := f(t)T_s \sum_{n=-\infty}^{\infty} \delta(t - nT_s) = T_s \sum_{n=-\infty}^{\infty} f(nT_s)\delta(t - nT_s).$$

Mathematicians consider the sequence of sampled values $\{f(nT_s)\}_{n=-\infty}^{\infty}$ as a vector in ℓ_2 (the space of all square-summable vectors). Electrical engineers like to continue to think of this sequence as a time signal. So to stay in the analog world, they use their beloved "Dirac impulse" as above. Informally, as they assert, "multiplication by an impulse train" in the time domain corresponds to convolution with an impulse train in the frequency domain. If the Fourier transform of the sampled sequence $f^*(t)$ is $F^*(\omega)$, then

$$F^*(\omega) = F(\omega) * \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_s) = \sum_{n=-\infty}^{\infty} F(\omega - n\omega_s).$$

Hence, again informally, the Fourier transform of the samples (considered as a time signal in the sense of the above representation) is an infinitely repeated replication of $F(\omega)$ at intervals of $\omega_s := 2\pi f_s$. The portion of the transform between $-\omega_s/2$ and $\omega_s/2$ is called the base band and all the other replication images. If $f(t)$ is bandlimited so that $F(\omega)$ is zero for $|\omega| > 2\pi f_c$, and if $f_s \geq 2f_c$ (Shannon's rate), then there is no overlap between successive replications. We have lost no information in the sampling process; if anything, we have picked up a lot of "superfluous information" at frequencies outside the range of interest.

To recover the original signal, we must remove the replication images. First $F(\omega)$ can be obtained from $F^*(\omega)$ by multiplying it by the characteristic function

$$\chi_{\omega_s/2}(\omega) = \begin{cases} 1 & \text{if } |\omega| \leq \omega_s/2, \\ 0 & \text{if } |\omega| > \omega_s/2. \end{cases}$$

This is done by an analog filter known as an interpolation filter or a low-pass filter. We are now back to $F(\omega)$, and the time function can be recovered via the inverse Fourier transform:

$$f(t) = \int_{-\infty}^{\infty} F(\omega) \mathrm{e}^{i\omega t} \mathrm{d}\omega.$$

In essence engineers view the Shannon sampling theorem in terms of an impulse train and a low-pass filter in the following way [30, 32]:

$$f(t) \overset{\text{sampler}}{\longmapsto} f^*(t) \overset{\text{low pass filter}}{\longmapsto} f(t).$$

First the signal $f \in B_\pi$ is sampled at the integers to convert $f(t)$ into the impulse train

$$f^*(t) = \sum_{n=-\infty}^{\infty} f(n)\delta(t-n),$$

where $\delta(t-n)$ is the Dirac delta function (impulse) at $t = n$. This is still expressed as an analog signal. This then is transmitted through an ideal low-pass filter which passes all frequencies of absolute value less than π and blocks all others. This converts $f^*(t)$ back into $f(t)$.

We denote the sampling map by $S : f \longmapsto f^*$ and the low-pass filter map by $P : f^* \longmapsto f$

$$B_\pi \overset{\text{S}}{\longmapsto} \boxed{\text{unknown space}} \overset{\text{P}}{\longmapsto} B_\pi.$$

But the above procedure has some mathematical difficulties which are not resolved in the engineering approach:

- The sampler S takes f into f^*, which is out of the space of bandlimited functions; indeed, f^* is not a signal with finite energy.
- In what sense does the impulse train series converge? One may prove that the series converges in the sense of tempered distributions to a generalized function.
- The map P recovers f at least formally since

$$P(\delta(t-n)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{-i\omega t} \mathrm{e}^{i\omega n} \mathrm{d}\omega = \frac{\sin \pi(t-n)}{\pi(t-n)}.$$

  If P is continuous (in some topology), then

$$(Pf^*)(t) = \sum_{n=-\infty}^{\infty} f(n) P(\delta(t-n)) = \sum_{n=-\infty}^{\infty} f(n) \frac{\sin \pi(t-n)}{\pi(t-n)}.$$

  However, since $\mathscr{S}'$ (the space of all tempered distributions, see the next section) is not a Hilbert space, we do not know if P is a continuous operator.
- Still another difficulty! P is not even well defined on $\mathscr{S}'$. Indeed,

$$Pg = \mathscr{F}^{-1}(\chi_{[-\pi,\pi]} \hat{g}),\ g \in \mathscr{S}',$$

where $\mathscr{F}^{-1}$ is the inverse Fourier transform of a tempered distribution. So P corresponds under the Fourier transform to the multiplication of the Fourier transform $\hat{g}$ in $\mathscr{S}'$ by the characteristic function of $[-\pi,\pi]$. Unfortunately the characteristic function is not a multiplier in $\mathscr{S}'$. Hence we need to restrict ourselves to a subspace of $\mathscr{S}'$ in which $\chi_{[-\pi,\pi]}$ is a multiplier.

These issues have been resolved by Nashed and Walter [28]. They obtained a rigorous proof of the engineering approach, by considering the sampling map S as an operator from B_π to H^{-1} and the filtering map P as an operator (actually an orthogonal projection) from H^{-1} to B_π:

$$B_\pi \overset{\mathrm{S}}{\longmapsto} H^{-1} \overset{\mathrm{P}}{\longmapsto} B_\pi,$$

where H^{-1} is a Sobolev space, see the next section for the definition of Sobolev spaces. More importantly, by emulating and extending this proof, they obtained a general unifying approach for sampling theorems in reproducing kernel Hilbert spaces (RKHS) that include many earlier sampling theorems.

4.5 Function Spaces: Tempered Distributions, Sobolev Spaces and Reproducing Kernel Hilbert Spaces

4.5.1 The Space of Tempered Distributions

Let $\mathscr{S}$ be the space of all rapidly decreasing C^∞ functions on the real line $\mathbf{R}$, i.e., functions that satisfy

$$|g^{(k)}(t)| \le C_{p,k}(1+|t|)^{-p},\ t \in \mathbf{R},$$

for all $p,k = 0,1,2,\ldots$. Convergence on $\mathscr{S}$ may be defined by endowing $\mathscr{S}$ with the seminorms:

$$\mu_{p,k} := \sup_{t\in\mathbf{R}}(1+|t|)^p|g^{(k)}(t)|.$$

Then $g_n \to g$ in $\mathscr{S}$ whenever

$$(1+|t|)^p\big(g_n^{(k)}(t) - g^{(k)}(t)\big) \to 0$$

uniformly in $t \in \mathbf{R}$ for each p and $k \ge 0$ as $n \to \infty$. The set $\mathscr{S}$ is dense in $L^2 := L^2(\mathbf{R})$ (the space of all square-integrable functions on the real line). We observe that compactly supported C^∞ functions are contained in $\mathscr{S}$, and that the space $\mathscr{S}$ is complete with respect to the convergence of semi-norms $\mu_{p,k}, p,k \ge 0$.

A tempered distribution is an element in the dual space $\mathscr{S}'$ of $\mathscr{S}$, i.e., $\mathscr{S}'$ consists of all continuous linear functionals on $\mathscr{S}$. The definition of Fourier transform

$$\mathscr{F}f(\omega) := \int_{-\infty}^{\infty} f(t)\mathrm{e}^{-it\omega}\mathrm{d}t, \quad f \in \mathscr{S}$$

may be extended from $\mathscr{S}$ to $\mathscr{S}'$. The following examples of Fourier transforms on $\mathscr{S}'$ are needed in the derivation of a rigorous setting for the proof of the engineering approach:

$$\mathscr{F}\big(\delta(t-\alpha)\big) = \mathrm{e}^{-i\omega\alpha},$$

$$\mathscr{F}\left(\sum_{n\in\mathbf{Z}} \delta(t-2\pi n)\right) = \sum_{n\in\mathbf{Z}} \delta(\omega-n),$$

$$\mathscr{F}\left(\sum_{n\in\mathbf{Z}} a_n \mathrm{e}^{int}\right) = 2\pi \sum_{n\in\mathbf{Z}} a_n \delta(\omega-n),$$

and

$$\mathscr{F}\left(\sum_{n\in\mathbf{Z}} a_n \mathrm{e}^{int}\chi_{[-\pi,\pi]}(t)\right) = \sum_{n\in\mathbf{Z}} a_n \frac{\sin\pi(\omega-n)}{\pi(\omega-n)},$$

where $\{a_n\}_{n\in\mathbf{Z}} \in \ell^2$.

4.5.2 Sobolev Spaces

An important Hilbert space structure on certain subsets of $\mathscr{S}'$ is provided by a class of Sobolev spaces. For $r \in \mathbf{R}$, the Sobolev space H^r consists of all tempered distributions $f \in \mathscr{S}'$ such that

$$\int_{-\infty}^{\infty} |\hat{f}(\omega)|^2(\omega^2+1)^r \mathrm{d}\omega < \infty.$$

The inner product of f and g in H^r is defined by

$$\langle f,g\rangle_r := \int_{-\infty}^{\infty} \hat{f}(\omega)\overline{\hat{g}(\omega)}(\omega^2+1)^r \mathrm{d}\omega,$$

where $\hat{f} := \mathscr{F}f$ is the Fourier transform of f. The space H^r is complete with respect to this inner product. For $r=0$, H^0 is just $L^2(\mathbf{R})$ by Parseval identity. For $r=1,2,\ldots$, H^r is the usual Sobolev space of functions that are $(r-1)$-times differentiable and whose rth derivative is in $L^2(\mathbf{R})$. For $r=-1,-2,\ldots$, H^r contains all tempered distributions with point support of order r. Thus the Dirac delta $\delta \in H^{-1}$, and δ', the distributional derivative of δ, belongs to H^{-2}.

4.5.3 Reproducing Kernel Hilbert Spaces

A Hilbert space H of complex-valued functions on a set Ω is called a RKHS if all the evaluation functionals $H \ni f \longmapsto f(t) \in \mathbf{C}$ are continuous (bounded) for each fixed $t \in \Omega$; i.e., there exist a positive constant C_t for each $t \in \Omega$ such that $|f(t)| \le C_t\|f\|$ for all $f \in H$.

By Riesz representation theorem, for each $t \in \Omega$ there exists a unique element k_t such that $f(t) = \langle f, k_t \rangle$ for all $f \in H$. The *reproducing kernel* $k(\cdot,\cdot) : \Omega \times \Omega \longmapsto \mathbf{C}$ of a RKHS H is defined by $k(s,t) = \langle k_s, k_t \rangle, s,t \in \Omega$.

We summarize some basic properties of RKHS that are particularly relevant to signal processing, wavelet analysis, and approximation theory:

- $k(s,t) = \overline{k(t,s)}$ for all $t,s \in \Omega$.
- $k(s,s) \geq 0$ for all $s \in \Omega$. Furthermore, if $k(t_0,t_0) = 0$ for some $t_0 \in \Omega$, then $f(t_0) = 0$ for all $f \in H$.
- $|k(s,t)| \leq \sqrt{k(s,s)}\sqrt{k(t,t)}$ for all $s,t \in \Omega$.
- The reproducing kernel $k(s,t)$ on $\Omega \times \Omega$ is a nonnegative definite symmetric kernel. Conversely by Aronszajn–Moore theorem, every nonnegative definite symmetric function $k(\cdot,\cdot)$ on $\Omega \times \Omega$ determines a unique Hilbert space H_k for which $k(\cdot,\cdot)$ is a reproducing kernel [5]. Here a complex-valued function F on $\Omega \times \Omega$ is said to be *positive definite* if for any n points $t_1, \ldots, t_n \in \Omega$, the matrix $A := (F(t_i,t_j))_{1 \leq i,j \leq n}$ is nonnegative definite, i.e., $u^*Au = \sum_{i,j=1}^n \overline{u_i} F(t_i,t_j) u_j \geq 0$ for all $u = (u_1, \ldots, u_n) \in \mathbf{C}^n$.
- A closed subspace $\tilde{H}$ of a RKHS H is also a RKHS. Moreover, the orthogonal projector P of H onto $\tilde{H}$ and the reproducing kernel $\tilde{k}(s,t)$ of the RKHS $\tilde{H}$ are related by $Pf(s) = \langle f, \tilde{k}_s \rangle, s \in \Omega$ for all $f \in H$ where $\tilde{k}_s = Pk_s$.
- If a RKHS space H with kernel $k(\cdot,\cdot)$ has direct orthogonal decomposition $H = H_1 \oplus H_2$ for some complementary orthogonal closed subspaces H_1 and H_2, then $k = k_1 + k_2$, where k_1, k_2 are reproducing kernels of the reproducing kernel Hilbert spaces H_1 and H_2, respectively.
- In a RKHS, the element representing a given bounded linear functional ϕ can be expressed by means of the reproducing kernel: $\phi(f) = \langle f, h \rangle$, where $h = \phi(k)$. Similarly for a bounded linear operator L on H to H, we have that $Lf(t) = \langle Lf, h \rangle = \langle f, L^*h \rangle$.
- Every finite-dimensional function space is a RKHS H with reproducing kernel $k(s,t) = \sum_{i=1}^n u_i(s)\overline{u_i(t)}$, where $\{u_i\}_{i=1}^n$ is an orthonormal basis for H. (Notice that the sum in the above definition of the kernel k is invariant under the choice of orthonormal basis).
- The space W^1, that contains all functions $f \in L^2[0,1]$ such that f is absolutely continuous and f', which exists almost everywhere, is in $L^2[0,1]$, and $f(0) = 0$, is a RKHS with kernel $k(s,t) = \min(s,t)$ under the inner product $\langle f, g \rangle = \int_0^1 f(t)g(t)\mathrm{d}t$.
- Sobolev space $H^s, s > 1/2$, is a reproducing kernel Hilbert space.
- Let H be a separable RKHS, then its reproducing kernel $k(\cdot,\cdot)$ has the expansion:

$$k(s,t) = \sum_{n=1}^{\infty} \varphi_n(t)\overline{\varphi_n(s)},$$

where $\{\phi_n\}_{n=1}^{\infty}$ is an orthonormal basis for H. We remark that for a general separable Hilbert space H, $\sum_{n=1}^{\infty} \varphi_n(t)\overline{\varphi_n(s)}$ is not a reproducing kernel and also

that ϕ_n's do not generally corresponds to sampling expansions. If they do, i.e.,if $\varphi_n(t) = k(t_n, t)$ for some sequence $\{t_n\}$, then we have that $f(t) = \sum_{n=1}^{\infty} f(t_n)\varphi_n(t)$, this constitutes a sampling theorem.

- If the reproducing kernel $k(s,t)$ of a RKHS H is continuous on $\Omega \times \Omega$, then H is a space of continuous functions (uniformly continuous on a bounded Ω). This follows from
$$|f(t) - f(s)| = |\langle f, k_t - k_s\rangle| \leq \|f\|\|k_t - k_s\|$$
and $\|k_t - k_s\|^2 = k(t,t) - 2k(t,s) + k(s,s)$ for all $s,t \in \Omega$.
- Strong convergence in a RKHS H implies pointwise convergence and uniform convergence on compact sets, because
$$|f(t) - f_n(t)| = |\langle f - f_n, k_t\rangle| \leq \|f - f_n\|\sqrt{k(t,t)}.$$
- $L_2[a,b]$, the space of all square-integrable functions on the interval $[a,b]$, is not a RKHS. Indeed, point evaluation is not well defined. Each function $f \in L^2[a,b]$ is actually an equivalence class of functions equal to each other almost everywhere. Thus the"value" at a point has no meaning since any point has measure zero.

4.6 Rigorous Justification of the Engineering Approach to Sampling

This section involves a search for function spaces in which the mathematical difficulties described in Sect. 4.3 are resolved. Clearly we want to work with a *subspace* H of the space $\mathscr{S}'$ of tempered distributions. We require that $\delta \in H$ and the convolution signal of the impulse train must converge in H under a mild condition on $\{f(t_n)\}_{n=-\infty}^{\infty}$. As remarked in Sect. 4.4, the characteristic function is not a multiplier in the space $\mathscr{S}'$, and hence the space H must also have the property that the characteristic function is a multiplier in H. Finally, the sampling map S and the low-pass filter map P must be well defined on the appropriate spaces, and their composition PS is the identity map:

- The characteristic functions are multipliers in the space of Fourier transforms of the elements in $H^r(\mathbf{R})$, which can be identified with the image under the Fourier transform of $L^2_r(\mathbf{R})$, the space of square-integrable functions with respect to the measure $d\mu(x) = (1 + |x|^2)^r dx$. So we may consider H^r with $r > -1/2$; specifically we take H^{-1}, since $\delta \in H^{-1}$.
- We consider the sampling map $Sf = f^*$ as a map onto H^{-1}. Then the partial sums of the sampled impulse train $\sum_{n=-N}^{N} f(n)\delta(t-n)$ belong to H^{-1} and converges in the norm of H^{-1} to f^*. This proposition does not require the signal to be bandlimited, but the next result does.

- When we consider projector onto the space H^{-1}, $B_\pi \subset L_2 \subset H^{-1}$ and B_π is closed in the topology of H^{-1}. Hence we can define the orthogonal projection of H^{-1} onto B_π. The reproducing kernel of B_π enables us to compute this projection easily. In fact,

$$P\delta_a(t) = k(t,a),\ t \in \mathbf{R}$$

and

$$Pf^* = \lim_{N\to\infty} \sum_{n=-N}^{N} f(n)k(\cdot,n)$$

in the norm of H^{-1} [28]. The above series converges to $Pf^*(t)$ by the continuity of the orthogonal projector and to $f(t)$ by the sampling theorem. Thus $Pf^* = f$.

The above ideas provide a mathematical proof of the engineering arguments, but they also suggest important extensions to other RKHS, as discussed in [28].

4.7 Sampling in Reproducing Kernel Hilbert Spaces

At the outset, the expansions (4.2) and (4.3) mentioned in introduction section appear markedly different, or at least not seem to be related. The expansion (4.2) holds for any complete orthonormal sequence in a separable Hilbert space, and sampling points $\{t_n\}$ could have no meaning in this context. On the other hand, the sampling expansion does not require the sampling sequence $\{S_n(t)\}_{n=1}^{\infty}$ to be orthonormal.

The expansions (4.2) and (4.3) can indeed be related. In finite-dimensional spaces the expansions are based on different choices of orthonormal basis. One with respect to inner product of two continuous functions, and the other is based on discrete orthogonality or biorthogonality of the sequences.

The expansions (4.2) and (4.3) may also be related in some special cases of orthonormal sequences in certain infinite-dimensional spaces. Let f be a signal defined on the real line. Suppose that there exists a reproducing kernel $k(t,s)$ such that for some real numbers $\{t_n\}_{n=1}^{\infty}$, the sequence $f_n(s) = k(t_n,s), n \geq 1$, is a complete orthonormal sequence, then $c_n = \langle f, f_n\rangle = f(t_n)$. The series expansion $\sum_{n=1}^{\infty} c_n f_n$ then becomes a sampling expansion, i.e., it states how to recover $f(t)$ from the sample values $\{f(t_n)\}_{n=1}^{\infty}$. For example, for π-bandlimited signals, $k(t,s) = \frac{\sin\pi(t-s)}{\pi(t-s)}$ and $S_n(t) = k(t,n)$ is an orthonormal basis for B_π, and the expansion result mentioned above reduces to the Whittaker–Shannon–Kotel'nikov sampling expansion [42].

In [28], the authors introduced an approach that provided general sampling theorem for functions in RKHS. The sampling theorems are for functions in a general RKHS with reproducing kernel $k(t,s)$, which is closed in the Sobolev space $H^{-p}, p > 1/2$. The sampling functions $S_n, n \geq 1$, need not be an orthogonal system, and the theory allows nonorthogonal sampling sequences. Then the system

$\{S_n := k(t_n,\cdot)\}_{n=1}^{\infty}$ has to satisfy a biorthogonality condition, i.e., $S_n(t_m) = \delta_{mn}$ for all $m,n \geq 1$, and the sampling points must satisfy a density-type condition. This general setup includes sampling theorems related to other transforms than the Fourier transform as in the classical theory. For example, Sturm–Liouville, Jacobi, and Laguerre transforms are among the examples discussed in [28], as well as sampling using frames. Also for the orthogonal case, several error analyses, such as truncation, aliasing, jittering, and amplitude error, are discussed in details.

Now we state a representative sampling theorem for signals in a reproducing kernel Hilbert space [28].

Theorem 2. *Let $H \subset L^2(\mathbf{R})$ be a reproducing kernel Hilbert space that is closed in the Sobolev space H^{-1} and under differentiation. Assume that its reproducing kernel $k(s,t)$ is continuous and has the zero sequence $\{t_n\}$ which is a set of uniqueness for H, and assume that $\{t_n\}$ tends to infinity as n tends to infinity. If $f \in H$ satisfies $f(t)/k(t,t) = O(t^{-2})$, then the sampled sequence*

$$f^*(t) = \sum_n \frac{f(t_n)}{k(t_n,t_n)}\delta(t-t_n)$$

converges in the sense of H^{-1} and its orthogonal projection onto H equals to $f(t)$, and the series

$$f(t) = \sum_n \frac{f(t_n)}{k(t_n,t_n)}k(t_n,t)$$

converges uniformly on sets for which $k(t,t)$ is bounded.

4.8 Sampling in Shift-Invariant Spaces

A *shift-invariant space* generated by a square-integrable function ϕ is given by

$$V_2(\phi) := \left\{ \sum_{n\in\mathbf{Z}} c_n\phi(\cdot - n) : \sum_{n\in\mathbf{Z}} |c_n|^2 < \infty \right\}.$$

Shift-invariant spaces have been shown to be realistic for modeling signals with smoother spectrum, and also suitable for taking into account real acquisition and reconstruction devices. The notion of shift-invariant spaces arises in approximation theory, wavelet theory, and sampling theory.

For the generator ϕ of a shift-invariant space, we usually assume that $\{\phi(\cdot - n) : n \in \mathbf{Z}\}$ consisting of all integer shifts of the generator ϕ is a Riesz basis for $V_2(\phi)$; i.e., there exist positive constants A and B such that

$$A\sum_{n\in\mathbf{Z}} |c_n|^2 \leq \left\| \sum_{n\in\mathbf{Z}} c_n\phi(\cdot - n)\right\|_2^2 \leq B\sum_{n\in\mathbf{Z}} |c_n|^2.$$

The Paley–Wiener space B_π is a shift-invariant space generated by the sinc function $\mathrm{sinc}(t)$, and the integer shifts of the sinc function form an orthonormal basis for the Paley–Wiener space B_π. The following is a representative sampling theorem for signals in a shift-invariant space $V_2(\phi)$ established in [40].

Theorem 3. *Let ϕ be a real continuous function such that $\sup_{t\in\mathbf{R}}|\phi(t)|(1+|t|)^{1+\varepsilon}<\infty$ for some $\varepsilon>0$, $\hat{\phi}^*(\omega):=\sum_{n\in\mathbf{Z}}\phi(n)\mathrm{e}^{-in\omega}\neq 0$ for all $\omega\in\mathbf{R}$, and $\{\phi(t-n):\, n\in\mathbf{Z}\}$ is an orthonormal basis for $V_2(\phi)$. Then any signal $f\in V_2(\phi)$ can be stably reconstructed from its samples $\{f(n)\}_{n\in\mathbf{Z}}$ on the integer lattice. Moreover,*

$$f(t)=\sum_{n\in\mathbf{Z}} f(n)\tilde{\phi}(t-n)\quad \text{for all } f\in V_2(\phi),$$

where $\tilde{\phi}\in V_2(\phi)$ is defined by $\hat{\tilde{\phi}}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^*(\omega), \omega\in\mathbf{R}$.

The reader may refer to [1, 4, 35, 37, 41] and references therein for some fundamental issues to sampling theory in shift-invariant spaces.

4.9 Sampling in Unitarily Translation-Invariant Hilbert Spaces

In this section, we consider sampling theorems on a unitarily translation-invariant RKHS generated from a single function. To be more specific, the RKHS H_ϕ has the reproducing kernel

$$k_\phi(t,s)=\int_{-\infty}^{\infty}\phi(u-t)\phi(u-s)\mathrm{d}u \tag{4.10}$$

generated by a function $\phi\in L^1(\mathbf{R})\cap L^2(\mathbf{R})$, whose Fourier transform does not have real zeros. Examples of such a generating function ϕ includes $(\sigma^2+t^2)^{-1}$, $\mathrm{e}^{-\sigma^2t^2}$, and $\mathrm{e}^{-\sigma|t|}$ where $\sigma>0$. The RKHS H_ϕ with reproducing kernel k_ϕ in (4.10) is given by

$$H_\phi=\{f:\|f\|_{H_\phi}<\infty\}, \tag{4.11}$$

where

$$\|f\|_{H_\phi}:=\left(\frac{1}{2\pi}\int_{\mathbf{R}}|\hat{f}(\omega)|^2/|\hat{\phi}(\omega)|^2\mathrm{d}\omega\right)^{1/2}.$$

Theorem 4 ([38]). *Let ϕ be an integrable function on the real line such that its Fourier transform $\hat{\phi}$ does not have real zeros and $\int_{\mathbf{R}}|\phi(t)|^2(1+t^2)^\alpha\mathrm{d}t<\infty$ for some $\alpha>1$, k_ϕ be the reproducing kernel in (4.10), and $\cdots<\lambda_{-2}<\lambda_{-1}<\lambda_0=0<\lambda_1<\lambda_2<\cdots$ be sampling points with $\lambda_{j+1}-\lambda_j\geq\varepsilon>0$ for all $j\in\mathbf{Z}$. Denote by $\mathscr{X}$ the closed subspace of the reproducing kernel Hilbert space H_ϕ in (4.11) spanned by $k_\phi(\cdot,t_j), j\in\mathbf{Z}$. Then the sampling operator*

$$\mathscr{X}\ni f\longmapsto (f(t_j))_{j\in\mathbf{Z}}$$

is stable in the sense that there exist positive constants A and B such that

$$A\|f\|_{H_\phi} \le \|(f(t_j))_{j\in\mathbf{Z}}\|_2 \le B\|f\|_{H_\phi} \quad \text{for all } f \in \mathscr{X}.$$

Moreover, the sampling expansion

$$f(t) = \sum_{j\in\mathbf{Z}} f(t_j)\frac{k_\phi(t,t_j)}{k_\phi(0,0)}$$

is valid for all $f \in \mathscr{X}$.

4.10 Sampling Signals with Finite Rate of Innovation

Signals with *finite rate of innovation* are those signals that can be determined by finitely many samples per unit of time [39]. The concept of signals with finite rate of innovation was introduced and studied by Martin Vetterli and his school. Prototype examples of signals with finite rate of innovation include delta pulses, narrow pulses in ultrawide band communication, mass spectrometry data in medical diagnosis, and splines with (non-)uniform knots. They also include bandlimited signals and time signals in shift-invariant spaces, which are discussed in the previous sections.

A common feature of signals with finite rate of innovation is that they have a parametric representation with a finite number of parameters per unit time. So we may model a signal f with finite rate of innovation as a superposition of impulse response of varying positions, amplitudes, and widths [34], i.e.,

$$f(t) = \sum_{\lambda\in\Lambda} c_\lambda \phi_\lambda(t-\lambda), \tag{4.12}$$

where each $\lambda \in \Lambda$ represents the innovative position of the signal, ϕ_λ is the impulse response of the signal-generating device at the innovative position λ, and c_λ is the amplitude of the signal at the innovation position λ. Thus the function space

$$V_p(\Phi) := \left\{ \sum_{\lambda\in\Lambda} c(\lambda)\phi_\lambda(\cdot-\lambda) :\ (c(\lambda))_{\lambda\in\Lambda} \in \ell^p(\Lambda) \right\}, 1 \le p \le \infty, \tag{4.13}$$

could be suitable for modeling signals with finite rate of innovation.

Sampling theory for signals with finite rate of innovation has been demonstrated to be important for accurate time estimation of ultraband communication, registration of multiview images, pattern recognition, quantification of spectra, etc. The following is a sampling theorem for signals with finite rate of innovation when the innovative position of the signal and the impulse response of the signal-generating device at the innovative position are given [8, 33]:

Theorem 5. *Let Λ, Γ be relatively separated subsets of $\mathbf{R}$, $\Phi = \{\phi_\lambda : \lambda \in \Lambda\}$ be a family of continuous functions on $\mathbf{R}$ such that $\sup_{t\in\mathbf{R},\lambda\in\Lambda} |\phi_\lambda(t)|(1+|t|)^\alpha < \infty$ for some $\alpha > 1$, and the space $V_2(\Phi)$ be as in (4.13). Assume that Φ is a Riesz basis of $V_2(\Phi)$, and that Γ is a stable ideal sampling set for $V_2(\Phi)$, i.e., there exist positive constants A and B such that*

$$A\|f\|_2 \le \|(f(\gamma))_{\gamma\in\Gamma}\|_2 \le B\|f\|_2 \quad \text{for all } f \in V_2(\Phi).$$

Then there exists a displayer $\tilde{\Psi} = \{\tilde{\psi}_\gamma : \gamma \in \Gamma\}$ such that

$$\sup_{t\in\mathbf{R},\gamma\in\Gamma} |\psi_\gamma(t)|(1+|t|)^\alpha < \infty$$

and

$$f(t) = \sum_{\gamma\in\Gamma} f(\gamma)\tilde{\psi}_\gamma(t-\gamma) \quad \text{for all } f \in V_2(\Phi).$$

Now we consider nonlinear and highly challenging problem of how to identify innovative positions and amplitudes of a signal with finite rate of innovation. For the stability for identification, the innovation positions should be separated from each other and the amplitudes at innovation positions should be above a certain level. So we may model those signals with finite rate of innovation as superposition of impulse response of active and nonactive generating devices located at unknown neighbors of a uniform grid. In [36] we assume, after appropriate scaling, that signals live in a perturbed shift-invariant space

$$V_{\infty,\oslash} := \left\{ \sum_{n\in\mathbf{Z}} c_n \varphi(\cdot - n - \sigma_n) \ : \ (c_k)_{k\in\mathbf{Z}} \in \ell^\infty_\oslash(\mathbf{Z}) \right\}$$

with unknown perturbation $\sigma := (\sigma_k)_{k\in\mathbf{Z}}$, where

$$\ell^\infty_\oslash(\mathbf{Z}) = \left\{ c := (c_k)_{k\in\mathbf{Z}} : \ \|c\|_{\ell^\infty_\oslash} := \sup_{c_k\neq 0} |c_k| + |c_k|^{-1} < \infty \right\}.$$

A negative result for sampling in the perturbed shift-invariant space $V_{\infty,\oslash}$ is that not all signals in such a space can be recovered from their samples if φ satisfies the popular Strang-Fix condition. The reason is that one cannot determine the jitter σ_0 of the signal $\sum_{k\in\mathbf{Z}} \varphi(\cdot - k - \sigma_0), \sigma_0 \in \mathbf{R}$, as it has constant amplitudes and is identical for all $\sigma_0 \in \mathbf{R}$. On the positive side, it is shown in [36] that any signal h in a perturbed shift-invariant space with unknown (but small) jitters can be recovered exactly from its average samples $\langle h, \psi_m(\cdot - k)\rangle, 1 \le m \le M, k \in \mathbf{Z}$, provided that the generator φ of the perturbed shift-invariant space and the average samplers $\psi_m, 1 \le m \le M$, satisfy the following condition:

$$\operatorname{rank} \begin{pmatrix} [\widehat{\nabla\varphi}, \widehat{\psi}_1](\xi) & \cdots & [\widehat{\nabla\varphi}, \widehat{\psi}_M](\xi) \\ [\hat{\varphi}, \widehat{\psi}_1](\xi) & \cdots & [\hat{\varphi}, \widehat{\psi}_M](\xi) \end{pmatrix} = 2 \quad \text{for all } \xi \in [-\pi, \pi]. \tag{4.14}$$

Here the *bracket product* $[f,g]$ of two square-integrable functions f and g is given by $[f,g](\xi)=\sum_{l\in\mathbf{Z}} f(\xi+2l\pi)\overline{g(\xi+2l\pi)}$.

Theorem 6. *Let φ and $\psi_1,\dots,\psi_M$ satisfy (4.14) and have the following regularity and decay properties:*

$$\sup_{t\in\mathbf{R}}\left(|\varphi(t)|+|\varphi'(t)|+|\varphi''(t)|+\sum_{m=1}^{M}|\psi_m(t)|\right)(1+|t|)^{\alpha}<\infty \tag{4.15}$$

for some $\alpha>1$. Then for any $L\ge 1$, there exists a positive number $\delta_1\in(0,1/2)$ such that any signal $h(t)=\sum_{k\in\mathbf{Z}} c_k\varphi(t-k-\sigma_k)$ in the space $V_{\infty,\varnothing}$ with $\|(c_k)_{k\in\mathbf{Z}}\|_{\ell^\infty_\varnothing}\le L$ and $\|(\sigma_k)_{k\in\mathbf{Z}}\|_\infty\le\delta_1$ could be reconstructed from its average sample data $\langle h,\psi_m(\cdot-k)\rangle, 1\le m\le M, k\in\mathbf{Z}$, in a stable way.

4.11 Sampling in Reproducing Kernel Banach Subspaces of L^p

Let $1\le p\le\infty$. A bounded linear operator T on $L^p(\mathbf{R})$ is said to be an *idempotent* operator if $T^2=T$. Denote the range space of the idempotent operator T on $L^p(\mathbf{R})$ by V_p; i.e.,

$$V_p:=\big\{Tf:\ f\in L^p(\mathbf{R})\big\}. \tag{4.16}$$

The Paley–Wiener space, finitely generated shift-invariant spaces, p-integrable spline spaces, spaces modeling signals with finite rate of innovation, and L^p itself are the range space of some idempotent operators.

Denote the *Wiener amalgam space* by

$$W^1:=\left\{f\in L^1(\mathbf{R}):\ \|f\|_{W^1}:=\left\|\sup_{-1/2\le z<1/2}|f(\cdot+z)|\right\|_1<\infty\right\}$$

and the *modulus of continuity* of a kernel function K on $\mathbf{R}\times\mathbf{R}$ by

$$\omega_\delta(K)(s,t):=\sup_{-\delta\le z_1,z_2\le\delta}|K(s+z_1,t+z_2)-K(x,y)|.$$

A *sampling set* Γ in this chapter means a relatively separated discrete subset of $\mathbf{R}$; i.e.,

$$B_\Gamma(\delta):=\sup_{t\in\mathbf{R}}\sum_{\gamma\in\Gamma}\chi_{[-\delta,\delta]}(t-\gamma)<\infty \tag{4.17}$$

for some $\delta>0$, where χ_E is the characteristic function on a set E. A sampling set Γ is said to have *gap* $\delta>0$ if

$$A_\Gamma(\delta):=\inf_{t\in\mathbf{R}}\sum_{\gamma\in\Gamma}\chi_{[-\delta,\delta]}(t-\gamma)\ge 1 \tag{4.18}$$

[1,3,4]. If we assume that the idempotent operator T is an integral operator

$$Tf(s) = \int_{\mathbf{R}} K(s,t)f(t)\mathrm{d}t, \quad f \in L^p(\mathbf{R}), \tag{4.19}$$

whose measurable kernel K has certain off-diagonal decay and regularity,

$$\left\| \sup_{z\in\mathbf{R}} |K(\cdot+z,z)| \right\|_{W^1} < \infty \tag{4.20}$$

and

$$\lim_{\delta\to 0} \left\| \sup_{z\in\mathbf{R}} |\omega_\delta(K)(\cdot+z,z)| \right\|_{W^1} = 0, \tag{4.21}$$

then,

- V_p is a reproducing kernel subspace of $L^p(\mathbf{R})$; i.e., for any $t \in \mathbf{R}$ there exists a positive constant C_t such that

$$|f(t)| \le C_t \|f\|_{L^p(\mathbf{R})} \quad \text{for all } f \in V_p. \tag{4.22}$$

- The kernel K satisfies the "reproducing kernel property":

$$\int_{\mathbf{R}} K(s,z)K(z,t)\mathrm{d}z = K(s,t) \quad \text{for all } s,t \in \mathbf{R}. \tag{4.23}$$

- $K(\cdot,t) \in V$ for any $t \in \mathbf{R}$.
- $V_p := \{\sum_{\lambda\in\Lambda} c(\lambda)\phi_\lambda(t-\lambda) : \ (c(\lambda))_{\lambda\in\Lambda} \in \ell^p(\Lambda)\}$, where Λ is a relative separated discrete subset of $\mathbf{R}$ and $\Phi = \{\phi_\lambda\}_{\lambda\in\Lambda} \subset V_p$ is localized in the sense that there exists a function h in the Wiener amalgam space W^1 such that ϕ_λ is dominated by h for every $\lambda \in \Lambda$, i.e.,

$$|\phi_\lambda(t)| \le h(t) \quad \text{for all } \lambda \in \Lambda \text{ and } t \in \mathbf{R}. \tag{4.24}$$

- Signals in V_p have finite rate of innovation.
- For $p = 2$, an idempotent operator T with kernel K satisfying symmetric condition $K(x,y) = \overline{K(y,x)}$ is a projection operator onto a closed subspace of L^2. In this case, the idempotent operator T and its kernel K is uniquely determined by its range space V_2 onto L^2.

The following sampling problem in the reproducing kernel space V_p is established in [25].

Theorem 7. *Let $1 \le p \le \infty$, T be an idempotent integral operator whose kernel K satisfies (4.20) and (4.21), V be the reproducing kernel subspace of $L^p(\mathbf{R})$ associated with the operator T, and $\delta_0 > 0$ be so chosen that*

$$r_0 := \left\| \sup_{z\in\mathbf{R}} |\omega_{\delta_0/2}(K)(\cdot+z,z)| \right\|_{L^1(\mathbf{R})} < 1. \tag{4.25}$$

Then any signal f in V can be reconstructed in a stable way from its samples $f(\gamma), \gamma \in \Gamma$, taken on a relatively separated subset Γ of $\mathbf{R}$ with gap δ_0.

Similar conclusion to the one in the above sampling theorem has been established in [12, Sect. 7.5] when the kernel K of the idempotent operator T satisfies the symmetric condition $K(x,y) = \overline{K(y,x)}$.

4.12 Convolution Sampling in Reproducing Kernel Banach Subspaces of L^p

In this section, we consider convolution sampling for signals in certain reproducing kernel subspaces of $L^p, 1 \le p \le \infty$. Here *convolution sampling* of a signal is ideal sampling of the convoluted signal taken on a sampling set. Precisely, given an integrable convolutor ψ and a sampling set Γ, the convolution sampling of a signal f includes two steps: Convoluting ψ with the signal f,

$$\psi * f(t) := \int_{-\infty}^{\infty} f(s)\psi(t-s)\mathrm{d}s,$$

and then sampling the convoluted signal $\psi * f$ at the sampling set Γ,

$$f \overset{\text{convoluting}}{\longmapsto} f * \psi \overset{\text{sampling}}{\longmapsto} \{(f * \psi)(\gamma)\}_{\gamma\in\Gamma}.$$

The data obtained by the above convolution sampling procedure is given by $\{(f * \psi)(\gamma)\}_{\gamma\in\Gamma}$. In [27], it is shown that any signal in the reproducing kernel subspace V_p associated with an idempotent operator can be stably reconstructed from its convolution samples taken on a sampling set with small gap if and only if the convolution procedure is stable on that space.

Theorem 8. *Let $1 \le p \le \infty, \psi_1, \ldots, \psi_L$ be integrable functions on the real line, V_p be the reproducing kernel subspace of L^p in (4.16), and set $\Psi = (\psi_1, \ldots, \psi_L)^T$. Assume that the kernel K of the idempotent operator T associated with the reproducing kernel space V_p satisfies (4.20) and (4.21). Then the following two statements are equivalent:*

(i) Ψ is a stable convolutor on V_p; i.e.,

$$0 < \inf_{g\in V_p, \|g\|_p=1} \sum_{l=1}^{L} \|\psi_l * g\|_p \le \sup_{g\in V_p, \|g\|_p=1} \sum_{l=1}^{L} \|\psi_l * g\|_p < \infty.$$

(ii) Ψ *is a stable convolution sampler on* V_p *for all sampling sets having sufficiently small gap; i.e., there exists* $\delta_0 > 0$ *such that*

$$0 < \inf_{0 \neq f \in V_p} \frac{\sum_{l=1}^{L} \left\| \left(\psi_l * f(\gamma)\right)_{\gamma \in \Gamma} \right\|_p}{\|f\|_p} \leq \sup_{0 \neq f \in V_p} \frac{\sum_{l=1}^{L} \left\| \left(\psi_l * f(\gamma)\right)_{\gamma \in \Gamma} \right\|_p}{\|f\|_p} < \infty$$

holds for any sampling set Γ *satisfying* $1 \leq A_\Gamma(\delta) \leq B_\Gamma(\delta) < \infty$ *for some* $\delta \in (0, \delta_0)$.

The equivalence in Theorem 8 was considered in [4] under the assumption that the reproducing kernel space V_p is a finitely-generated shift-invariant space.

4.13 Reproducing Kernel Hilbert Space Induced by Sampling Expansions

As indicated earlier, both the sampling map

$$f \longmapsto (f(t_n))_{n=1}^{\infty}$$

and the inverse map

$$(f(t_n))_{n=1}^{\infty} \longmapsto f$$

need to be continuous in a setting where sampling expansions are to be used. Thus the evaluation functional $E_t f := f(t)$ needs to be continuous for all t. Equivalently, the signal resides in a RKHS, even though this RKHS may not explicitly be identified. In [28], the authors have shown that, under very mild conditions, many versions of sampling theorems hold for RKHS. In [29], the authors asked whether an RKHS exists for each sampling theorem and showed that the answer is affirmative when a sampling sequence satisfies minimal properties. The starting point is an abstract notion of a sampling expansion.

Definition: Let f be a function belonging to a class $\mathscr{F}$ of continuous functions on $\Omega \subset \mathbf{R}$. A sampling theorem is associated with $\mathscr{F}$ if there is a sequence of sampling pairs $\{(S_n, t_n)\}$ of functions $S_n \in \mathscr{F}$ and points $t_n \in \Omega$ such that

- $S_n(t_k) = \delta_{nk}$, where δ_{nk} is the Kronecker delta.
- For each $f \in \mathscr{F}$, the sequence $\{f(t_n)\} \in \ell^2$, i.e., $\sum_n |f(t_n)|^2 < \infty$.
- The set $\{t_n\}$ is a set of uniqueness for $\mathscr{F}$.
- For each $\{b_n\} \in \ell^2$ the series $\sum_n b_n S_n(t)$ converges pointwise in Ω.

Then the authors construct a RKHS associated with sampling expansion as follows [29]:

Theorem 9. *Let H_0 be the Hilbert space consisting of $\mathscr{F}$ with the inner product $\langle f,g\rangle_0 = \sum_n f(t_n)\overline{g(t_n)}$. Then H_0 satisfies the following:*

- *H_0 is a reproducing kernel Hilbert space with $k(t,s) = \sum_n \overline{S_n(t)}S_n(s)$ as its reproducing kernel.*
- *$\{S_n\}$ is an orthogonal basis for H_0.*
- *The sampling expansion $g(t) = \sum_n g(t_n)S_n(t)$ holds for any $g \in H_0$.*

4.14 Sampling in Reproducing Kernel Banach Spaces

A *reproducing kernel Banach space* is a Banach space B of functions on a set Ω such that the evaluation functions $f \to f(t)$ is continuous for each $t \in \Omega$ [5]. The range space V_p of an idempotent integral operator is a reproducing kernel Banach space when the kernel of the idempotent operator satisfies certain regularity conditions. In this section, we investigate sampling in a reproducing kernel Banach space.

Let $1 \le p \le \infty$ and B be a Banach space with norm denoted by $\|\cdot\|_B$. We say that a countable subset Λ of Ω is a *p-sampling set* for the Banach space B if

$$0 < \inf_{f\in B,\|f\|_B=1} \|(f(\lambda))_{\lambda\in\Lambda}\|_p \le \sup_{f\in B,\|f\|_B=1} \|(f(\lambda))_{\lambda\in\Lambda}\|_p < \infty, \tag{4.26}$$

and a countable collection of elements $g_\lambda, \lambda \in \Lambda$, in the dual space of B to be a *p-frame* if

$$0 < \inf_{\|f\|_B=1} \|(g_\lambda(f))_{\lambda\in\Lambda}\|_p \le \sup_{\|f\|_B=1} \|(g_\lambda(f))_{\lambda\in\Lambda}\|_p < \infty, \tag{4.27}$$

i.e., the analysis operator $T : B \ni f \longmapsto (g_\lambda(f))_{\lambda\in\Lambda} \in \ell^p$ is bounded from both above and below [2]. Similarly to sampling in a RKHS, for a reproducing kernel Banach space B of functions on a set Ω, a countable subset Λ of Ω is a p-sampling set for the space B if and only if the corresponding evaluation functionals $h_\lambda, \lambda \in \Lambda$, form a p-frame for the space B. Moreover, in [14] it is shown that a reconstruction formula always exists.

Theorem 10. *Let $1 \le p,q \le \infty$ satisfy $1/p + 1/q = 1$, B be a reproducing kernel Banach space of functions on a set Ω, and $\Lambda \subset \Omega$ be a p-sampling set. Then there exists a collection of functions $S_\lambda(t), \lambda \in \Lambda$, such that:*

- *$(S_\lambda(t))_{\lambda\in\Lambda}$ is q-summable for every $t \in \Omega$.*
- *$(\eta_\lambda)_{\lambda\in\Lambda}$ is a p-frame for the range space of the sampling operator $S : B \ni f \longmapsto (f(\lambda))_{\lambda\in\Lambda} \in \ell^p$, where $\eta_\lambda = (S_{\lambda'}(\lambda))_{\lambda'\in\Lambda}$.*

- *Every signal f in the reproducing kernel Banach space B has the following sampling expansion:*

$$f(t) = \sum_{\lambda \in \Lambda} f(\lambda) S_\lambda(t),\ t \in \Omega,$$

with the pointwise convergence.

4.15 Average Sampling in L^2

In this section, we consider very general sampling procedure where the samples are obtained by inner products between time signal and sampling functionals. More precisely, given a time signal f living in a Hilbert space H, its average sample y_γ at the location $\gamma \in \Gamma$ is obtained by taking the inner product between the signal f and the sampling functional $\psi_\gamma(\cdot - \gamma)$ at the location γ; i.e., the sampling procedure on H via the average sampler $\Psi = (\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$ is a linear operator from H to $\ell^2(\Gamma)$:

$$S : H \ni f \longmapsto \{y_\gamma := \langle f, \psi_\gamma(\cdot - \gamma)\rangle\}_{\gamma \in \Gamma} \in \ell^2(\Gamma). \tag{4.28}$$

We restrict ourselves to consider *well-localized samplers* $\Psi = (\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$, which means that Γ is a relatively separated subset of $\mathbf{R}$ and the sampling functionals ψ_γ are dominated by a function h in the Wiener amalgam space W^1; i.e., $|\psi_\gamma(t)| \le h(t)$ for all $t \in \mathbf{R}$ and $\gamma \in \Gamma$. The reasons for considering well-localized samplers are twofold:

- At each position $\gamma \in \Gamma$, we locate an acquisition device, and hence it is reasonable to assume that there are finitely many such acquisition devices in any unit intervals, which in turn implies that Γ is relatively separated.
- We use the sampling functional ψ_γ to reflect the characteristic of the acquisition device at the location γ, and hence the sampling functional ψ_γ should essentially be supported in a neighborhood of the sampling location γ, which can be described by the dominance by a function h with fast decay at infinity.

It is well known that signals with finite energy do not have finite rate of innovation. In [26], we show that any signal f with finite energy could be determined by its samples $\langle f, \psi_\gamma(\cdot - \gamma)\rangle, \gamma \in \Gamma$ for some well-localized sampler $(\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$, but could not be recovered in a stable way from the samples $\langle f, \psi_\gamma(\cdot - \gamma)\rangle, \gamma \in \Gamma$ for any well-localized sampler $(\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$.

Theorem 11. *(i) There is a well-localized sampler $(\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$ such that any function $f \in L^2$ is uniquely determined by its samples $\langle f, \psi_\gamma(\cdot - \gamma)\rangle, \gamma \in \Gamma$.*

(ii) There does not exist a well-localized sampler $(\psi_\gamma(\cdot - \gamma))_{\gamma \in \Gamma}$ such that the sampling operator S in (4.28) is stable for $H = L^2$ in the sense that there exist positive constants A and B such that $A\|f\|_2^2 \le \sum_{\gamma \in \Gamma} |\langle f, \psi_\gamma(\cdot - \gamma)\rangle|^2 \le B\|f\|_2^2$ for all $f \in L^2$.

We remark that functions $\psi_\gamma, \gamma \in \Gamma$, in the well-localized sampler in the first conclusion of Theorem 11 cannot be selected to be supported in a fixed compact set, but it is possible to let elements $\psi_\gamma, \gamma \in \Gamma$, in the well-localized sampler to be independent on $\gamma \in \Gamma$. This is closely related to the spectral problem: the density of the collection of exponentials $\{\exp(i\gamma t)\}_{\gamma\in\Gamma}$ in a weighted L^2 space [31]. In [26], we conjecture that there is not a determining sampler $\{\psi_\gamma(\cdot-\gamma)|\ \gamma\in\Gamma\}$ such that $\|\psi_\gamma\|_2 = 1$ and $|\psi_\gamma(x)| \le C\exp(-\varepsilon|x|)$ for some positive constants C, ε and a relatively separated subset Γ of $\mathbf{R}$.

Dedication. This chapter is dedicated to Professor Gilbert Walter on the occasion of his 80th birthday:

- In appreciation of his friendship and important contributions to Mathematical Analysis and Applications.
- With admiration of the novel and clever ways in which he has brought together ideas from classical and modern analysis to advance our understanding of generalized functions, wavelets, and signal processing.

Acknowledgments The second author is partially supported by the National Science Foundation (DMS-1109063).

References

1. Aldroubi A, Gröchenig K (2001) Nonuniform sampling and reconstruction in shift-invariant space, SIAM Rev 43:585–620
2. Aldroubi A, Sun Q, Tang W-S (2001) p-frames and shift-invariant subspaces of L^p. J Fourier Anal Appl 7:1–21
3. Aldroubi A, Sun Q, Tang W-S (2004) Nonuniform average sampling and reconstruction in multiply generated shift-invariant spaces. Constr Approx 20:173–189
4. Aldroubi A, Sun Q, Tang W-S (2005) Convolution, average sampling, and a Calderon resolution of the identity for shift-invariant spaces. J Fourier Anal Appl 22:215–244
5. Aronszajn N (1950) Theory of reproducing kernels. Trans Am Math Soc 68:337–404
6. Benedetto JJ, Ferreira PJSG (eds.) (2001) Modern sampling theory: mathematics and applications. Birkhuse, Boston
7. Benedetto JJ, Zayed AI (eds.) (2003) Sampling, wavelets, and tomography. Birkhauser, Boston
8. Bi N, Nashed MZ, Sun Q (2009) Reconstructing signals with finite rate of innovation from noisy samples. Acta Appl Math 107:339–372.
9. Butzer PL (1983) A survey of the Whittaker-Shannon sampling theorem and some of its extensions. J Math Res Exposition 3:185–212
10. Butzer PL, Splettstöβer W, Stens RL (1988) The sampling theorem and linear prediction in signal analysis. Jahresber Deutsch Math-Verein 90:1–60
11. Butzer PL, Stens RL (1992) Sampling theory for not necessarily band-limited functions: a historical overview. SIAM Rev 34:40–53
12. Fornasier M, Rauhut H (2005) Continuous frames, function spaces, and the discretization problem. J Fourier Anal Appl 11:245–287
13. Haddad RA, Parsons TW (1991) Digital signal processing: theory, applications and hardware. Computer Science Press
14. Han D, Nashed MZ, Sun Q (2009) Sampling expansions in reproducing kernel Hilbert and Banach spaces. Numer Funct Anal Optim 30:971–987

15. Higgins JR (1985) Five short stories about the cardinal series. Bull Am Math Soc 12:45–89
16. Higgins JR (1996) Sampling theory in fourier and signal analysis volume 1: foundations. Oxford University Press, Oxford
17. Higgins JR, Stens RL (2000) Sampling theory in fourier and signal analysis: volume 2: advanced topics. Oxford Science Publications, Oxford
18. Ismail M, Nashed Z, Zayed A, Ghaleb A (eds.) (1995) Mathematical analysis, wavelets and signal processing (contemporary mathematics), vol 190. American Mathematical Society, Providence, RI
19. Jerri JA (1977) The Shannon sampling theorem—its various extensions and applications: a tutorial review. Proc IEEE 65:1565–1596
20. Larson D, Massopust P, Nashed Z, Nguyen MC, Papadakis M, Zayed A (eds.) (2008) Frames and operator theory in analysis and signal processing (Contemporary Mathematics), vol 451. American Mathematical Society, Providence, RI
21. Marks II RJ (1991) Introduction to shannon sampling and interpolation theory. Springer, Berlin
22. Marks RJ (1993) Advanced topics in shannon sampling and interpolation theory. Springer, Berlin, Heidelberg
23. Marvasti FA (ed) (2001) Nonuniform Sampling: theory and practice (information technology: transmission, processing, and storage). Plenum Pub Corp, New York
24. Nashed MZ, Scherzer O (eds.) (2002) Inverse problems, image analysis and medical imaging (Contemporary Mathematics), vol 313. American Mathematical Society, Providence, RI
25. Nashed MZ, Sun Q (2010) Sampling and reconstruction of signals in a reproducing kernel subspace of $L^p(R^d)$. J Funct Anal 258:2422–2452
26. Nashed MZ, Sun Q, Tang W-S (2009) Average sampling in L^2. Can Acad Sci Paris, Ser I 347:1007–1010
27. Nashed MZ, Sun Q, Xian J Convolution sampling and reconstruction of signals in a reproducing kernel subspace. Proc Am Math Soc (to appear)
28. Nashed MZ, Walter GG (1991) General sampling theorems for functions in reproducing kernel Hilbert spaces. Math Contr Signals Syst 4:363–390
29. Nashed MZ, Walter GG (1995) Reproducing kernel Hilbert spaces from sampling expansions. Contemporary Math 190:221–226
30. Oppenheim AV, Schafer RW (1989) Digital signal processing. Prentice-Hall Inc., New Jersey
31. Olevskii A, Ulanovskii A (2004) Almost integer translates. Do nice generators exist? J Fourier Anal Appl 10:93–104
32. Papoulis A (1977) Generalized sampling expansion. IEEE Trans Circ Syst 24:652–654
33. Sun Q (2006) Non-uniform sampling and reconstruction for signals with finite rate of innovations. SIAM J Math Anal 38:1389–1422
34. Sun Q (2008) Frames in spaces with finite rate of innovation. Adv Comput Math 28:301–329
35. Sun Q (2010) Local reconstruction for sampling in shift-invariant spaces. Adv Comput Math 32:335–352
36. Sun Q (2011) Localized nonlinear functional equations and two sampling problems in signal processing. submitted
37. Unser M (2000) Sampling—50 years after Shannon. Proc IEEE 88:569–587
38. van der Mee C, Nashed MZ, Seatzu S (2003) Sampling expansions and interpolation in unitarily translation invariant reproducing kernel Hilbert spaces. Adv Comput Math 19:355–372
39. Vetterli M, Marziliano P, Blu T (2002) Sampling signals with finite rate of innovation. IEEE Trans Signal Process 50:1417–1428
40. Walter GG (1992) A sampling theorem for wavelet subspaces. IEEE Trans Inform Theor 38:881–884
41. Walter GG, Shen X (2000) Wavelets and other orthogonal systems, 2nd edn. CRC Press, USA
42. Yao K (1967) Applications of reproducing kernel Hilbert spaces-bandlimited signal models. Inform Contr 11:429–444
43. Zayed A (1993) Advances in Shannon's sampling theory. CRC Press, USA

Chapter 5
Coprime Sampling and Arrays in One and Multiple Dimensions

P.P. Vaidyanathan and Piya Pal

Abstract This chapter gives an overview of the concept of coprime sampling. The basic idea is that a continuous-time (or spatial) signal is sampled simultaneously by two sets of samplers, with sampling rates $1/MT$ and $1/NT$ where M and N are coprime integers and $T > 0$. One of the results is that it is possible to estimate the autocorrelation of the signal at a much higher rate ($= 1/T$) than the total sampling rate. Thus, any application which is based on autocorrelation will benefit from such sampling and reconstruction. One example is in array processing, in the context of estimation of direction of arrival (DOA) of sources. Traditionally, an array with L sensors would be able to identify $L-1$ independent sources, but with a pair of coprime arrays, one can identify $O(L^2)$ sources. It is also shown how to use two DFT filter banks, one in conjunction with each sampling array, to produce a much denser tiling of the frequency domain than each filter bank would individually be able to do. This chapter also discusses the extension of coprime sampling to multiple dimensions by using sampling geometries that are defined based on lattices. In this context the generation of coprime pairs of integer matrices is a very interesting mathematical problem and is dealt with in detail. The use of coprime samplers in system identification is also elaborated upon. A brief review of fractionally spaced equalizers in digital communications, in the context of coprime sampling, is included.

5.1 Introduction

Coprime sampling was recently introduced in [18] for the case of one-dimensional signals and extended to multiple dimensions in [19]. The basic idea is that a continuous-time (or spatial) signal is sampled simultaneously by two sets of

P.P. Vaidyanathan (✉) • P. Pal
California Institute of Technology, Pasadena, CA 91125, USA
e-mail: ppvnath@systems.caltech.edu; piyapal@caltech.edu

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_5,

samplers at rates $1/NT$ and $1/MT$, where M and N are coprime integers. One of the results is that it is possible to estimate the autocorrelation of the signal at the much higher rate $1/T$. Thus, any application which is based on autocorrelation will benefit from such sampling and reconstruction.

One example is in array processing, in the context of estimation of directions of arrival (DOA) of sources. Traditionally, an array with L sensors would be able to identify $L-1$ independent sources, but with a pair of coprime arrays, one can identify $O(L^2)$ sources. It is also possible to use two DFT (Discrete Fourier Transform) filter banks, one in conjunction with each sampling array, to produce a much denser tiling of the frequency domain than each filter bank would individually be able to do. Coprime sampling can also be used to estimate the frequencies of sinusoids buried in noise. The highest frequency that can be estimated is proportional to $1/T$ even though the number of samples per unit time is only $(1/NT)+(1/MT)$.

Coprime sampling can be extended to multiple dimensions by using a pair of samplers on lattice geometries. An important mathematical problem that comes up in this context is the generation of pairs of integer matrices $\mathbf{M}$ and $\mathbf{N}$ which are commuting and coprime. Coprime sampling also has interesting applications in system identification and in channel equalization.

5.1.1 Outline

In this chapter we review coprime sampling and some of its applications. In Sect. 5.2 we introduce coprime samplers and spatial arrays for the case of one-dimensional signals. The application in the estimation of direction of arrival of uncorrelated sources is described in Sect. 5.3, and Sect. 5.4 explains the application of DFT filter banks in conjunction with coprime arrays. For multidimensional signals, coprime sampling theory based on lattice geometries is introduced in Sect. 5.5, and many properties of multidimensional arrays are reviewed in Sect. 5.5.1. The use of two-dimensional DFT filter banks in conjunction with coprime arrays is described in Sect. 5.5.2. An important mathematical problem that comes up in this context is the generation of pairs of integer matrices which are commuting and coprime. This is discussed at length in Sect. 5.6. The application of coprime sampling in the identification of linear systems from input-output measurements is briefly described in Sect. 5.7, and applications in fractionally spaced equalizers are elaborated in Sect. 5.8.

Historically, coprime sampling has in the past been used for identifying sinusoids in noise (see references in [27, 28]). Coprime pulsing has also been employed for the resolution of range ambiguities in radar [16]. Coprime arrays, their coarrays, and applications, as discussed in this chapter, were introduced in [18], and the two-dimensional extensions were given in [19].

5.1.2 Notations

Here are some of the standard acronyms used throughout the chapter: (a) FIR (finite duration impulse response), (b) IIR (infinite duration impulse response), (c) DFT and IDFT (discrete fourier transform and its inverse), (d) LTI (linear time invariant), and (e)MIMO (multiple input multiple output). Boldfaced letters indicate vectors and matrices. $\mathbf{A}^T, \mathbf{A}^*$, and $\mathbf{A}^\dagger$, respectively, denote the transpose, conjugate, and transpose-conjugate of a matrix. The z-transform of a sequence $x(n)$ is denoted as $X(z)$, that is, $X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n}$. The greatest common divisor (gcd) of a set of integers $a, b, c, \ldots$ is denoted as

$$(a, b, c, \ldots). \tag{5.1}$$

The quantity $W_N = e^{-j2\pi/N}$, where $j = \sqrt{-1}$, arises in the expressions involving the DFT, and the subscript is dropped if there is no ambiguity. The term polynomial matrix refers to an expression of the form

$$\mathbf{H}(z) = \sum_{n=0}^{J} \mathbf{h}(n)z^{-n}, \tag{5.2}$$

where $\mathbf{h}(n)$ are matrices. Thus a polynomial matrix is nothing but the transfer function of a causal MIMO LTI system [23].

5.2 Coprime Arrays in One Dimension

Consider Fig. 5.1 which shows two uniform samplers operating in parallel on a continuous-time signal. The rates of the samplers are $1/MT$ and $1/NT$, where $M > 1$ and $N > 1$ are integers, and the sampled sequences are

$$x(Mn_1) = x_c(MTn_1), \quad x(Nn_2) = x_c(NTn_2), \tag{5.3}$$

where $x_c(t)$ is the underlying continuous-time signal being sampled. Imagine that $x_c(t)$ is wide-sense stationary (WSS) with an autocorrelation

$$R_c(\tau) = E[x_c(t)x_c^*(t-\tau)] \tag{5.4}$$

and that we are interested in obtaining the samples

$$R(k) = R_c(kT) \tag{5.5}$$

with sample spacing T. This sample rate $1/T$ is higher than the sampling rates of the individual samplers. If M and N are coprime integers (i.e., have no common factors),

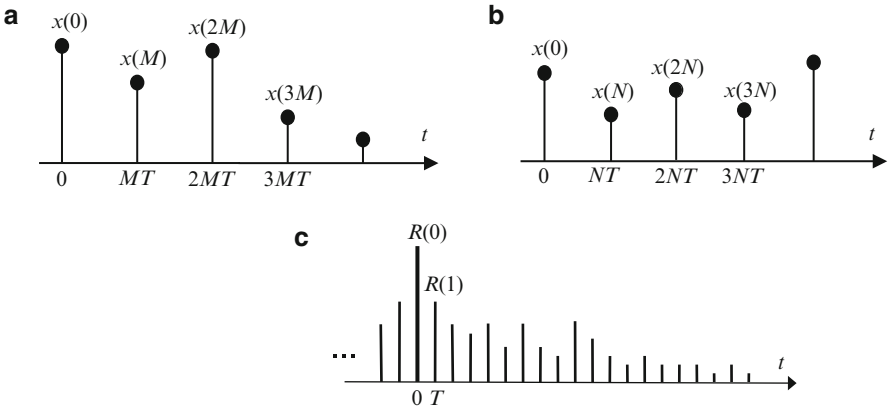

Fig. 5.1 **(a)**, **(b)** Two uniform samplers operating in parallel on a signal $x_c(t)$ The rates are $1/MT$ and $1/NT$ where M and N are integers **(c)** The autocorrelation of $x_c(t)$ sampled at the dense spacing T

then it can be shown that the densely sampled sequence $R(k)$ can be estimated from the sparse set of samples $x(Mn_1)$ and $x(Nn_2)$ by time-domain averaging, no matter how large M and N are. To see this observe that

$$R(k) = E[x(Mn_1)x^*(Nn_2)] = E[x(n)x^*(n-k)], \tag{5.6}$$

where

$$k = Mn_1 - Nn_2. \tag{5.7}$$

If M and N are coprime, then given any integer k, there exist integers n_1, n_2 such that (5.7) is satisfied. Thus, by performing a time-domain averaging such as

$$R(k) \approx \frac{1}{J}\sum_{n=0}^{J-1} x(M(n_1+nN)x^*(N(n_2+nM)), \tag{5.8}$$

we can estimate $R(k)$ (assuming that the process $x(n)$ is ergodic). Considering that there are two samplers, the total rate of sampling is

$$\frac{1}{NT} + \frac{1}{MT} \quad \text{samples per unit time.} \tag{5.9}$$

The autocorrelation can therefore be estimated at a much higher sampling rate $1/T$ than the rate at which the signals are sampled. Notice that the total duration of time for which samples have to be collected in order to perform the averaging in (5.8) is approximately equal to MNJ. This affects the latency or delay involved in

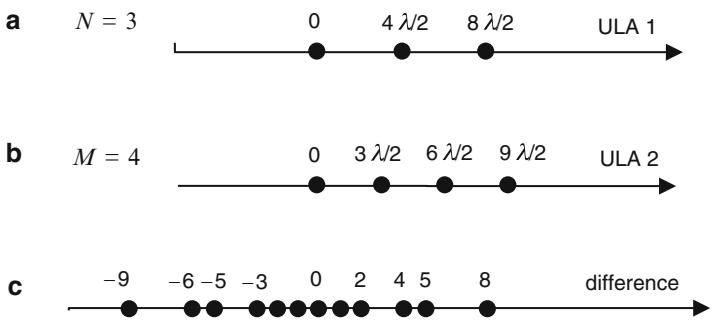

Fig. 5.2 **(a)**, **(b)** A pair of sparse coprime arrays in space (c) The difference set $k = Mn_1 - Nn_2$

computing the autocorrelations. The price paid for sparsity (large M,N) is precisely this latency. The choice of the rate $1/T$ depends on how finely we want the autocorrelation sampled (e.g., Nyquist rate of $R_c(\tau)$).

5.2.1 Spatial Arrays

If the sampling is in space, then this leads to applications in array processing, such as the estimation of the direction of arrival (DOA) for electromagnetic waves [24]. In such applications, the variable T is typically $\lambda/2$ where λ is the wavelength of interest. Figure 5.2(a), (b) shows an example of spatial arrays which are sparse and coprime. The number of array elements is finite in this example: there are N elements with spacing $M\lambda/2$ and M elements with spacing $N\lambda/2$. Figure 5.2(c) shows the set of all differences $Mn_1 - Nn_2$ in the example. The actual locations of the coarray elements will be at the points

$$(Mn_1 - Nn_2)\lambda/2. \tag{5.10}$$

In the case of space arrays, the averaging (5.8) is replaced with **snapshot averaging**. For example, if $x_l(n\tau)$ and $x_m(n\tau)$ are the samples of the outputs of the lth and mth sensors at time $n\tau$, then

$$R(k) \approx \frac{1}{J}\sum_{n=0}^{J-1} x_l(n\tau)x_m^*(n\tau), \tag{5.11}$$

where $k = l - m$. Here J is the number of snapshots, and τ is the time-domain sample spacing at the outputs of the spatial sensors. Thus, by snapshot averaging, the autocorrelation $R(k)$ can be computed for all values of k in (5.7) which can be achieved for the specified ranges of n_1 and n_2. It is assumed here that the signal has spatial wide-sense stationarity, that is, the autocorrelation depends only on the

difference $l-m$. The fact that the two arrays are sparse also means that there is less **mutual coupling** between adjacent sensor elements in space. This is an important advantage in practical implementations of the system.

5.2.2 Coarrays

The set of all achievable values of the difference (5.7) will be referred to as the difference coarray generated by the sampling arrays $\{Mn_1\}$ and $\{Nn_2\}$. Thus, *the difference coarray is the set of lags k at which the autocorrelation* $R(k) = R_c(kT)$ *can be estimated*. More precisely, we have the following definition.

Definition 1. Given any set $\mathscr{N}$ of numbers (e.g., sensor positions), the set of all distinct differences

$$n_l - n_m, \tag{5.12}$$

where $n_l, n_m \in \mathscr{N}$, is said to be the **difference coarray** of the set $\mathscr{N}$.

Thus, if k belongs in the coarray, then so does $-k$, and the coarray is automatically symmetric. For example, if $\mathscr{N}$ is the union of all the sensor positions in the two coprime arrays, then the coarray would have all *cross differences* of the form

$$Mn_1 - Nn_2 \quad \text{and} \quad Nn_2 - Mn_1 \tag{5.13}$$

and self-differences of the form

$$Mn_1 - Mn_1' \quad \text{and} \quad Nn_2 - Nn_2'. \tag{5.14}$$

A number of remarks are now in order:

1. If the sets of integers n_1 and n_2 defining the sensor positions Mn_1 and Nn_2 are unrestricted (i.e., belong to the range $-\infty < n_1, n_2 < \infty$), then it is clear that all the self differences are already included in the cross differences. It turns out that even if the sets of integers n_1 and n_2 defining the sensor positions Mn_1 and Nn_2 are restricted to certain finite sets, it is often true that all the self differences are included in the set of cross differences. For example, if $0 \le n_1 \le N-1$ and $0 \le n_2 \le M-1$, then this can be verified to be the case [18]. For convenience we shall therefore never explicitly refer to the self-differences, and take it for granted that the complete coarray is just the union of the sets $Mn_1 - Nn_2$ and $Nn_2 - Mn_1$.
2. In the context of coprime arrays we sometimes distinguish between $Mn_1 - Nn_2$ and $Nn_2 - Mn_1$ and refer to their union as a **symmetric coarray** for added clarity.[1]

[1]This is a slight misuse of notation because the coarray of a set $\mathscr{N}$, by definition, is a symmetric set [see (5.12)].

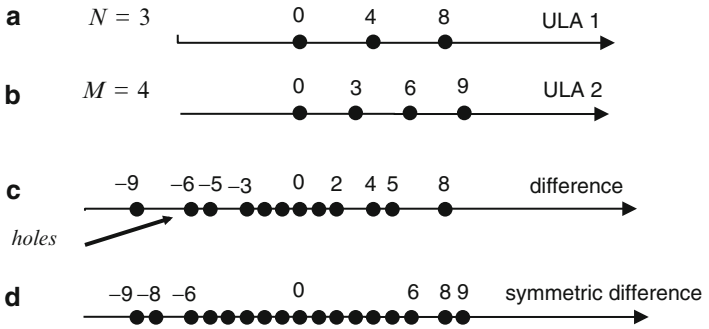

Fig. 5.3 **(a)**, **(b)** Example of coprime ULAs with $N = 3, M = 4$, (c) Difference set $k = Mn_1 - Nn_2$, and (d) Symmetric difference All distances are in multiples of $\lambda/2$

3. Note that the elements in the set $\mathcal{N}$ can also be two- or three-dimensional vectors, as in the case of higher-dimensional arrays [7].
4. Historically, the use of coarrays for the computation of autocorrelations, and hence for applications in DOA estimation has been considered earlier in the context of minimum redundancy arrays (MRAs) [1, 3, 8, 13], but coarrays of *coprime* arrays have only been considered recently [18].

5.2.3 Arrays with Extra Elements

In the equation

$$k = Mn_1 - Nn_2, \tag{5.15}$$

although k takes on all integer values if n_1 and n_2 are allowed to take all integer values, in practice the ranges of n_1 and n_2 are limited. For example, if n_1 and n_2 are in the ranges $0 \le n_1 \le N-1, 0 \le n_2 \le M-1$, then k takes MN distinct values in the range

$$-N(M-1) \le k \le M(N-1). \tag{5.16}$$

Since this range has room for nearly $2MN$ integers, the coarray has *holes*. Figure 5.3(a), (b) shows an example of two spatial uniform linear arrays (ULAs) with $N = 3$ and $M = 4$. (All distances are in multiples of $\lambda/2$.) The difference set $Mn_1 - Nn_2$ contains integers from -9 to 8, with holes at $-8, -7, -4, 3, 6$, and 7 (Fig. 5.3c). If we also include the negative of this set, $Nn_2 - Mn_1$, then we get the symmetric difference set shown in Fig. 5.3(d). This has a uniform stretch of all integers from -6 to 6, and the holes are at ± 7. This is precisely the coarray of the set of all sensors in the union of the two arrays in this example:

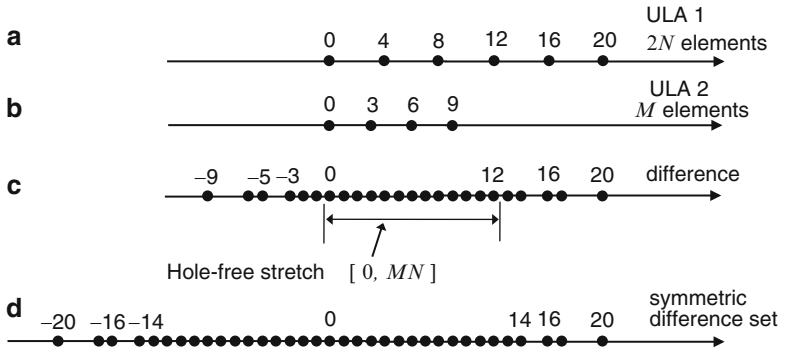

Fig. 5.4 **(a)**, **(b)** Example of coprime ULAs with $2N$ and M elements where $N = 3, M = 4$, **(c)** Difference set $k = Mn_1 - Nn_2$, and (d) Symmetric difference. All distances are in multiples of $\lambda/2$

It is shown in [18] that if n_1 and n_2 are in the extended range

$$0 \leq n_1 \leq 2N-1, \quad 0 \leq n_2 \leq M-1, \tag{5.17}$$

then k takes all the MN contiguous values in the range $0 \leq k \leq MN-1$ (and some values outside). The element $k = MN$ is also automatically included (just set $n_1 = N, n_2 = 0$). In particular, there are *no holes* in the range $0 \leq k \leq MN$. Figure 5.4 shows what happens to the example with $N = 3$ and $M = 4$ when $2N$ sensors are used instead of N. As seen from part (c), there is a hole-free range $0 \leq k \leq MN$; in fact, it is slightly longer in this example ($-3 \leq k \leq 14$). The symmetric difference set now stretches from -14 to 14 (Fig. 5.4d). Alternatively, if n_1 is in the range $0 \leq n_1 \leq N-1$ and n_2 in the extended range $-(M-1) \leq n_2 \leq M-1$, then also k takes all the MN contiguous values in the range $0 \leq k \leq MN-1$.

It should be mentioned here that there is another array geometry called the nested array geometry, which can create coarrays with $O(N^2)$ elements starting from $O(N)$ sensors. The nested array was developed in [11] and has the advantage that it can produce a coarray without holes and has larger number of distinct coarray elements compared to the coprime arrays. The coprime arrays on the other hand offer greater spacing between sensors which helps to reduce mutual coupling between them.

5.3 Application in DOA Estimation

Consider Fig. 5.5(a) which schematically shows a monochromatic plane wave arriving at an angle of θ with respect to the vertical. Figure 5.5(b) shows a pair of coprime ULAs, with elements spaced apart by $M\lambda/2$ and $N\lambda/2$. There are $2N$ elements in the first array and M elements in the second [as in (5.17)]. Figure 5.5(c) shows the symmetric coarray with a ULA stretch in the range $-A \leq n \leq A$ where

$$A = MN, \tag{5.18}$$

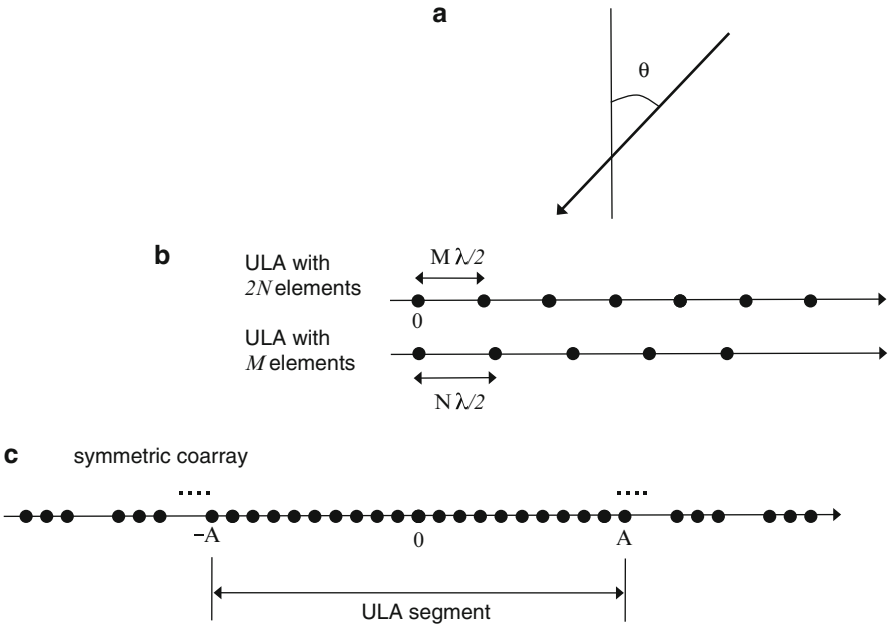

Fig. 5.5 **(a)** Schematic of a plane wave arriving at angle θ, **(b)** A pair of coprime ULAs, and **(c)** Symmetric coarray or virtual array with aperture A

with virtual elements spaced apart by $\lambda/2$. We say that the coarray has **aperture** A, that is, there are some missing sensors for $|n| > A$, but the region $-A \leq n \leq A$ is filled with virtual sensors.

Now, there are standard algorithms for the estimation of the direction of arrival (DOA) based on the measured autocorrelation set $R(k)$. The MUSIC algorithm [24], for example, can estimate the DOAs of J independent sources from a knowledge of $R(k), -J \leq k \leq J$. Typically, the so-called MUSIC spectrum is computed and plotted and shows peaks at angles where there are sources. If we have a single ULA with N elements, then since the coarray is an ULA in the range $-(N-1) \leq k \leq N-1$, the MUSIC algorithm can identify up to $N-1$ sources. If the array is a union of two coprime arrays as in our case (with $2N$ and M elements, respectively, as in [5.17)], then $R(k)$ can be estimated for

$$-MN \leq k \leq MN. \tag{5.19}$$

Based on this one would expect that we can identify up to MN independent sources. This is indeed the case, but the details of the development of the MUSIC algorithm for estimation based on coprime coarrays (or coarrays of any other array geometry) are rather nontrivial and have been developed in [10] (based on an earlier work by the same authors in the context of nested arrays [11]).

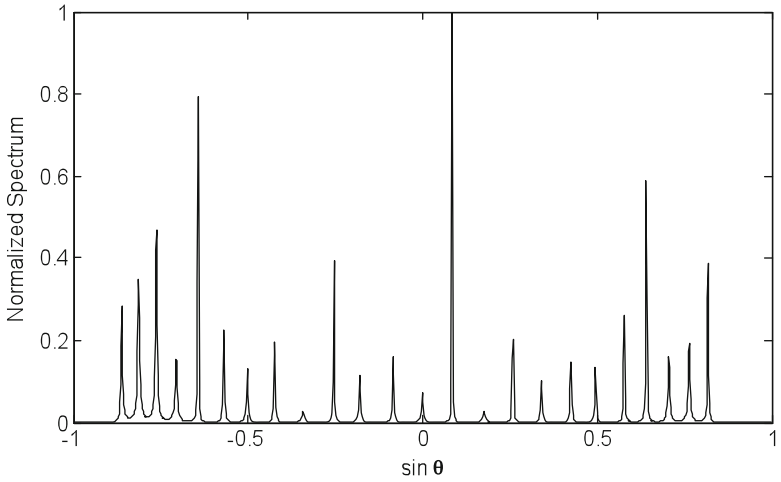

Fig. 5.6 MUSIC spectrum based on the coprime coarray generated from $M = 7, N = 5$. Up to 35 independent sources can be identified. The figure shows an example with 24 independent narrowband sources. The *SNR* is 0 dB

Consider an example where $M = 7$ and $N = 5$. The number of physical sensor elements is $M + 2N = 17$. We have $A = MN = 35$ which shows that up to 35 independent narrowband sources can be identified. For the case where there are 24 independent sources, Fig. 5.6 shows the MUSIC spectrum computed as in [10]. The signal to noise ratio at the sensor outputs is assumed to be 0 dB in this example. The figure clearly shows all the 24 peaks corresponding to the 24 sources.

Coprime sampling can also be used to estimate the frequencies of sinusoids buried in noise. The idea is similar in principle to DOA estimation. The highest frequency that can be estimated is proportional to $1/T$ even though the number of samples per unit time is only $(1/NT) + (1/MT)$. For further details see [18].

5.4 Coprime DFT Filter Banks and Beamforming

We now consider the application of DFT filter banks [23] at the outputs of the coprime arrays (Fig. 5.7). In this system we combine the outputs of the sensors using IDFT matrices. This creates a set of M filters $H_k(z)$ in the first array and a set of N filters $G_\ell(z)$ in the second array. We will see that each filter $H_k(z)$ has multiple passbands (multiple beams or grating lobes in the array-processing context), because of the large interelement spacing $N\lambda/2$. The same remark applies to $G_\ell(z)$. Assuming M and N are coprime, it was shown in [18] that by taking the MN products of these sets of M and N DFT outputs and performing statistical averaging, it is possible to simulate the behavior of an MN-point DFT filter bank in the power-spectrum domain! Thus, the power-spectral frequency domain can

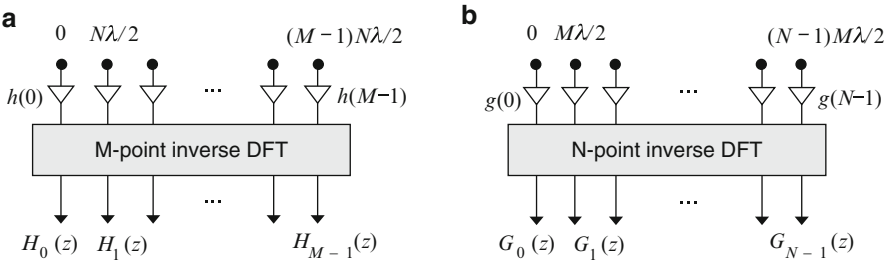

Fig. 5.7 A pair of *DFT* (Discrete Fourier Transform) filter banks generating M beams from array one and N beams from array two

be densely tiled with MN narrowband filters, each with a single passband (even though the two individual filter banks have only M and N filters, respectively, each producing multiple passbands). This result was derived in [18] (and briefly reviewed in [19]). In the context of array processing, since the DFT domain can be mapped to the angle of arrival domain [24], this result has direct bearing on two-dimensional beamforming.

To describe the details, consider Fig. 5.7(a), where the 0th filter is obtained by adding the outputs of the taps $h(n)$. The associated beam pattern is [24]

$$\sum_{n=0}^{M-1} h(n)e^{-j\alpha n}, \tag{5.20}$$

where $\alpha = 2\pi d \sin\theta/\lambda$, d is the interelement spacing and λ is the wavelength of the monochromatic wave impinging on the array from the direction of arrival θ. Traditionally, the interelement spacing is $d = \lambda/2$ (unlike in Fig. 5.7). Let the corresponding α be denoted as ω:

$$\omega = \pi \sin\theta. \tag{5.21}$$

For the sparse array since the interelement spacing $d = N\lambda/2$, the beampattern $H_0(z)$ is

$$H(e^{j\omega N}) = \sum_{n=0}^{M-1} h(n)e^{-jN\omega n}. \tag{5.22}$$

We can regard this as a "transfer function" $H(z^N)$. Similarly, the 0th filter $G_0(z)$ from the second array produces $G(z^M)$. The outputs of $H_0(z)$ and $G_0(z)$, in response to two sources at ω_1 and ω_2, are

$$y_1 = c_1 H(e^{j\omega_1 N}) + c_2 H(e^{j\omega_2 N}) \tag{5.23}$$

and

$$y_2 = c_1 G(e^{j\omega_1 M}) + c_2 G(e^{j\omega_2 M}). \tag{5.24}$$

Now

$$\begin{aligned} y_1 y_2^* = {} & |c_1|^2 H(e^{j\omega_1 N}) G^*(e^{j\omega_1 M}) + |c_2|^2 H(e^{j\omega_2 N}) G^*(e^{j\omega_2 M}) \\ & + c_1 c_2^* H(e^{j\omega_1 N}) G^*(e^{j\omega_2 M}) + c_1^* c_2 H(e^{j\omega_2 N}) G^*(e^{j\omega_1 M}). \end{aligned} \tag{5.25}$$

The "cascade effect" is represented by the first two terms, but there are cross terms (last two terms). If $c_1(n)$ and $c_2(n)$ are zero-mean uncorrelated processes, then the time average (snapshot average) of $c_1(n)c_2^*(n)$ becomes negligible, and the first two terms dominate. In this case the product beamformer behaves like a cascade $H(e^{j\omega N})G^*(e^{j\omega M})$. Next, in Fig. 5.7, the first IDFT operator produces an array bank of M filters

$$H_k(z) = H(z^N W_M^k) = \sum_{n=0}^{M-1} h(n) z^{-nN} W_M^{-nk}, \tag{5.26}$$

$0 \le k \le M-1$, where $W_M = e^{-j2\pi/M}$. Similarly, the second IDFT operator produces

$$G_\ell(z) = G(z^M W_N^\ell) = \sum_{n=0}^{N-1} g(n) z^{-nM} W_N^{-n\ell}, \tag{5.27}$$

$0 \le \ell \le N-1$. Assume that $H(e^{j\omega})$ and $G(e^{j\omega})$ are ideal low pass:

$$H(e^{j\omega}) = \begin{cases} 1 & for |\omega| < \pi/M \\ 0 & \text{otherwise.} \end{cases} \tag{5.28}$$

$$G(e^{j\omega}) = \begin{cases} 1 & for |\omega| < \pi/N \\ 0 & \text{otherwise.} \end{cases} \tag{5.29}$$

Notice that $H(e^{j\omega N})$ has N passbands, with each passband having width $2\pi/MN$, and $G(e^{j\omega M})$ has M passbands, with each passband having width $2\pi/MN$. Since $H_k(e^{j\omega})$ and $G_\ell(e^{j\omega})$ are shifted versions of $H(e^{j\omega N})$ and $G(e^{j\omega M})$, they are also multiple passband filters. Now consider the product

$$F_{\ell k}(e^{j\omega}) = G_\ell^*(e^{j\omega}) H_k(e^{j\omega}), \tag{5.30}$$

$0 \le \ell \le N-1, 0 \le k \le M-1$. Assuming M and N are coprime, it has been shown in [18] that the following are true:

1. Each of the MN filters $F_{\ell k}(e^{j\omega})$ has only one passband, and its width is $2\pi/MN$.
2. Moreover, no two of the filters $F_{\ell k}(e^{j\omega})$ have overlapping passbands, so the MN passbands completely tile $0 \le \omega < 2\pi$.

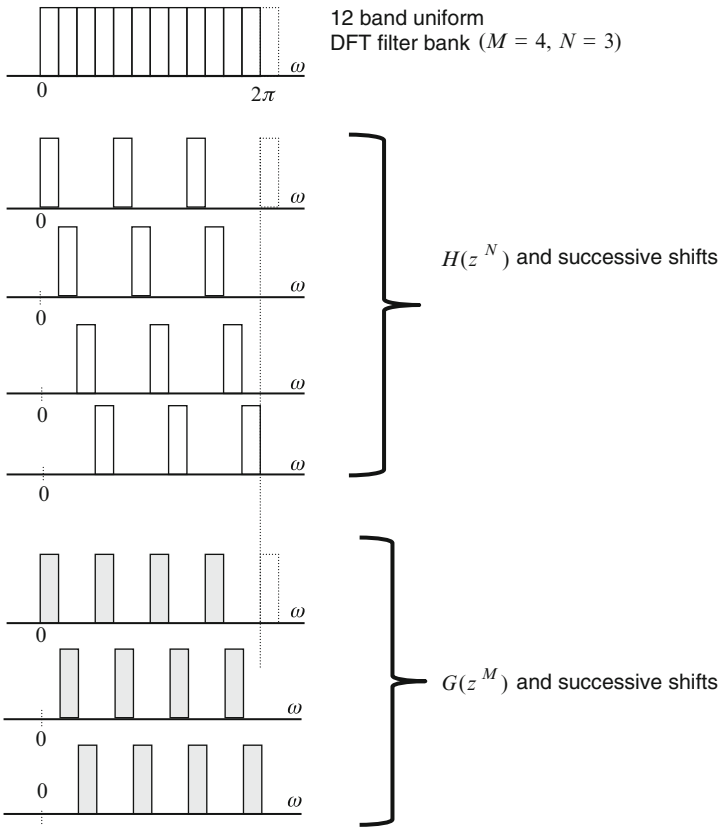

Fig. 5.8 The filters $H(z^N)$ and $G(z^M)$ and various shifted versions. Here $M = 4$ and $N = 3$

Figure 5.8 presents an example for $M = 4$ and $N = 3$. The filter $G(z^4)$ has four passbands (each with width $2\pi/12$) and three distinct shifted versions $G(z^4 W_3^\ell)$. The filter $H(z^3)$ has three passbands (each with width $2\pi/12$), and four distinct shifted versions $H(z^3 W_4^k)$. Each shifted version $G(z^4 W_3^\ell)$ overlaps with any shifted version $H(z^3 W_4^k)$ in precisely one passband. Furthermore, the twelve combinations of (k, ℓ) produce twelve distinct filters

$$H(z^3 W_4^k) G(z^4 W_3^\ell), \tag{5.31}$$

covering $0 \le \omega < 2\pi$ (shown at the top of the figure). Summarizing, we get the effect of a MN-band DFT filter bank by combining one M-band filter bank operating on M sensors and one N-band filter bank operating on N sensors, if the sensor spacings are as in Fig. 5.7, and M, N are coprime. This is achieved by performing the snapshot averaging operation, required to eliminate the cross terms in (5.25).

Simulation Example with Snapshot Averaging. In practice when the filters are not ideal, the band suppressions described above are not exact. Thus, even though the shifted versions $G(e^{j\omega M}W_N^{\ell})$ and $H(e^{j\omega N}W_M^{k})$ ideally have only one overlapping band, in practice, the "nonoverlapping bands" have overlap at the band edges. The extent of this overlap can be reduced if the filters are replaced with

$$H(z) = \sum_{n=0}^{M_1-1} h(n)z^{-n}, \quad G(z) = \sum_{n=0}^{N_1-1} g(n)z^{-n}, \tag{5.32}$$

where $M_1 > M$ and $N_1 > N$. This is equivalent to extending the antenna arrays so that there are $M_1 > M$ antenna elements in the first array and $N_1 > N$ antenna elements in the second. With appropriate choice of M_1 and N_1, the undesirable overlaps between bands of $G(e^{j\omega M}W_N^{\ell})$ and $H(e^{j\omega N}W_M^{k})$ can be reduced to any specified extent, the obvious price paid being the extra number of antenna elements. Figure 5.9 shows beams simulated by snapshot averaging of the products of the form (5.25), obtained from pairs of filters from the two filter banks. In this example $M_1 = M + 5$ and $N_1 = N + 5$. Notice how the performance improves as we go from 100 to 700 snapshots. This simulation was performed by applying a signal from a specific angle θ, measuring the filter bank outputs, averaging the products, and repeating this for a dense set of values of θ. The filter taps $h(n)$ and $g(n)$ can be designed using standard FIR filter design methods. In this example, the Remez exchange method for equiripple FIR design was used [9]. More examples can be found in [18].

5.5 Coprime Sampling in Multiple Dimensions

In two dimensions, the most general form of uniform sampling is sampling on a lattice (or a shifted version thereof). Given a square matrix $\mathbf{A}$, the set of all points of the form

$$\mathbf{t} = \mathbf{An}, \tag{5.33}$$

where $\mathbf{n}$ is an integer vector, is said to be the lattice generated by $\mathbf{A}$, and is denoted as $LAT(\mathbf{A})$ [4], [23]. Here we consider two lattices $LAT(\mathbf{M})$ and $LAT(\mathbf{N})$ generated by **integer** matrices $\mathbf{M}$ and $\mathbf{N}$. Figures 5.10(a), (b) shows two examples of such lattices. The quantity

$$\rho(\mathbf{M}) = \frac{1}{\det\mathbf{M}} \tag{5.34}$$

is equal to the density of the lattice points (number of lattice points per unit area). The region indicated as $FPD(\mathbf{M})$ in the figure (abbreviation for *fundamental parallelepiped*) is the set of all real vectors of the form

$$\mathbf{Mx}, \quad \mathbf{x} \in [0,1)^D \quad (FPD(\mathbf{M})). \tag{5.35}$$

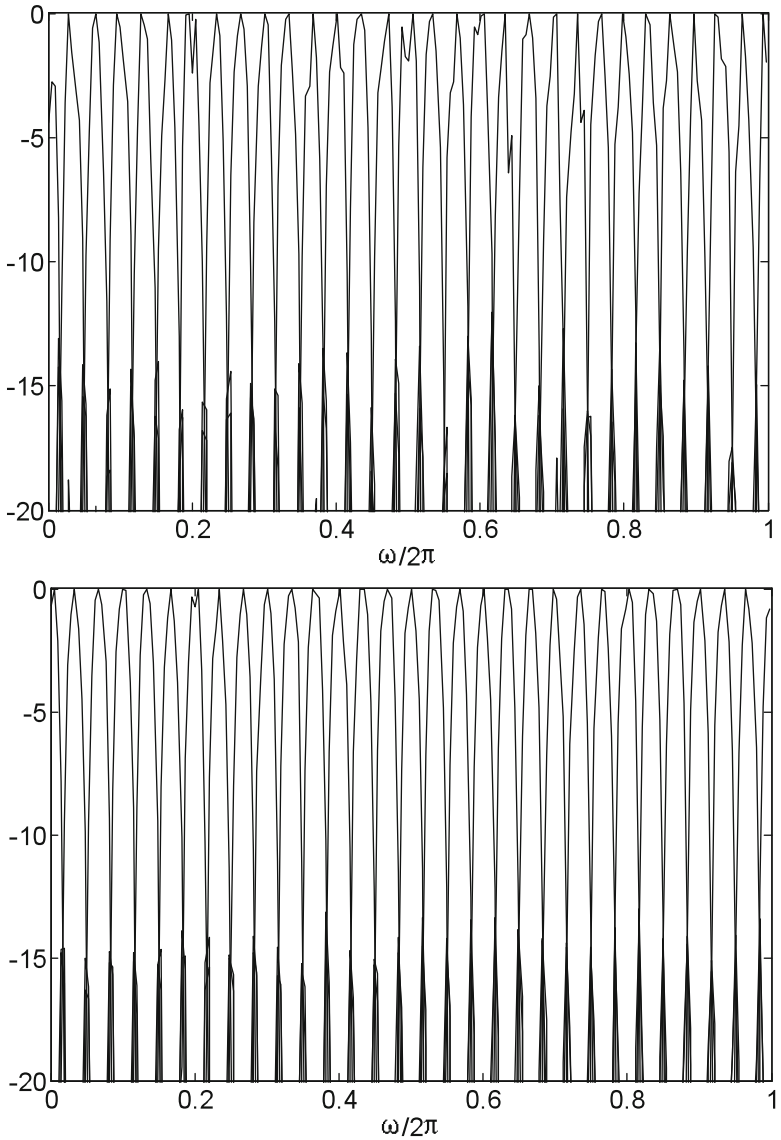

Fig. 5.9 Examples of simulated beams with $M = 6, N = 5$. Remez weights used, with $M + 5$ and $N + 5$ sensor elements. Number of snapshots $= 100$ (top) and 700 (bottom)

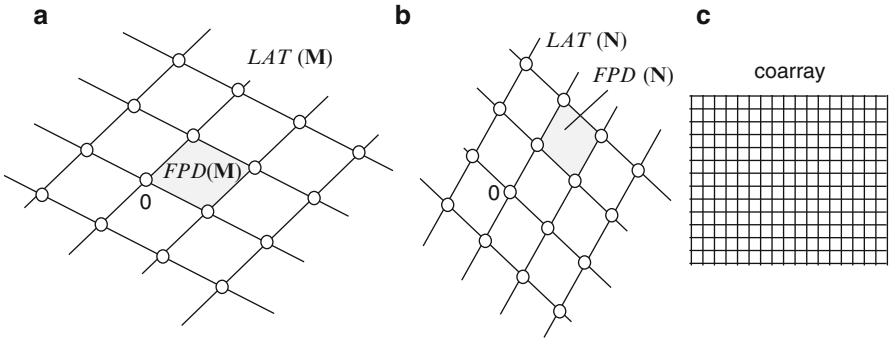

Fig. 5.10 Two-dimensional lattice arrays generated by coprime matrices and their difference coarray

The notation $\mathbf{x} \in [0,1)^D$ means that $\mathbf{x}$ is a real vector whose components are in the range $0 \leq x_i < 1$. The area of the $FPD(\mathbf{M})$ is equal to $\det\mathbf{M}$. So there are $\det\mathbf{M}$ integer vectors in it. The *symmetric parallelepiped* $SPD(\mathbf{M})$ is similarly defined as the set of points

$$\mathbf{Mx}, \quad \mathbf{x} \in (-1,1)^D \quad (SPD(\mathbf{M})). \tag{5.36}$$

We say that the sampling lattices are coprime if the matrices $\mathbf{M}$ and $\mathbf{N}$ are coprime in a certain sense to be defined below. Under this condition, many of the properties exhibited by coprime samplers in one dimension can be made to work in two dimensions also as shown in [19]. For example, it is possible to choose $\mathbf{M}$ and $\mathbf{N}$ such that the difference coarray, now defined as the set of all integer vectors of the form

$$\mathbf{k} = \mathbf{Mn}_1 - \mathbf{Nn}_2, \tag{5.37}$$

includes *all* the integers.[2] Thus, if $\det\mathbf{M}$, $\det\mathbf{N}$ are very large compared to unity, then the coarray is much denser than the sparse coprime set of samples. In practice $\mathbf{n}_1$ and $\mathbf{n}_2$ are restricted to be in a finite range, and in that case the negative of (5.37) has to be explicitly included to obtain a symmetric difference set.

We now come to the definition of coprimality. First, we say that the integer matrix $\mathbf{L}$ is a left factor of $\mathbf{M}$ if we can write $\mathbf{M} = \mathbf{LM}_1$ for some integer matrix $\mathbf{M}_1$.

[2] Strictly speaking, the difference coarray is the set of all integers of the form $\mathbf{m}_1 - \mathbf{m}_2$, $\mathbf{m}_i \in \mathscr{N}$, where $\mathscr{N}$ is in the union of all the elements in the two arrays. Thus, the coarray would include the cross differences $\mathbf{Mn}_1 - \mathbf{Nn}_2$, $\mathbf{Nn}_2 - \mathbf{Mn}_1$ and the self-differences $\mathbf{Mn}_1 - \mathbf{Mn}_1'$ and $\mathbf{Nn}_2 - \mathbf{Nn}_2'$ as in Sect. 5.2.2. For simplicity we sometimes refer to the elements $\mathbf{Mn}_1 - \mathbf{Nn}_2$ as the coarray and the union of the elements $\mathbf{Mn}_1 - \mathbf{Nn}_2$ and $\mathbf{Nn}_2 - \mathbf{Mn}_1$ as the "symmetric coarray" as in Sect. 5.2.2, since the self-differences are usually included in these cross differences.

The integer matrices $\mathbf{M}$ and $\mathbf{N}$ are said to be **left coprime** if they do not have a common left factor $\mathbf{L}$ other than unimodular matrices (i.e., matrices with $\det \mathbf{L} = \pm 1$). Right coprimality is similarly defined.

When $\mathbf{M}$ and $\mathbf{N}$ are left coprime, the lattices $LAT(\mathbf{M})$ and $LAT(\mathbf{N})$ define a coprime pair of samplers. Next, if $\mathbf{M}$ and $\mathbf{N}$ commute in multiplication, that is,

$$\mathbf{MN} = \mathbf{NM}, \tag{5.38}$$

then it can be shown [2], [19] that $\mathbf{M}$ and $\mathbf{N}$ are also right coprime. In this case we simply say that **$\mathbf{M}$ and $\mathbf{N}$ are coprime**.

5.5.1 Properties of Multidimensional Coarrays

We now mention a number of important properties of coarrays constructed from lattices. The following result is proved in [19].

Lemma 1 (Difference-set element range). *Assuming* $\mathbf{n}_1 \in FPD(\mathbf{N})$ *and* $\mathbf{n}_2 \in FPD(\mathbf{M})$ *and assuming* $\mathbf{MN} = \mathbf{NM}$, *the elements (5.37) belong in* $SPD(\mathbf{MN})$.

This is similar to the statement in one dimension that if $0 \le n_1 \le N-1$ and $0 \le n_2 \le M-1$, then the elements $Mn_1 - Nn_2$ belong in the range $-MN < k < MN$. The theorem below shows, in particular, that we can choose $\mathbf{M}$ and $\mathbf{N}$ such that there are $\det(\mathbf{MN})$ distinct elements in the set of differences $\mathbf{Mn}_1 - \mathbf{Nn}_2$, as $\mathbf{n}_1$ and $\mathbf{n}_2$ vary over

$$\mathbf{n}_1 \in FPD(\mathbf{N}), \quad \mathbf{n}_2 \in FPD(\mathbf{M}). \tag{5.39}$$

So the difference set (5.37) has at least $\det(\mathbf{MN})$ freedoms. The coarray has more (because of the negated differences $\mathbf{Nn}_2 - \mathbf{Mn}_1$), and they all belong in $SPD(\mathbf{MN})$. The following result is proved in [19].

Theorem 1 (Coprime coarrays in MD). *Assume the* $D \times D$ *nonsingular integer matrices* $\mathbf{M}$ *and* $\mathbf{N}$ *are commuting and coprime. Then:*

1. *Given any integer vector* $\mathbf{k}$, *there exist integer vectors* $\mathbf{n}_1$ *and* $\mathbf{n}_2$ *such that (5.37) holds. So the coarray contains all integer vectors when* $\mathbf{n}_1$ *and* $\mathbf{n}_2$ *vary over all integers.*
2. *Let the integer vectors* $\mathbf{n}_1, \mathbf{n}_1'$ *be restricted to* $FPD(\mathbf{N})$ *and* $\mathbf{n}_2, \mathbf{n}_2'$ *restricted to* $FPD(\mathbf{M})$. *Then*

$$\mathbf{Mn}_1 - \mathbf{Nn}_2 \neq \mathbf{Mn}_1' - \mathbf{Nn}_2', \tag{5.40}$$

   as long as $(\mathbf{n}_1, \mathbf{n}_2) \neq (\mathbf{n}_1', \mathbf{n}_2')$.
3. *The integer vectors* $\mathbf{Mn}_1$ *and* $\mathbf{Nn}_2$ *(i.e., the array elements) are distinct when the integers* $\mathbf{n}_1$ *and* $\mathbf{n}_2$ *are such that* $\mathbf{n}_1 \in FPD(\mathbf{N})$ *and* $\mathbf{n}_2 \in FPD(\mathbf{M})$, *unless* $\mathbf{n}_1 = \mathbf{n}_2 = \mathbf{0}$.

It can also be shown that if $\mathbf{n}_1 \in FPD(2\mathbf{N})$ and $\mathbf{n}_2 \in FPD(\mathbf{M})$ (or alternatively, $\mathbf{n}_1 \in FPD(\mathbf{N})$ and $\mathbf{n}_2 \in SPD(\mathbf{M})$), then all $\mathbf{k} \in FPD(\mathbf{MN})$ can be generated by $\mathbf{k} = \mathbf{M}\mathbf{n}_1 - \mathbf{N}\mathbf{n}_2$ (similar to the hole-free range of MN elements in the 1D case). All detailed proofs can be found in [19].

5.5.2 Dense 2D Frequency Tiling with Coprime DFT Filter Banks

In the one-dimensional case we explained how we can start with a coprime pair of arrays and use two DFT filter banks to obtain a dense tiling of the frequency domain by snapshot averaging (Sect. 5.4). Each DFT filter bank produces filters with multiple passbands, but when we multiply a pair of filters, one from each filter bank, then these multiple bands "mysteriously disappear" (as described in Sect. 5.4), and there remains a single band. This band is very narrow (as if it is coming from an MN-point DFT filter bank), and its position depends on the choice of the filter pair. Since there are MN filter pairs, there are MN such bands and they tile the entire frequency axis $0 \leq \omega < 2\pi$. In the 2D case, we can do something similar. We will only give a general outline here, because the details are very involved. All these details can be found in [19]. Consider the two sparse arrays on lattices defined as follows:

- Array one is on $LAT(\mathbf{N})$ with elements located at $\mathbf{Nn}, \mathbf{n} \in FPD(\mathbf{M})$, and has $\det\mathbf{M}$ elements. This is similar to having M elements spaced apart by $N\lambda/2$ in the 1D case (Fig. 5.7).
- Array two is on $LAT(\mathbf{M})$ with elements located at $\mathbf{Mm}, \mathbf{m} \in FPD(\mathbf{N})$, and has $\det\mathbf{N}$ elements. This is similar to having N elements spaced apart by $M\lambda/2$ in the 1D case (Fig. 5.7).

There are sophisticated ways to construct so-called $\mathbf{M}$-DFT filter banks for Array one by using multidimensional DFTs [4], [23]. Such a filter bank would have $\det\mathbf{M}$ filters. In this filter bank, since the sensors are on $LAT(\mathbf{N})$, the $\det\mathbf{M}$ filter responses are appropriately shifted versions of

$$H(\mathbf{N}^T\boldsymbol{\Omega}) = \sum_{\mathbf{n}} h(\mathbf{n}) e^{-j(\mathbf{N}^T\boldsymbol{\Omega})^T\mathbf{n}}. \tag{5.41}$$

The expressions for the filters are

$$H_{\mathbf{k}}(\mathbf{N}^T\boldsymbol{\Omega}) = H\Big(\mathbf{N}^T(\boldsymbol{\Omega} - 2\pi\mathbf{N}^{-T}\mathbf{M}^{-T}\mathbf{k})\Big), \tag{5.42}$$

$\mathbf{k} \in FPD(\mathbf{M}^T)$. Note that the filter number has been indexed by a vector $\mathbf{k}$ for convenience. Since $FPD(\mathbf{M}^T)$ has room for $|\det\mathbf{M}|$ integer vectors, there are $|\det\mathbf{M}|$

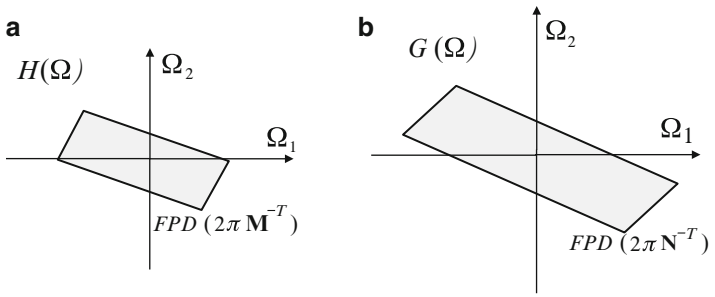

Fig. 5.11 **(a)**,**(b)** Passband regions of low-pass filters $H(\mathbf{\Omega})$ and $G(\mathbf{\Omega})$ in the two 2D DFT filter banks

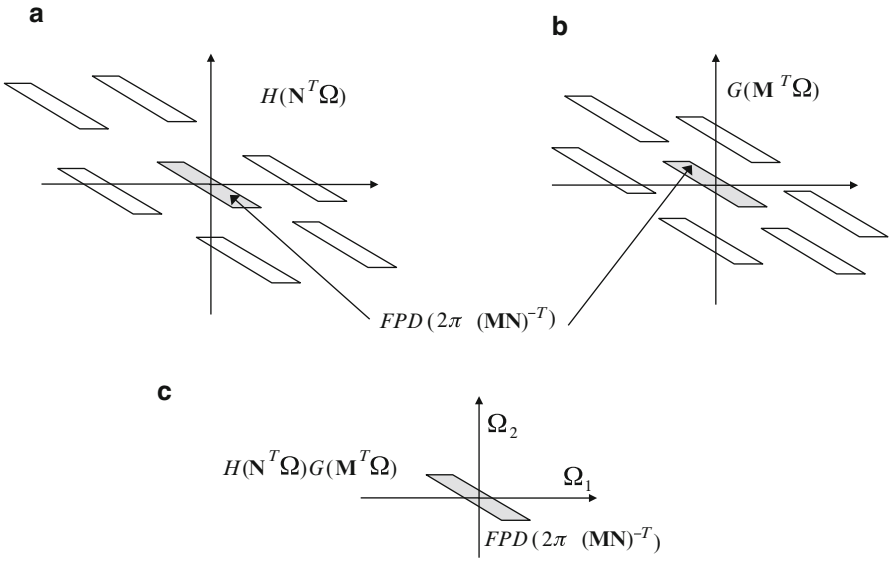

Fig. 5.12 **(a)**,**(b)** passband regions of the filters $H(\mathbf{N}^T\mathbf{\Omega})$ and $G(\mathbf{M}^T\mathbf{\Omega})$ and **(c)** the product filter

such filters. Each filter has multiple passbands (det$\mathbf{N}$ passbands to be precise) in the 2D frequency plane $[0,2\pi)^2$ because of the dependence on $\mathbf{N}^T\mathbf{\Omega}$ rather than $\mathbf{\Omega}$. Similarly, from the second array, we create an $\mathbf{N}$-DFT filter bank which has det$\mathbf{N}$ filters

$$G_{\mathbf{j}}(\mathbf{M}^T\mathbf{\Omega}) = G\Big(\mathbf{M}^T(\mathbf{\Omega} - 2\pi\mathbf{M}^{-T}\mathbf{N}^{-T}\mathbf{j})\Big), \tag{5.43}$$

$\mathbf{j} \in FPD(\mathbf{N}^T)$, and each filter has det$\mathbf{M}$ passbands. Figure 5.11 shows examples of the passband supports of $H(\mathbf{\Omega})$ and $G(\mathbf{\Omega})$, and Fig. 5.12(a), (b) gives a qualitative idea of what the multiple passband supports in $H(\mathbf{N}^T\mathbf{\Omega})$ and $G(\mathbf{M}^T\mathbf{\Omega})$ look like. The det$\mathbf{M}$ filters $H_{\mathbf{k}}(\mathbf{N}^T\mathbf{\Omega})$ and the det$\mathbf{N}$ filters $G_{\mathbf{j}}(\mathbf{M}^T\mathbf{\Omega})$ are shifted versions of the

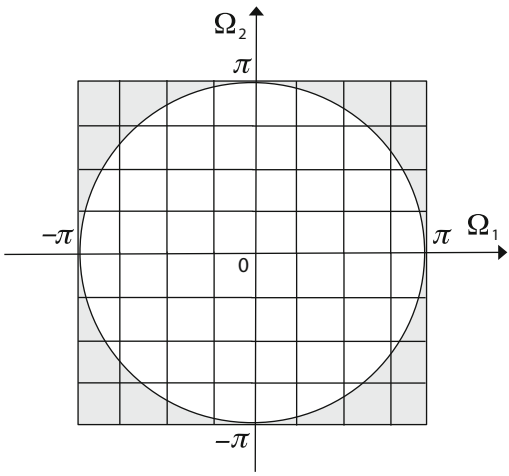

Fig. 5.13 The dense tiling of frequency plane obtained by combining two coprime filter banks. This can be achieved by using an adjugate pair of lattice generators. See text

filters $H(\mathbf{N}^T\mathbf{\Omega})$ and $G(\mathbf{M}^T\mathbf{\Omega})$, similar in principle to the filters in Fig. 5.8. Assuming that $\mathbf{M}$ and $\mathbf{N}$ are commuting and coprime, a number of results have been proved in [19] (all frequency responses are assumed to be real for simplicity):

1. If we multiply any of the $\det\mathbf{M}$ filters $H_{\mathbf{k}}(\mathbf{N}^T\mathbf{\Omega})$ with any of the $\det\mathbf{N}$ filters $G_{\mathbf{j}}(\mathbf{M}^T\mathbf{\Omega})$, then there is exactly one overlapping passband, and its shape is precisely $FPD(2\pi(\mathbf{MN})^{-T})$ (analogous to a bandpass filter with passband width $2\pi/MN$ in 1D). For the product $H(\mathbf{N}^T\mathbf{\Omega})G(\mathbf{M}^T\mathbf{\Omega})$, this is demonstrated in Fig. 5.12(c).
2. For the $\det\mathbf{MN}$ different combinations of the product $H_{\mathbf{k}}(\mathbf{N}^T\mathbf{\Omega})G_{\mathbf{j}}(\mathbf{M}^T\mathbf{\Omega})$, these passbands have different center frequencies. If we take the union of all the passband supports generated by these $\det\mathbf{MN}$ combinations of $H_{\mathbf{k}}(\mathbf{N}^T\mathbf{\Omega})$ and $G_{\mathbf{j}}(\mathbf{M}^T\mathbf{\Omega})$, then the result is the entire $2D$ frequency plane $[0,2\pi)^2$, or equivalently $[-\pi,\pi)^2$.

Thus, the product of a pair of filters produces a single passband whose area is only

$$\frac{1}{\det(\mathbf{MN})} \tag{5.44}$$

times the 2D frequency plane, and the union of all the $\det(\mathbf{MN})$ products tiles the entire frequency plane. In this way we can create a dense tiling of the frequency plane starting from sparse 2D lattice arrays.[3]

In fact we can even obtain a simple **rectangular** tiling by starting from non-rectangular lattices as demonstrated in Fig. 5.13 (the circular disk represents the so-called visible region in the array-processing context; see Sect. 5.5.3). The advantage of achieving such a rectangular tiling with the help of non-rectangular

[3]The product filters are physically realized as in the one-dimensional case, by multiplying the filter outputs as in (5.25) and performing snapshot averages.

(lattice) arrays is that there is more freedom in choice of geometry for the arrays. For example, it is possible to maximize the minimum distance between the sensors (see Sect. VI of [19]) and reduce mutual coupling. It is shown in [19] that such rectangular tilings can be achieved by making $\mathbf{N}$ equal to the adjugate of the matrix $\mathbf{M}$ (which in turn can be an arbitrary nonsingular lattice generator). We will discuss adjugates again in Sect. 5.6.

5.5.3 *Case of Monochromatic Plane Waves*

For the special case where a plane monochromatic wave is being received by the array (as in a simple array-processing setting, [24]), the "frequency vector" $\boldsymbol{\Omega}$ has the following physical meaning:

$$\boldsymbol{\Omega} = \begin{bmatrix} \Omega_1 \\ \Omega_2 \end{bmatrix} = \frac{2\pi d \sin\phi}{\lambda} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}. \tag{5.45}$$

Here d is a fixed scalar with spatial dimension, ϕ denotes the elevation angle, and θ the azimuthal angle. With $d = \lambda/2$ we have

$$\boldsymbol{\Omega} = \begin{bmatrix} \Omega_1 \\ \Omega_2 \end{bmatrix} = \pi \sin\phi \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}. \tag{5.46}$$

Note that the two components Ω_1 and Ω_2 are coupled, and

$$\Omega_1^2 + \Omega_2^2 = \pi^2 \sin^2\phi. \tag{5.47}$$

If we draw a circle with radius $\pi \sin\phi$ in the (Ω_1, Ω_2) plane, then all points on this circle represent the same elevation angle ϕ. At any point on this circle, the angle θ represents the azimuthal angle. This is shown in Fig. 5.14(a). In the rectangular frequency region defined by $-\pi \le \Omega_1, \Omega_2 < \pi$, the region which maps into meaningful (θ, ϕ) pairs is a disk with radius π, centered at the origin. This disk can therefore be regarded as the *visible region* of the (Ω_1, Ω_2) plane and is shown in Fig. 5.14(b).

5.6 Commuting Coprime Matrices

In this section we elaborate on the mathematical problem of constructing integer matrices $\mathbf{M}$ and $\mathbf{N}$ which are commuting and coprime. Throughout the section the greatest common divisor (gcd) of integers $a, b, c, \ldots$ is denoted as $(a, b, c, \ldots)$. Most of the results reviewed here are from [12] and [20].

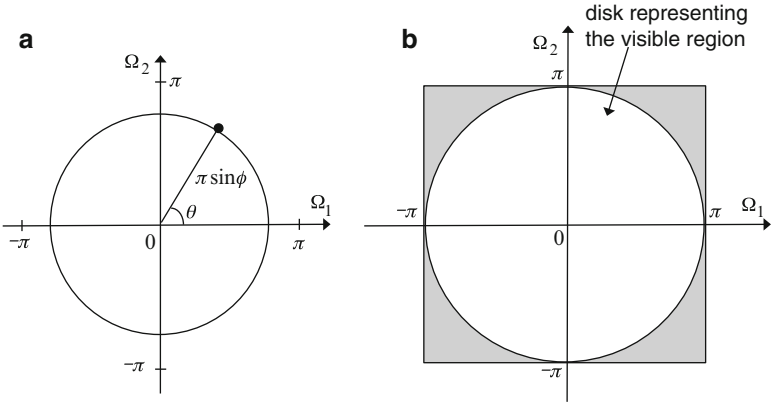

Fig. 5.14 **(a)** The relation between the frequencies (Ω_1, Ω_2) and the angles θ, ϕ in the context of array processing. **(b)** The visible region of the (Ω_1, Ω_2) plane in the context of array processing for plane monochromatic waves. See text

5.6.1 Some Important Families of Coprime Integer Matrices

Coprimality conditions for a number of 2×2 integer matrix pairs are presented in [12]. The results were based on the use of Bezout's identity for integer matrices which says that if the integer matrices $\mathbf{P}$ and $\mathbf{Q}$ are left coprime, there exist integer matrices $\mathbf{A}$ and $\mathbf{B}$ such that

$$\mathbf{PA} + \mathbf{QB} = \mathbf{I}. \tag{5.48}$$

See [19] for further details on Bezout's identity for integer matrices. In this section we shall briefly summarize some of the main results from [12]. First consider two *circulant matrices*

$$\mathbf{P} = \begin{bmatrix} p & q \\ q & p \end{bmatrix}, \quad \mathbf{P}_1 = \begin{bmatrix} p_1 & q_1 \\ q_1 & p_1 \end{bmatrix}. \tag{5.49}$$

These always commute automatically ($\mathbf{PP}_1 = \mathbf{P}_1\mathbf{P}$). The following result was proved in [12].

Theorem 2. *The circulant matrices (5.49) are coprime if and only if*

$$(p+q, p_1+q_1) = 1 \quad \text{and} \quad (p-q, p_1-q_1) = 1. \tag{5.50}$$

That is, the two DFT coefficients of the top row of $\mathbf{P}$ *should be coprime to the corresponding DFT coefficients from* $\mathbf{P}_1$.

Next, consider the two *skew-circulant* matrices

$$\mathbf{P} = \begin{bmatrix} p & q \\ -q & p \end{bmatrix}, \quad \mathbf{P}_1 = \begin{bmatrix} p_1 & q_1 \\ -q_1 & p_1 \end{bmatrix}. \tag{5.51}$$

These commute automatically. The following result was proved in [12].

Theorem 3. *The skew-circulant matrices (5.51) are coprime if and only if*

$$(p^2+q^2, p_1^2+q_1^2, p_1q-q_1p) = 1, \tag{5.52}$$

and furthermore, this condition can be rewritten in the form $(p^2+q^2, p_1^2+q_1^2, pp_1+qq_1) = 1$.

The special case where two matrices are adjugates of each other is particularly interesting.[4] The circulant adjugate pair has the form

$$\mathbf{P} = \begin{bmatrix} p & q \\ q & p \end{bmatrix}, \quad \mathbf{P}_1 = \begin{bmatrix} p & -q \\ -q & p \end{bmatrix}, \tag{5.53}$$

whereas the skew-circulant adjugate pair has the form

$$\mathbf{P} = \begin{bmatrix} p & q \\ -q & p \end{bmatrix}, \quad \mathbf{P}_1 = \begin{bmatrix} p & -q \\ q & p \end{bmatrix}. \tag{5.54}$$

These matrix pairs are obtained by setting $p_1 = p$ and $q_1 = -q$ in the general forms. It can then be shown [12] that coprimality conditions (5.50) for the circulant pair simplify to the single condition

$$(p+q, p-q) = 1, \tag{5.55}$$

which can be restated as follows:

$$(p,q) = 1, \quad \text{and} \quad p \text{ and } q \text{ have opposite parity.} \tag{5.56}$$

It can also be shown that the skew-circulant adjugate pair (5.54) is coprime if and only if (5.55) holds. Thus, surprisingly, the condition (5.52) is equivalent to (5.55) whenever $p_1 = p$ and $q_1 = -q$. The circulant adjugate pair and the skew-circulant adjugate pair therefore share the same coprimality condition.

[4] The adjugate $\widehat{\mathbf{P}}$ of $\mathbf{P}$ is the matrix of cofactors [5] which arises, for example, in the expression for the inverse $\mathbf{P}^{-1} = \widehat{\mathbf{P}}/\det\mathbf{P}$. In the 2×2 case, $\mathbf{P}$ is also the adjugate of $\widehat{\mathbf{P}}$, so we have an adjugate pair.

Another result proved in [12] is that the triangular matrices

$$\mathbf{P} = \begin{bmatrix} p & q \\ 0 & s \end{bmatrix}, \quad \mathbf{P}_1 = \begin{bmatrix} p_1 & q_1 \\ 0 & s_1 \end{bmatrix}, \tag{5.57}$$

are left coprime if and only if

$$(s, s_1) = 1 \quad \text{and} \quad (p, p_1, q_1 s - q s_1) = 1. \tag{5.58}$$

For the special case of adjugate pairs

$$\mathbf{P} = \begin{bmatrix} p & q \\ 0 & s \end{bmatrix}, \quad \mathbf{P}_1 = \begin{bmatrix} s & -q \\ 0 & p \end{bmatrix}, \tag{5.59}$$

the coprimality condition (5.58) reduces to

$$(p, s) = 1. \tag{5.60}$$

Next, the 3×3 triangular matrix $\mathbf{P}$ and its adjugate $\widehat{\mathbf{P}}$

$$\mathbf{P} = \begin{bmatrix} p & q & r \\ 0 & s & t \\ 0 & 0 & u \end{bmatrix}, \quad \widehat{\mathbf{P}} = \begin{bmatrix} su & -qu & qt - rs \\ 0 & pu & -pt \\ 0 & 0 & ps \end{bmatrix} \tag{5.61}$$

can be shown to be coprime [12] if and only if the diagonal elements of $\mathbf{P}$ are coprime in pairs, that is, $(p, s) = (p, u) = (s, u) = 1$.

5.6.2 Generating Coprime Pairs Based on gcd of Minors

We now present a general condition for left coprimality [20] which holds for square matrices of any size. This is a powerful result which allows us to generate a large class of coprime matrices of arbitrary sizes, as we shall demonstrate.

Theorem 4. *The $D \times D$ integer matrices $\mathbf{M}$ and $\mathbf{N}$ are left coprime if and only if the gcd Δ_D of all the $D \times D$ minors of the matrix*

$$\begin{bmatrix} \mathbf{M} & \mathbf{N} \end{bmatrix} \tag{5.62}$$

is equal to unity.

A similar result was also reported in some earlier papers in different contexts [26], [15]. A simple proof can be found in [20], and is based on the Smith-form decomposition [6], [23] of the matrix (5.62). The theorem implies in particular that the gcd Δ_i of all the $i \times i$ minors of (5.62) is also unity, for $1 \leq i \leq D-1$ (since Δ_i is a factor of Δ_{i+1}). We will present a number of coprime families of matrices in this section. But first, we mention some simple consequences of Theorem 4.

1. If there is a $D \times D$ minor equal to unity, then $\Delta_D = 1$, and we conclude immediately that $\mathbf{M}$ and $\mathbf{N}$ are coprime.
2. *Bipolar matrices*. If the elements of an integer matrix can only have the values 1 and -1, we call it a bipolar matrix. Here is a 3×3 example:

$$\begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}. \tag{5.63}$$

   For such matrices the 2×2 minors can only be $0, 2$, or -2. Since 2 is a common factor of such a set of numbers, the gcd of the 2×2 minors cannot satisfy the necessary condition $\Delta_2 = 1$. So a consequence of Theorem 4 is that *two bipolar matrices can never be coprime*.
3. *Mixing the columns*. From the set of $2D$ columns of the $D \times D$ matrices $\mathbf{P}$ and $\mathbf{Q}$, suppose we pick any set of D columns and form the matrix $\mathbf{P}_{new}$, and with the remaining D columns form the matrix $\mathbf{Q}_{new}$. Then $\mathbf{P}_{new}$ and $\mathbf{Q}_{new}$ are left coprime if and only if $\mathbf{P}$ and $\mathbf{Q}$ are left coprime. This is because the set of all $D \times D$ minors of the composite matrix $\begin{bmatrix} \mathbf{P} & \mathbf{Q} \end{bmatrix}$ is the same as the set of all $D \times D$ minors of the composite matrix $\begin{bmatrix} \mathbf{P}_{new} & \mathbf{Q}_{new} \end{bmatrix}$. For example, recall that the circulant and its adjugate given in (5.53) are coprime if and only if (5.55) holds. Now consider the skew circulant and its adjugate given in (5.54). Since the four columns in (5.54) are identical to the four columns in (5.53), it follows readily that the skew-circulant pair is coprime under the same condition (5.55). This is therefore a second way to look at the skew-circulant case, which was handled in [20] using a different technique, namely the use of Bezout's identity for integer matrices.

5.6.2.1 Generalized Circulants

We now consider a 3×3 integer matrix of the form

$$\mathbf{P} = \begin{bmatrix} p & q & r \\ \alpha r & p & q \\ \alpha q & \alpha r & p \end{bmatrix}, \tag{5.64}$$

where α, p, q, and r are integers. These are called **generalized circulants**. It is readily verified that two generalized circulants with the same α commute. They reduce to circulants for $\alpha = 1$ and skew circulants for $\alpha = -1$. It can be verified that the adjugate of $\mathbf{P}$ has the same form:

$$\widehat{\mathbf{P}} = \begin{bmatrix} p^2 - \alpha qr & \alpha r^2 - pq & q^2 - pr \\ \alpha(q^2 - pr) & p^2 - \alpha qr & \alpha r^2 - pq \\ \alpha(\alpha r^2 - pq) & \alpha(q^2 - pr) & p^2 - \alpha qr \end{bmatrix}. \tag{5.65}$$

The commuting matrices $\mathbf{P}$ and $\widehat{\mathbf{P}}$ are coprime if and only if the 3×3 minors of the matrix $\left[\mathbf{P}\ \widehat{\mathbf{P}}\right]$ have gcd $= 1$. The number of such minors is

$$\binom{2D}{D} = \binom{6}{3} = 20. \tag{5.66}$$

Even though the problem of identifying the gcd of these minors appears formidable, there is plenty of structure in the matrix and therefore in the minors. Based on these, the following result has been proved in [20].

Theorem 5. *The generalized circulant matrix* $\mathbf{P}$ *and its adjugate* $\widehat{\mathbf{P}}$ *are coprime if and only if*

$$\Big(\det\mathbf{P},\, 3(p^2 - \alpha qr),\, 3(\alpha r^2 - pq),\, 3\alpha(q^2 - pr)\Big) = 1. \tag{5.67}$$

Here $\det \mathbf{P} = p^3 + \alpha q^3 + \alpha^2 r^3 - 3\alpha pqr.$

Other equivalent ways of stating this are also given in [20]. From the above result, we deduce that a 3×3 circulant ($\alpha = 1$) is coprime to its adjugate if and only if

$$(\det\mathbf{P},\, 3(p^2 - qr),\, 3(r^2 - pq),\, 3(q^2 - pr)) = 1. \tag{5.68}$$

It is shown in [20] that this is equivalent to simultaneously satisfying the following two conditions:

$$(p + q + r,\, 3) = 1, \quad (p^2 - qr,\, q^2 - pr,\, r^2 - pq) = 1. \tag{5.69}$$

Similarly, a 3×3 skew circulant ($\alpha = -1$) is coprime to its adjugate if and only if

$$(p - q + r,\, 3) = 1, \quad (p^2 + qr,\, q^2 - pr,\, r^2 + pq) = 1. \tag{5.70}$$

In this section many of the examples are circulants, skew circulants, or generalized versions. The importance of circulants and skew circulants in 2D applications is explained in [19] in some detail.

5.6.2.2 Adjugates

Many of the examples considered in this section were adjugate pairs. In the DFT filter bank application described in Sect. 5.5.2, we mentioned that it is possible to obtain a dense tiling of the 2D frequency plane by starting from a coprime pair of DFT filter banks. This is done by performing snapshot averages of the products of pairs of outputs from the two filter banks as in Sect. 5.4. In the 2D case it was mentioned that the dense tiling can even be a rectangular tiling, if the two sparse lattices $LAT(\mathbf{M})$ and $LAT(\mathbf{N})$ are appropriately chosen. For example, if $\mathbf{N}$ is the

adjugate of **M**, this is indeed the case. This is the importance of adjugate pairs in 2D array processing. An interesting result on coprimality of adjugate pairs is the following.

Theorem 6. *Coprimality of adjugate pairs. Consider the integer matrix* **P** *and its adjugate* $\widehat{\mathbf{P}}$ *given by*

$$\mathbf{P} = \begin{bmatrix} p & q \\ r & s \end{bmatrix}, \quad \widehat{\mathbf{P}} = \begin{bmatrix} s & -q \\ -r & p \end{bmatrix}. \tag{5.71}$$

These are coprime if and only if

$$(ps - rq,\ s + p) = 1, \tag{5.72}$$

that is, if and only if the determinant of **P** *and the trace of* **P** *are coprime.*

This result was proved in [20], based on Theorem 4. In fact this result can be used to rederive the coprimality conditions for many of the 2×2 adjugate pairs mentioned earlier in Sect. 5.6.1.

5.7 System Identification

Consider Fig. 5.15 which shows a continuous-time linear time invariant (LTI) system with impulse response $h_c(t)$. Imagine we apply a discrete-time input $x(n)$ with sample spacing NT as shown, where $N > 0$ is an integer. Thus

$$y_c(t) = \sum_m x(m) h_c(t - mNT). \tag{5.73}$$

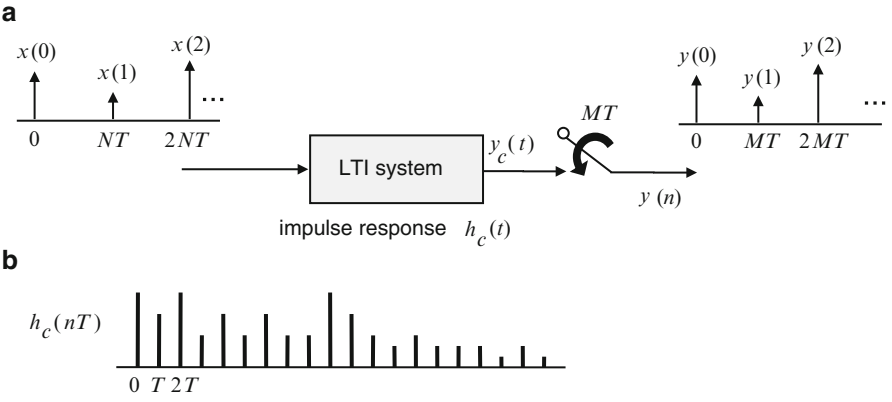

Fig. 5.15 **(a)** A continuous-time channel with input stream at sparse rate $1/NT$, and output sampled at the sparse rate $1/MT$. **(b)** Samples of the channel impulse response, with dense spacing T

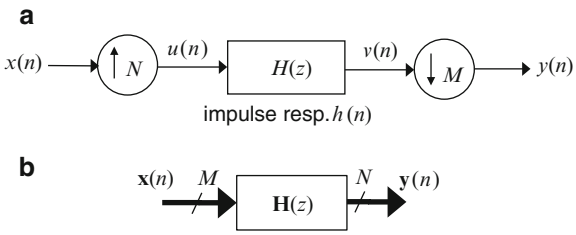

Fig. 5.16 **(a)** Representation of the system (5.76) using multirate building blocks, **(b)** its MIMO equivalent representation

In the figure the system output is sampled with a different spacing MT (for some integer $M > 0$) so that

$$y(n) = \sum_m x(m) h_c((Mn - Nm)T) = \sum_m x(m) h(Mn - Nm), \tag{5.74}$$

where

$$h(k) = h_c(kT) \tag{5.75}$$

is the sampled version of the impulse response, with a denser sample spacing T than that of the input and the output. Now assume that M and N are coprime. This means that given any integer k, there exist integers m and n such that $k = Mn - Nm$. That is, the specific samples $y(n)$ and $x(m)$ are related by $h(k)$. Thus every sample $h(k)$, taken at the dense spacing T, participates in the input-output relationship, even though the input and output samples are spaced apart sparsely (by NT and MT).

The main point we wish to make is that, from a knowledge of the sparse sequences $x(n)$ and $y(n)$ in Fig. 5.15(a), it is actually possible to identify the dense set of samples of the system $h(n)$ in Fig. 5.15(b) completely. The details can be found in [21]. To gain insight into this result, we start from the input-output relation

$$y(n) = \sum_m x(m) h(Mn - Nm) \tag{5.76}$$

which would have been a convolution if $M = N = 1$. But for arbitrary integers M, N (5.76) does not represent a convolution. It can however be represented as in Fig. 5.16(a), where the building blocks $\uparrow N$ and $\downarrow M$ are defined by the following input-output relationships:

$$u(n) = \begin{cases} x(n/N) & \text{if } n \text{ is a multiple of } N \\ 0 & \text{otherwise} \end{cases} \tag{5.77}$$

and

$$y(n) = v(Mn). \tag{5.78}$$

Here $\uparrow N$ is called an N-fold **expander** and $\downarrow M$ is called an M-fold **decimator**. These are time-varying building blocks [23], so the system in Fig. 5.16(a) is a linear

and time-varying system. However, it can be shown [23], [21] that this system can equivalently be represented by a multi-input multi-output (MIMO) LTI system. To explain this, define the vector

$$\mathbf{x}(n) = \begin{bmatrix} x(Mn) & x(Mn+1) & \dots & x(Mn+M-1) \end{bmatrix}^T \tag{5.79}$$

which is called the **blocked version** of $x(n)$ with block size M. If we imagine that $x(n)$ is divided into nonoverlapping blocks of M successive samples, then each $\mathbf{x}(n)$ represents one such block. Similarly, let

$$\mathbf{y}(n) = \begin{bmatrix} y(Nn) & y(Nn+1) & \dots & y(Nn+N-1) \end{bmatrix}^T \tag{5.80}$$

be the **blocked version** of $y(n)$ with block size N. It can then be shown that

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z). \tag{5.81}$$

That is, the system $\mathbf{x}(n) \mapsto \mathbf{y}(n)$ is a $N \times M$ MIMO LTI system with transfer matrix $\mathbf{H}(z)$. See Fig. 5.16(b). The elements $H_{km}(z)$ in this matrix are related to the polyphase components [23] of the scalar filter $H(z)$. To explain this, write $H(z)$ in the form

$$H(z) = \sum_{i=0}^{MN-1} z^{-i} E_i(z^{MN}). \tag{5.82}$$

This can always be done by defining the ith polyphase component as

$$E_i(z) = \sum_n h(MNn+i) z^{-n}. \tag{5.83}$$

For arbitrary M and N, the (k,m)th element $H_{k,m}(z)$ of the matrix $\mathbf{H}(z)$ can be determined as follows: first express the integer $Mk - Nm$ as

$$Mk - Nm = i_0 + MNi_1, \tag{5.84}$$

where $0 \le i_0 \le MN-1$ is the remainder modulo MN. With this choice i_1 is either 0 or -1. It can then be shown [21] that

$$H_{km}(z) = z^{i_1} E_{i_0}(z). \tag{5.85}$$

For example, suppose $M = 2$ and $N = 3$. Then, working through the above details, one can verify that

$$\mathbf{H}(z) = \begin{bmatrix} E_0(z) & z^{-1}E_3(z) \\ E_2(z) & z^{-1}E_5(z) \\ E_4(z) & E_1(z) \end{bmatrix}. \tag{5.86}$$

Even though Fig. 5.16(a) can always be represented as in Fig. 5.16(b), we have to be careful with the converse. Given an arbitrary $N \times M$ LTI system $\mathbf{H}(z)$ we can draw it as in Fig. 5.16(a) for some scalar LTI system $H(z)$ if and only if M and N are coprime [22].

The main insight we gain from the MIMO representation is that the identification of the system $H(z)$ in Fig. 5.16(a) is equivalent to the identification of the MIMO LTI system $\mathbf{H}(z)$ from measurements of its input and output. A direct proof that such identification can indeed be done can be found in [21].

5.8 Fractionally Spaced Equalizers

Before concluding this chapter we would like to briefly mention the use of coprime sampling rates in digital communication systems for the purpose of channel equalization. Consider Fig. 5.17(a) which shows a communication channel assumed to be an LTI system with impulse response $h_c(t)$. The channel input is at the rate $1/NT$ (symbol rate), and the output of the channel is sampled at a rate $1/MT$ at the receiver. So the receiver sampling rate is N/M times the symbol rate. With $N > M$ we therefore have an **oversampling receiver**. For example, $N = 2M$ implies oversampling by two, and $N = 3M/2$ implies oversampling by 1.5. It is well known that oversampling brings in some advantages [14], [17]. With no loss of generality we can assume M and N to be coprime.

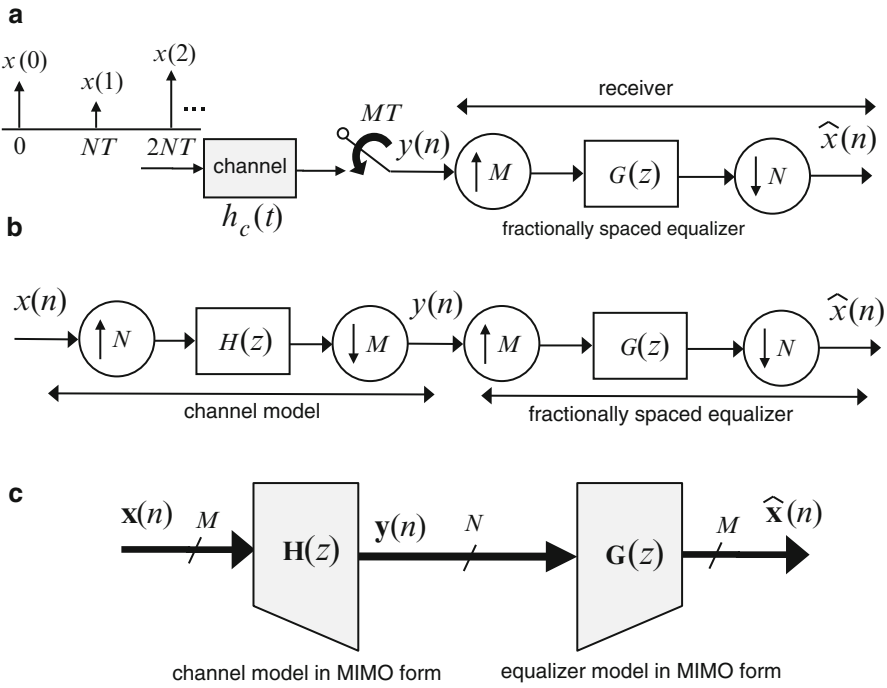

Fig. 5.17 **(a)** A channel $h_c(t)$ with output sampling rate $1/MT$ greater than the symbol rate $1/NT$ and the fractionally spaced equalizer $G(z)$, **(b)** equivalent multirate representation, and **(c)** equivalent MIMO LTI representation

The receiver has to process the samples $y(n)$ to produce the symbol-rate signal $\hat{x}(n)$ (equalized signal). This is done by using the sampling-rate-alternation system shown on the receiver side of Fig. 5.17(a), where $G(z)$ is called a fractionally spaced equalizer (FSE). The discrete-time equivalent of the channel with input and output rates $1/NT$ and $1/MT$ was already shown to be as in the left side of Fig. 5.17(b). Now, by using the fact that the system in Fig. 5.16(a) has the representation shown in Fig. 5.16(b), we can redraw the entire system as shown in Fig. 5.17(c). In this system $\mathbf{x}(n)$ and $\widehat{\mathbf{x}}(n)$ are the M-blocked versions of the transmitted and received symbol streams. If the equalizer $\mathbf{G}(z)$ can be designed such that $\widehat{\mathbf{x}}(n) = \mathbf{x}(n)$ (in absence of channel noise, which is not shown here), then we have a **zero-forcing equalizer**. From the figure it is clear that $\mathbf{G}(z)$ is a zero-forcing equalizer if and only if

$$\mathbf{G}(z)\mathbf{H}(z) = \mathbf{I}. \tag{5.87}$$

Since $\mathbf{H}(z)$ is $N \times M$ with $N > M$, we can partition $\mathbf{G}(z)$ and $\mathbf{H}(z)$ to rewrite this as follows:

$$\begin{bmatrix}\mathbf{G}_0(z) & \mathbf{G}_1(z)\end{bmatrix} \begin{bmatrix}\mathbf{H}_0(z) \\ \\ \mathbf{H}_1(z)\end{bmatrix} = \mathbf{I}, \tag{5.88}$$

where $\mathbf{G}_0(z)$ has M columns and $\mathbf{G}_1(z)$ has $N-M$ columns. Thus the zero-forcing condition is

$$\mathbf{G}_0(z)\mathbf{H}_0(z) + \mathbf{G}_1(z)\mathbf{H}_1(z) = \mathbf{I}. \tag{5.89}$$

For the case where $H(z)$ is FIR, $\mathbf{H}(z)$ is a polynomial matrix, that is, it has the form $\mathbf{H}(z) = \sum_{n=0}^{J} \mathbf{h}(n)z^{-n}$. In this case there will exist an FIR equalizer $\mathbf{G}(z)$ such that (5.89) holds, as long as $\mathbf{H}_0(z)$ and $\mathbf{H}_1(z)$ are right coprime. This follows from Bezout's identity for polynomial matrices [6], [23], which is similar to Bezout's identity for integer matrices (5.48) in Sect. 5.6.1).

In fact any partition with $r > 0$ columns for $\mathbf{G}_0(z)$ and $N - r > 0$ columns for $\mathbf{G}_1(z)$ can be used in (5.88). It can be shown [22] that as long as the polynomial matrix $\mathbf{H}(z)$ does not have a polynomial right factor $\mathbf{R}(z)$ other than unimodular matrices, there will exist polynomial $\mathbf{G}(z)$ to satisfy (5.87). For further details on these systems the reader should refer to [25], [26], and references therein.

5.9 Concluding Remarks

In this chapter we gave an overview of coprime sampling and its applications, both in one dimension and in multiple dimensions. The discussions have also opened up some interesting problems for future research. For example, recall that the autocorrelation can be estimated at the dense rate $1/T$ from the samples taken at the sparse rates $1/MT$ and $1/NT$. This is true regardless of how sparse the two

samplers are. However, there is obviously a price paid for making them arbitrarily sparse. For example, the window of time required to collect data, in order to get a certain accuracy in the autocorrelation estimate, will have to be too long if M and N are very large. In the case of spatial sampling, large M and N imply that there is less mutual coupling between sensors, which is good. On the other hand, arbitrarily large spacing means that more physical area or aperture is required. Such trade-offs have to be studied carefully.

In two dimensions, given a sampling density for the sparse samplers, namely $1/\det\mathbf{M}$ and $1/\det\mathbf{N}$, there are many lattices that can realize these densities. Thus the optimal choice of lattices is also an open problem. Many examples were given in Sect. 5.5 and in [19], but no detailed study has been made about "optimal" choices.

Another interesting problem is the possibility of errors in sensor locations. Since the dense coarray with spacing $\lambda/2$ is obtained by taking the differences of elements in a sparse array, errors in sensor positions affect those in the coarray. A very dense coarray with large "errors" will create errors in results (such as DOA estimates). Another issue is the analysis of the accuracy of DOA estimates, which depends on the number of snapshots. No such analysis is available for the case of estimates based on the dense coarray. With regard to the application of coprime sampling in system identification (Sect. 5.7), the application for channel identification in digital communications is an interesting problem that remains to be explored because, in such applications, the input $x(n)$ is known only for short stretches of time (pilot modes, during which no other signal transmission takes place). In short, it appears that several interesting problems and new directions have been opened up by the theory of coprime sampling.

Acknowledgments This work was supported in parts by the Office of Naval Research grant N00014-08-1-0709 and the California Institute of Technology.

References

1. Abramovich YI, Gray DA, Gorokhov AY, Spencer NK (1998) positive-definite toeplitz completion in DOA estimation for nonuniform linear antenna arrays. I. Fully augmentable arrays. IEEE Trans Signal Proc 46:2458–2471
2. Chen T, Vaidyanathan PP (1993) The role of integer matrices in multidimensional multirate systems. IEEE Trans Signal Proc SP-41:1035–1047
3. Chen CY, Vaidyanathan, PP (2008) Minimum redundancy MIMO radars. IEEE Int Symp Circ Syst 45–48
4. Dudgeon, DE, Mersereau, RM (1984) Multidimensional digital signal processing, Prentice Hall, Inc., Englewoods Cliffs
5. Horn RA, Johnson, CR (1985) Matrix analysis. Cambridge University Press, Cambridge
6. Kailath, T (1980) Linear Systems, Prentice Hall, Inc., Englewood Cliffs
7. Hoctor RT, Kassam, SA (1990) The unifying role of the coarray in aperture synthesis for coherent and incoherent imaging. Proceedings of the IEEE. 78:735–752
8. Moffet A (1968) Minimum-redundancy linear arrays. IEEE Trans Antenn Propag 16: 172–175

9. Oppenheim AV, Schafer RW (1999) Discrete time signal processing, Prentice Hall, Inc., Englewood Cliffs
10. Pal P , Vaidyanathan, PP (2011) Coprime sampling and the MUSIC algorithm. Proc of the 14th IEEE Digital Signal Processing Workshop, Sedona, AZ
11. Pal P, Vaidyanathan PP (2010) Nested arrays: a novel approach to array processing with enhanced degrees of freedom. IEEE Trans Signal Process 58: 4167–4181
12. Pal P, Vaidyanathan PP (2011) Coprimality of certain families of integer matrices, IEEE Trans Signal Process 59: 1481–1490
13. Pillai SU, Bar-Ness, Y, Haber F (1985) A new approach to array geometry for improved spatial spectrum estimation. Proc IEEE 73: 1522–1524
14. Proakis JG (1995) Digital communications. McGraw Hill, Inc., New York
15. Rajagopal R, Potter LC (2003) Multivariate MIMO FIR inverses. IEEE Trans Image Proc12: 458–465
16. Skolnik MI (2001) Introduction to radar systems. McGraw-Hill, NY
17. Treichler JR, Fijalkow I, Johnson CR (1996) Fractionally spaced equalizers: how long should they be? IEEE Signal Process Mag 13:65–81
18. Vaidyanathan PP, Pal P (2011) Sparse sensing with coprime samplers and arrays. IEEE Trans Signal Process 59:573–586
19. Vaidyanathan PP, Pal P (2011) Theory of sparse coprime sensing in multiple dimensions. IEEE Trans Signal Process 59:3592–3608
20. Vaidyanathan PP, Pal P (2011) A general approach to coprime pairs of matrices, based on minors. IEEE Trans Signal Process 59:3536–3548
21. Vaidyanathan PP, Pal P (2010) System identification with sparse coprime sensing. IEEE Signal Process Lett 17:823–826
22. Vaidyanathan PP, Pal P (2011) Coprime Sampling for System Stabilization with FIR Multirate Controllers. Proc IEEE Asilomar Conf on Signals, Systems, and Compuers
23. Vaidyanathan PP (1993) Multirate systems and filter bank. Prentice Hall, Inc., Englewood Cliffs
24. Van Trees HL (2002) Optimum array processing: part IV of detection, estimation and modulation theory, Wiley Interscience, NY
25. Vrcelj B, Vaidyanathan PP (2002) MIMO biorthogonal partners and applications, IEEE Trans Sig Process 50: 528–542
26. Vrcelj B, Vaidyanathan PP (2003) Fractional biorthogonal partners in channel equalization and signal interpolation. IEEE Trans Signal Process 51:1928–1940
27. Xia XG (1999) On estimation of multiple frequencies in undersampled complex valued waveforms. IEEE Trans Signal Process 47:3417–3419
28. Xia XG, Liu K (2005) A generalized Chinese remainder theorem for residue sets with errors and its application in frequency determination from multiple sensors with low sampling rates. IEEE Signal Process Lett 12:768–771

Chapter 6
Chromatic Expansions and the Bargmann Transform

Ahmed I. Zayed

Abstract Chromatic series expansions of bandlimited functions have recently been introduced in signal processing with promising results. Chromatic series share similar properties with Taylor series insofar as the coefficients of the expansions, which are called chromatic derivatives, are based on the ordinary derivatives of the function, but unlike Taylor series, chromatic series have more practical applications. The Bargmann transform was introduced in 1961 by V. Bargmann who showed, among other things, that the Bargmann transform is a unitary transformation from $L^2(\mathbb{R}^n)$ onto the Bargmann–Segal–Foch space $\mathfrak{F}$ on which Foch's operator solutions to some equations in quantum mechanics are realized.

The goal of this article is to survey results on chromatic derivatives and explore the connection between chromatic derivatives and series on the one hand and the Bargmann transform and the Bargmann–Segal–Foch space on the other hand.

6.1 Introduction

Chromatic derivatives and series expansions have recently been introduced by A. Ignjatovic in [14, 15] as an alternative representation to Taylor series for bandlimited functions, and they have been shown to be more useful in practical applications than Taylor series; see [4, 6, 11–13, 22, 25, 26].

The nth chromatic derivative $K^n[f](t_0)$ of an analytic function $f(t)$ at t_0 is a linear combination of the ordinary derivatives $f^{(k)}(t_0)$, $0 \leq k \leq n$, where the coefficients of the combination are based on systems of orthogonal polynomials.

Chromatic derivatives are intrinsically related to the Fourier transformation. They are constructed using the fact that, under the Fourier transformation, differentiation in the time domain corresponds to multiplication by powers of ω in the frequency domain.

A.I. Zayed (✉)
Department of Mathematical Sciences, DePaul University, Chicago, IL 60614, USA
e-mail: azayed@condor.depaul.edu

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_6,

In two recent papers [27, 28] we introduced more general types of chromatic derivatives and series that are better suited to handle integral transforms other than the Fourier transform. In [28] we presented two different methods to construct a differential operator L that gives rise to generalized chromatics derivatives and their associated integral transform. In the first method the operator L arises from certain Sturm–Liouville boundary-value problems, while in the second, it arises from initial-value problems involving differential operators of order n.

The states of a quantum mechanical system of n degrees of freedom are described by functions in configuration space variables $(q_1, \ldots, q_n)$ or in momentum space variables $(p_1, \ldots, p_n)$ or a complex combinations thereof, such as

$$\eta_k = 2^{-1/2}(q_k - ip_k), \quad \xi_k = 2^{-1/2}(q_k + ip_k).$$

In 1928 Fock [7] introduced the operator solution $\xi_k = \partial/\partial\eta_k$ of the commutation rule $[\xi_k, \eta_k] = 1$ that appears in quantum mechanics. In 1961, V. Bargmann [1–3] investigated in greater detail the function space $\mathfrak{F}$ on which Foch's solution is realized. He also studied the relationship between the space $\mathfrak{F}$, which is called the Bargmann–Segal–Foch space, and the Hilbert space of square integrable functions $L^2(\mathbb{R}^n)$. That relationship was established using an integral transform that is now known as the Bargmann transform. The Bargmann transform resurfaced again in recent years because of its connection with other important transforms such as the Gabor and Zak transforms [16, 17].

The purpose of this chapter is twofold: (1) to give an overview of chromatic derivatives and series in one and several variables and (2) to show that functions in the Bargmann–Segal–Foch space $\mathfrak{F}$ can be represented by chromatic series expansions. We utilize the Bargmann transform and chromatic derivatives to obtain an orthonormal basis for the space $\mathfrak{F}$. A salient feature of this basis is that all the elements of the basis are generated from one single function by applying successive chromatic differentiations to it. This is reminiscent of the wavelets and Gabor systems in which orthonormal bases of certain function spaces are generated from one single function by translation and dilation in the former case and by translation and modulation in the latter.

The notation we use in this chapter is standard. We denote the set of real numbers by $\mathbb{R}$, the integers by $\mathbb{Z}$, and the natural numbers by $\mathbb{N}$.

The chapter is organized as follows. In Sect. 6.2 we give a brief introduction to chromatic derivatives and series. A more general form of chromatic series, which will be used to derive the main results of the articles, is presented in Sect. 6.3. Because chromatic derivatives in higher dimensions are based on orthogonal polynomials in several variables, we will give a brief introduction to the theory of orthogonal polynomials in several variables in Sect. 6.4, followed by an introduction to chromatic derivatives and series in higher dimensions in Sect. 6.5. Sections 6.6 and 6.7 give a brief introduction to the Bargmann–Segal–Foch space and the Bargmann transform and some of their properties that will be used in the sequel. The main result is presented in Sect. 6.8.

6.2 Chromatic Derivatives and Series

In this section we describe the general idea of chromatic expansions associated with a family of orthogonal polynomials.

For the reader's convenience, we will briefly describe how chromatic series are constructed in one dimension and relegate the treatment in higher dimensions to Sect. 6.5. Let $W(\omega)$ be a nonnegative weight function such that all of its moments are finite, that is, such that

$$\mu_n = \int_{-\infty}^{\infty} \omega^n W(\omega) \mathrm{d}\omega < \infty.$$

Let $\{P_n(\omega)\}_{n=0}^{\infty}$ be the family of polynomials orthonormal with respect to $W(\omega)$:

$$\int_{-\infty}^{\infty} P_n(\omega) P_m(\omega) W(\omega) \mathrm{d}\omega = \delta_{\mathrm{m,n}},$$

and let $K^n(f) = P_n(i\frac{d}{dt})(f)$ be the corresponding linear differential operator obtained from $P_n(\omega)$ by replacing ω^k $(0 \le k \le n)$ with $i^k \frac{d^k}{dt^k}$. These differential operators are called *chromatic derivatives* associated with the family of orthogonal polynomials $\{P_n(\omega)\}$ because they preserve the spectral features of bandlimited signals.

Let $\psi(z)$ be the Fourier transform of the weight function $W(\omega)$,

$$\psi(z) = \int_{-\infty}^{\infty} e^{i\omega z} W(\omega) \mathrm{d}\omega.$$

Because $\psi(z)$ will be used in a Taylor-type expansion of functions analytic in a domain around the origin, we shall assume that $\limsup(\mu_n/n!)^{1/n} < \infty$, where, $\psi^{(n)}(0) = i^n \mu_n$. This condition implies that $\psi(z)$ is analytic around the origin. As shown in [10], this condition holds if and only if

$$\int_{-\infty}^{\infty} e^{c|\omega|} W(\omega) \mathrm{d}\omega < \infty$$

for some $c > 0$, and in this case $\psi(z)$ is analytic in the strip $S(c/2) = \{z : \Im(z) < c/2\}$.

The chromatic series expansion of $f \in C^{\infty}(\mathbb{R})$ is given by the following formal series:

$$f(z) \sim \sum_{n=0}^{\infty} K^n(f)(0) K^n(\psi)(z). \tag{6.1}$$

It has been shown in [10] that if $f(z)$ is analytic in the strip $S(c/2)$ and $\sum_{n=0}^{\infty} |K^n(f)(0)|^2$ converges, then for all $u \in \mathbb{R}$, the series (6.1) converges to $f(z)$, uniformly in every strip $\{z : |\Im(z)| < c/2 - \varepsilon\}$, for any $\varepsilon > 0$. Here it

should be emphasized that although chromatic series were originally introduced for bandlimited functions, the theory now applies to a much larger class of functions.

In the particular case, where $W(\omega) = \chi_{(-1,1)}$, the chromatic series associated with the Legendre polynomials converge in the whole complex plane, that is, the strip $S(c/2)$ is $\mathbb{C}$, and the set of entire functions for which $\sum_{n=0}^{\infty} |K^n(f)(0)|^2$ converges is precisely the set of L^2 functions whose Fourier transforms are finitely supported, that is, the set of bandlimited functions. For such functions the chromatic expansions converge uniformly on $\mathbb{R}$, and their truncated series are themselves bandlimited which is analogous to the Whittaker–Shannon sampling series [29]. This is in contrast to Taylor series whose truncated series are not. For this reason chromatic series have more practical applications in signal processing than Taylor series.

6.3 Chromatic Derivatives Associated with More General Differential Operators

In this section we summarize two generalizations of chromatic derivatives and their associated chromatic series.

6.3.1 Chromatic Derivatives Associated with a Sturm–Liouville Differential Operator

Consider the singular Sturm–Liouville boundary-value problem on the half line

$$Ly = -y'' + q(x)y = \lambda y, \quad 0 \le x < \infty, \tag{6.2}$$

$$y(0)\cos\alpha + y'(0)\sin\alpha = 0, \quad -\pi < \alpha \le \pi, \tag{6.3}$$

where $q(x) \in L^1(\mathbb{R}^+)$ is real valued. It is known that the condition $q \in L^1(\mathbb{R}^+)$ implies that the problem is in the limit point case at ∞ and that the spectrum is continuous [24]. In fact, there exists a non-decreasing function $\rho(\lambda)$ such that for all $f \in L^2(\mathbb{R}^+)$

$$\hat{f}(\lambda) = \int_0^{\infty} f(x)\phi(x,\lambda)dx \tag{6.4}$$

exists in the mean and defines a function $\hat{f}(\lambda)$ such that

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\lambda)\phi(x,\lambda)\mathrm{d}\rho(\lambda), \tag{6.5}$$

where $\phi(x,\lambda)$ is a solution of the differential equation (6.2) that satisfies the initial condition

$$\phi(0,\lambda) = \sin\alpha, \ \phi'(0,\lambda) = -\cos\alpha. \tag{6.6}$$

We call the integral transform (6.4) the ϕ-transform of f. Fix $0 < a < \infty$, and let $K^2(a)$ denote the set of all functions with supports in $[0,a]$ that are square integrable with respect to $d\rho$. In most cases of interest $d\rho$ is supported on a half line which, without loss of generality, we may take as $[0,\infty)$. For sufficient conditions for this to hold see [20, p.128]. The main result can be summarized in the following theorem whose proof can be found in [28].

Theorem 1. *Consider the boundary-value problem (6.2) and (6.3), and let $CK^2(a)$ denote the image of $K^2(a)$ under the transformation (6.5), i.e., $f \in CK^2(a)$ if and only if there exists $\hat{f} \in K^2(a)$ such that (6.5) holds. Then there exists a sequence of polynomials $\{p_n(\lambda)\}_{n=0}^{\infty}$ that are orthonormal with respect to $d\rho$ on $[0,a]$, and $p_n(\lambda)$ is of exact degree n. Furthermore, for any $f \in CK^2(a)$, we have for $\alpha \neq 0, \pi$*

$$f(x) = \frac{1}{(\sin\alpha)} \sum_{n=0}^{\infty} [p_n(L)f](0)\psi_n(x), \tag{6.7}$$

where

$$\psi_n(x) = \int_0^{\infty} p_n(\lambda)\phi(x,\lambda)\mathrm{d}\rho(\lambda),$$

and the series converges to f in the mean. Similar expressions exist for $\alpha = 0$ or π. The functions $\{\psi_n(x)\}$ are orthonormal on $[0,\infty)$ and satisfy the initial condition (6.6). The series (6.7) converges to $f(x)$ pointwise for $0 \leq x < \infty$. In fact, the series converges to f uniformly on compact subsets of $(0,\infty)$.

Definition 2. The nth generalized chromatic derivative of a function f associated with the differential operator L at $x = 0$ is defined as

$$K^n[f](0) = \langle \hat{f}, p_n \rangle_{d\rho} = \frac{1}{(\sin\alpha)} [p_n(L)f](0), \text{ for } \alpha \neq 0, \pi$$

where $\hat{f}$ is the ϕ transform of f. Analogous to (6.1), we define the generalized chromatic series expansion of f as

$$\sum_{n=0}^{\infty} K^n[f](0)K^n[\psi](x), \text{ where } \psi_n(x) = K^n[\psi](x),$$

and

$$\psi(x) = \int_0^a \phi(x,\lambda)\mathrm{d}\rho(\lambda).$$

6.3.2 Chromatic Series Associated with More General Integral Transforms

In this section we briefly introduce another generalization of chromatic series that is more intrinsically related to integral transforms other than the Fourier transform. We

begin with an integral transform and assume that the kernel of the transform arises from an initial-value problem associated with a linear differential operator with variable coefficients. It should be noted that the kernels of most classical integral transforms, such as the Fourier, Laplace, Hankel, and Legendre transforms, possess this property.

Consider the integral transform

$$\int_{J_1} \hat{f}(\lambda)\phi(x,\lambda)\mathrm{d}\lambda, \tag{6.8}$$

where J_1 is either $[0,\infty)$ or $(-\infty,\infty)$. Assume that the kernel function $\phi(x,\lambda)$ satisfies the differential equation

$$L\phi(x,\lambda) = \lambda\phi(x,\lambda), \quad x \in I \tag{6.9}$$

on some interval I, where

$$L = q_n(x)\frac{\mathrm{d}^n}{dx^n} + q_{n-1}(x)\frac{\mathrm{d}^{n-1}}{dx^{n-1}} + \ldots + q_1(x)\frac{\mathrm{d}}{dx} + q_0(x)$$

for some continuous functions $q_i(x)$ on I and $q_n(x) \neq 0$ on I.

Let $a \in I$ and assume that $\phi(x,\lambda)$ is continuous in λ and $\phi(a,\lambda) \neq 0$. Without loss of generality, we may take $\phi(a,\lambda) = 1$. If $\phi(a,\lambda) = 0$, we may take ϕ to satisfy $\frac{\partial\phi(x,\lambda)}{\partial x}\Big|_{x=a} = 1$. Such ϕ always exists as one of the fundamental solutions of the initial-value problem [21]. We will focus on the case $J_1 = [0,\infty)$

Let $w(\lambda) > 0$ be a weight function on $0 \leq \lambda \leq b$. Let $\{p_n(\lambda)\}_{n=0}^{\infty}$ be a complete orthonormal system of polynomials in $L^2([0,b),w(\lambda))$ with respect to the weight function $w(\lambda)$.

Consider $f(x)$ the integral transform (6.8) of $\hat{f}$,

$$f(x) = \int_0^b \hat{f}(\lambda)\phi(x,\lambda)\mathrm{d}\lambda, \quad x \in I. \tag{6.10}$$

Applying the operator L to both sides of (6.10), we have

$$Lf(x) = \int_0^b \lambda\hat{f}(\lambda)\phi(x,\lambda)\mathrm{d}\lambda.$$

and hence

$$p_n(L)f(x) = \int_0^b p_n(\lambda)\hat{f}(\lambda)\phi(x,\lambda)\mathrm{d}\lambda.$$

Moreover, we have $[p_n(L)f](a) = \int_0^b p_n(\lambda)\hat{f}(\lambda)\mathrm{d}\lambda = \langle \hat{f}, p_n \rangle$. Thus, formally, we have

$$f(x) = \sum_{n=0}^{\infty} K^n[f](a)\psi_n(x),$$

where the series converges pointwise for $x \in I$,

$$\psi_n(x) = K^n[\psi](x) = \int_0^b p_n(\lambda) w(\lambda) \phi(x,\lambda) \mathrm{d}\lambda, \quad x \in I, \tag{6.11}$$

with

$$\psi(x) = \int_0^b w(\lambda) \phi(x,\lambda) \mathrm{d}\lambda,$$

and $K^n[f](a) = [p_n(L)f](a)$ is the nth chromatic derivative of f associated with the integral transform (6.8) and the system of orthonormal polynomials $\{p_n\}$.

6.4 Orthogonal Polynomials Expansions in Several Variables

6.4.1 Orthogonal Polynomials in Several Variables

Let $\mathbb{N}$ denote the set of nonnegative integers and α be a multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$. We use the notation

$$\alpha! = \alpha_1! \alpha_2! \ldots \alpha_d!, \quad |\alpha| = \alpha_1 + \ldots + \alpha_d,$$

and $\delta_{\alpha,\beta} = \delta_{\alpha_1,\beta_1} \ldots \delta_{\alpha_d,\beta_d}$. For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, we define the monomial $x^\alpha = x_1^{\alpha_1} \ldots x_d^{\alpha_d}$ and $|\alpha|$ is the degree of x^α. A polynomial P in d variables is a linear combination of the form

$$P(x) = \sum_\alpha c_\alpha x^\alpha,$$

where c_α are complex numbers. We denote the set of all polynomials in d variables by Π^d and the set of all polynomials of degree at most n by Π_n^d. The set of all homogenous polynomials of degree n will be denoted by P_n^d

$$P_n^d = \left\{ P : P(x) = \sum_{|\alpha|=n} c_\alpha x^\alpha \right\}.$$

Every polynomial in d variables can be written as a linear combination of homogenous polynomials

$$P(x) = \sum_{k=0}^n \sum_{|\alpha|=k} c_\alpha x^\alpha.$$

It is known that the dimension r_n^d of P_n^d is

$$r_n^d = \binom{n+d-1}{n}.$$

For a fixed d, we may write $r_n = r_n^d$. In one variable, monomials are ordered according to their degrees as $1, x, x^2, \ldots$; however, in several variables, such natural order does not exist. Therefore, we will use the lexicographic order, that is, $\alpha > \beta$, if the first nonzero entry in the difference $\alpha - \beta = (\alpha_1 - \beta_1, \alpha_2 - \beta_2, \ldots, \alpha_d - \beta_d)$ is positive.

If $< , >$ is an inner product on Π^d, we say that a polynomial P is orthogonal to a polynomial Q if $\langle P, Q\rangle = 0$. A polynomial P is called an orthogonal polynomial if P is orthogonal to all polynomials of lower degree, that is, $\langle P, Q\rangle = 0$, for all $Q \in \Pi^d$ with $\deg Q < \deg P$. Denote by V_n^d the space of orthogonal polynomials of degree exactly n:

$$V_n^d = \left\{ P \in \Pi_n^d : \langle P, Q\rangle = 0, \text{ for all } Q \in \Pi_{n-1}^d \right\};$$

the dimension of V_n^d is the same as that of P_n^d.

A multi-sequence $s : \mathbb{N}^d \to I\!R$ is written as $s = (s_\alpha)_{\alpha \in \mathbb{N}^d}$, and for each multi-sequence, we define a linear functional on Π^d by

$$\mathcal{L}(x^\alpha) = s_\alpha, \quad \alpha \in \mathbb{N}^d .$$

Let the elements of the set $\left\{\alpha \in \mathbb{N}^d : |\alpha| = n\right\}$ be arranged as $\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(r_n)}$ according to the lexicographic order. Let $\mathbf{x}^n$ denote the column vector

$$\mathbf{x}^n = (x^\alpha)_{|\alpha|=n} = \left(x^{\alpha(j)}\right)_{j=1}^{r_n};$$

i.e., $\mathbf{x}^n$ is a vector whose elements are the monomials x^α for $|\alpha| = n$, arranged in the lexicographic order.

Define the vector moments $\mathbf{s}_k = \mathcal{L}\left(\mathbf{x}^k\right)$ and

$$\mathbf{s}_{(k)+(j)} = \mathcal{L}\left(\mathbf{x}^k \left(\mathbf{x}^j\right)^T\right),$$

which is a matrix of size $r_k^d \times r_j^d$ whose elements are $\mathcal{L}\left(x^{\alpha+\beta}\right)$ for $|\alpha| = k, |\beta| = j$.

Define the matrix

$$M_{n,d} = \left(s_{(k)+(j)}\right)_{k,j=0}^{n} \text{ and } \Delta_{n,d} = \det M_{n,d} ;$$

$M_{n,d}$ is called a moment matrix and its elements are $\mathcal{L}\left(x^{\alpha+\beta}\right)$ for $|\alpha| \le n, |\beta| \le n$. If $\{P_\alpha\}_{|\alpha|=n}$ is a sequence of polynomials in Π_n^d we get the column polynomial vector $\mathbb{P}_n = (P_{\alpha^{(1)}}, \ldots, P_{\alpha^{(r_n)}})^T$, where $\alpha^{(1)}, \ldots, \alpha^{(r_n)}$ is the lexicographic order in $\left\{\alpha \in \mathbb{N}^d : |\alpha| = n\right\}$. For more details on the subsequent discussion, see [5].

Definition 3. Let $\mathcal{L}$ be a moment functional. A sequence of polynomials $\{P_\alpha\}_{|\alpha|=n}$ in Π_n^d is said to be orthogonal with respect to $\mathcal{L}$ if

$$\mathcal{L}\left(\mathbf{x}^m \mathbb{P}_n^T\right) = 0 \text{ for } n > |m| \text{ and } \mathcal{L}\left(\mathbf{x}^n \mathbb{P}_n^T\right) = s_n,$$

where s_n is an invertible matrix of size $r_n^d \times r_n^d$.

By definition $\mathcal{L}\left(\mathbf{x}^m \mathbb{P}_n^T\right) = 0$ if and only if $\mathcal{L}\left(x^\beta P_\alpha\right) = 0,\ |\beta| = m, |\alpha| = n$. Hence, each P_α is orthogonal to any polynomial of lower degree. It is known that if $\mathcal{L}$ is a moment functional and $\mathbb{P}_n$ is orthogonal as defined before, then $\{\mathbb{P}_0, \mathbb{P}_1, \ldots, \mathbb{P}_n\}$ is a basis for Π_n^d. Hence, there exists matrices $\mathbf{c}_k$ of size $r_n^d \times r_k^d$ such that

$$\mathbf{x}^n = \mathbf{c}_n \mathbb{P}_n + \mathbf{c}_{n-1} \mathbb{P}_{n-1} + \ldots + \mathbf{c}_0 \mathbb{P}_0. \tag{6.12}$$

It is also known that for a given moment functional $\mathcal{L}$, a system of orthogonal polynomials exists if and only if $\Delta_{n,d} \neq 0$.

A moment linear functional $\mathcal{L}$ is said to be positive definite if $\mathcal{L}\left(p^2\right) > 0$ for all $p \in \Pi^d$, $p \neq 0$. The associated sequence $\{s_\alpha\}$ with $\mathcal{L}$ will be also called positive definite.

If $p = \sum a_\alpha x^\alpha$, then $\mathcal{L}(p) = \sum a_\alpha s_\alpha$ and $\mathcal{L}(p^2) = \sum_{\alpha,\beta} a_\alpha a_\beta s_{\alpha+\beta} > 0$ for every sequence $a = (a_\alpha)$ for which $a_\alpha = 0$, except for finitely many multi-indices α. This implies that $\mathcal{L}$ is positive definite if and only if for every tuple $(\beta^{(1)}, \ldots, \beta^{(r)})$, $1 \leq j \leq r$, the matrix $(s_{\beta^{(i)}+\beta^{(j)}})_{i,j=1}^r$ has positive determinant.

It is known [5] that if $\mathcal{L}$ is positive definite, then $\Delta_{n,d} > 0$, and there exists a system of orthogonal polynomials with respect to $\mathcal{L}$. In fact, in this case, there exists an orthonormal basis with respect to $\mathcal{L}$, i.e., there exists a sequence of vector polynomials $\{\mathbb{P}_n\}$ such that

$$\mathcal{L}\left(\mathbb{P}_m \mathbb{P}_n^T\right) = 0 \ \text{ if } \ m \neq n, \ \ \mathcal{L}\left(\mathbb{P}_n \mathbb{P}_n^T\right) = I_{r_n},$$

where I_{r_n} is the identity matrix of size $r_n^d \times r_n^d$.

Let M denote the set of nonnegative Borel measures on $\mathbb{R}^d$ having moments of all orders. Thus, $\mu \in M$ if

$$\int_{\mathbb{R}^d} |x|^\alpha \, \mathrm{d}\mu < \infty \ \text{ for all } \ \alpha \in \mathbb{N}^\mathrm{d},$$

and we call $s_\alpha = \int_{\mathbb{R}^d} x^\alpha d\mu$ the moments of μ. For such a measure $\mu \in M$, we have a moment functional $\mathcal{L}$ defined for polynomials $P \in \Pi^d$ by

$$\mathcal{L}(P) = \int_{\mathbb{R}^d} P(x) \mathrm{d}\mu(\mathrm{x}).$$

If $\mathrm{d}\mu(\mathrm{x}) = \mathrm{W}(\mathrm{x}) dx$ and $W(x)$ is a nonnegative weight function, then $\mathcal{L}$ is positive definite, that is, $\mathcal{L}(P^2) > 0$ for any $0 \neq P \in \Pi^d$. It is known that if $s = (s_\alpha)$ is a multi-sequence, then there exists $\mu \in M$ such that $s_\alpha = \int x^\alpha \mathrm{d}\mu(\mathrm{x})$ if and only if the associated moment functional $\mathcal{L}$ is nonnegative on the set of nonnegative polynomials. That is, $\mathcal{L}(P) \geq 0$ for any $P \in \Pi_+^d = \left\{P \in \Pi^d : P(x) \geq 0\right\}$.

6.4.2 Orthogonal Polynomial Expansions

Let $\mathcal{L}$ be a positive moment functional (hence positive definite) and let $\{\mathbb{P}_n\}$ be the system of orthonormal polynomials associated with it. From the above discussion it follows that there exists a measure $\mu \in M$ such that

$$\mathcal{L}(x^\alpha) = s_\alpha = \int_{\mathbb{R}^d} x^\alpha \mathrm{d}\mu(\mathrm{x}).$$

We assume that μ is absolutely continuous so that $d\mu = W(x)dx$, with W being nonnegative so that $\mathcal{L}_W(f) = \int f(x)W(x)dx$ is positive definite. We also assume that for some $c > 0$,

$$\int_{\mathbb{R}^d} e^{c\|x\|} W(x)\mathrm{d}(\mathrm{x}) < \infty, \tag{6.13}$$

so that polynomials are dense in $L^2(d\mu)$ [5]. Condition (6.13) is satisfied, for example, if μ is compactly supported.

We adopt the following notation: If $P(x_1,\ldots,x_d)$ is a polynomial in $x_1,\ldots,x_d$, then the polynomial $P\left(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_d}\right)$ will be denoted by $P\left(\frac{\partial}{\partial x}\right)$, where x^α is replaced by $\frac{\partial^{|\alpha|}}{\partial x^\alpha}$ where α is a multi-index. More explicitly, if $\alpha = (\alpha_1,\ldots,\alpha_d)$, then

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_1} .. x_d^{\alpha_d}, \quad |\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_d,\ \alpha_i \in \mathbb{N}$$

is replaced by

$$\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_d^{\alpha_d}}.$$

Let $\{P_\alpha^n\}_{|\alpha|=0}^\infty$ denote the sequence of orthonormal polynomials with respect to $\mathcal{L}_W$. Let $L_W^2(\mathbb{R}^d)$ denote the space of all square integrable functions with respect to the weight function W. For any function $f \in L_W^2(\mathbb{R}^d)$, consider its generalized Fourier expansion with respect to $\{P_\alpha^n\}_{|\alpha|=0}^\infty$,

$$f \backsim \sum_{n=0}^\infty \sum_{|\alpha|=n} c_\alpha^n(f) P_\alpha^n, \quad \text{with} \quad c_\alpha^n(f) = < f, P_\alpha^n >_W = \int f(x) P_\alpha^n(x) W(x) dx. \tag{6.14}$$

If we use the vector notation, we have

$$f \backsim \sum_{n=0}^\infty \mathbf{c}_n^T(f)\mathbb{P}_n, \quad \text{with} \quad \mathbf{c}_n(f) = < f, \mathbb{P}_n >_W = \int f(x)\mathbb{P}_n(x) W(x) dx;$$

here $\mathbf{c}_n$ is a column vector with components c_α^n and $|\alpha| = n$. It follows that (6.14) can be written as

$$f \backsim \sum_{n=0}^\infty \mathrm{Proj}_{V_n^d} f, \quad \text{where} \quad \mathrm{Proj}_{V_n^d} f = \mathbf{c}_n^T(f)\mathbb{P}_n,$$

which leads to

$$\text{Proj}_{V_n^d} f(x) = \int f(y)\mathbf{P}_n(x,y)W(x)dx, \quad \text{with} \quad \mathbf{P}_n(x,y) = \mathbb{P}_n^T(x)\mathbb{P}_n(y).$$

6.5 Multidimensional Chromatic Derivatives

Let $z = (z_1, z_2, \ldots, z_d) \in \mathbb{C}^d$ and define the inner product

$$< z, w >= \sum_{k=1}^{d} z_k \overline{w}_k, z, w \in \mathbb{C}^d \text{ so that } \|z\|^2 = \sum_{k=1}^{d} |z_k|^2$$

If $z_k = x_k + iy_k$, define $\Re z = (x_1, x_2, \ldots, x_d)$ and $\Im z = (y_1, y_2, \ldots, y_d)$ so that $z = \Re z + i\Im z$. Moreover,

$$\|\Re z\|^2 = \sum_{k=1}^{d} x_k^2, \quad \|\Im z\|^2 = \sum_{k=1}^{d} y_k^2.$$

For every real $a > 0$, we let $S_d(a) = \{z \in \mathbb{C}^d \ : \ \|\Im(z)\| < a\}$.

Definition 4. Let $f : \mathbb{C}^d \to \mathbb{C}$; the n^{th} chromatic derivatives $\mathbb{K}^n(f)$ of $f(z)$ with respect to the polynomials $\{P_\alpha^n\}$ is defined as

$$\mathbb{K}^n(f) = \mathbb{P}_n\left(-i\frac{\partial}{\partial z}\right)(f),$$

where $\mathbb{P}_n$ is the column vector defined before.

Hence, the n^{th} chromatic derivative of f is a column vector with r_n^d components, with each component being a linear combination of partial derivatives. It is easy to check that if $\omega \in I\!R^d$ is fixed, then[1]

$$K_\alpha^n(e^{i\langle\omega,z\rangle}) = P_\alpha^n(\omega)e^{i\langle\omega,z\rangle}. \tag{6.15}$$

For more details on the results of this section, see [9].

Proposition 5. *Let $\varphi(\omega) \in L_W^2(I\!R^d)$ and define a corresponding function $f_\varphi : S_d(c/2) \to \mathbb{C}$ by a $W(\omega)$-weighted Fourier transform of φ:*

$$f_\varphi(z) = \int_{I\!R^d} \varphi(\omega)e^{i\langle\omega,z\rangle} W(\omega)\mathrm{d}\omega. \tag{6.16}$$

[1] If $x, \omega \in I\!R^d$ are both real vectors, then we denote their scalar product by $x.\omega$. If at least one of u, z is complex, we denote their scalar product by $\langle u, z\rangle$.

Then $f_\varphi(z)$ is analytic on $S_d(c/2)$, and for all n and $z \in S_d(c/2)$,

$$\mathbb{K}^n[f_\varphi](z) = \int_{I\!R^d} \mathbb{P}_n(\omega)\, \varphi(\omega)\, e^{i\langle \omega, z\rangle}\, W(\omega) \mathrm{d}\omega. \tag{6.17}$$

Proof. Let $z \in \mathbb{C}^d$ and $z = x + iy$, where $x, y \in I\!R^d$, with $\|y\| < c/2 - \varepsilon$ for some $\varepsilon > 0$; then for all $\alpha \in \mathbb{N}^d$,

$$\begin{aligned}
&\int_{I\!R^d} |(i\omega)^\alpha \varphi(\omega) e^{i\langle \omega, z\rangle}| W(\omega) \mathrm{d}\omega \\
&\leq \left(\int_{I\!R^d} |\varphi(\omega)|^2 W(\omega) \mathrm{d}\omega \int_{\mathbb{R}^d} \omega^{2\alpha} \mathrm{e}^{2(\omega.y)} \mathrm{W}(\omega) \mathrm{d}\omega \right)^{1/2} \\
&\leq \|\varphi\|_W \left(\int_{I\!R^d} \omega^{2\alpha} e^{2\|\omega\|\|y\|} W(\omega) \mathrm{d}\omega \right)^{1/2}
\end{aligned}$$

Choose $M_\alpha > 0$ such that for all ω,

$$\omega^{2\alpha} = \omega_1^{2\alpha_1} \dots \omega_k^{2\alpha_k} \leq \|\omega\|^{2\alpha_1} \dots \|\omega\|^{2\alpha_k} = \|\omega\|^{2|\alpha|} < M_\alpha e^{\varepsilon\|\omega\|},$$

then

$$\int_{I\!R^d} |(i\omega)^\alpha \varphi(\omega) e^{i\langle \omega, z\rangle}| W(\omega) d\omega \leq M_\alpha^{1/2} \|\varphi\|_W \left(\int_{I\!R^d} e^{\|\omega\|(c-\varepsilon)} W(\omega) d\omega \right)^{1/2} < \infty.$$

The claims now follow from (6.12), (6.13), and (6.15).

It is known that if condition (6.13) holds, then $\{\mathbb{P}_n(\omega)\}$ is a complete system in $L^2_W(I\!R^d)$.

Proposition 6. *Let $\varphi(\omega) \in L^2_W(I\!R^d)$; if for some fixed $u \in S_d(c/2)$, the function $\varphi(\omega) e^{i\langle \omega, u\rangle}$ belongs to $L^2_W(I\!R^d)$, then in $L^2_W(I\!R^d)$, we have*

$$\varphi(\omega) e^{i\langle \omega, u\rangle} = \sum_{n=0}^{\infty} [\mathbb{K}^n[f_\varphi](u)]^T\, \mathbb{P}_n(\omega), \tag{6.18}$$

and for f_φ given by (6.16), we have

$$\sum_{n=0}^{\infty} \sum_{|\alpha|=n} |K_\alpha^n[f_\varphi](u)|^2 = \|\varphi(\omega) e^{i\langle \omega, u\rangle}\|_W^2 < \infty. \tag{6.19}$$

Proof. Let f_φ be given by (6.16). By Proposition 5, (6.17) holds. However, if $\varphi(\omega) e^{i\langle \omega, u\rangle}$ belongs to the space $L^2_W(I\!R^n)$, then (6.17) asserts that the projection of $\varphi(\omega) e^{i\langle \omega, u\rangle}$ onto the vector $\mathbb{P}_n(\omega)$ is equal to $\mathbb{K}^n[f_\varphi](u)$:

$$\langle \varphi(\omega) e^{i\langle \omega, u\rangle}, \mathbb{P}_n(\omega)\rangle_W = \mathbb{K}^n[f_\varphi](u). \tag{6.20}$$

Since $\{\mathbb{P}_n(\omega)\}_{n\in\mathbb{N}}$ is a complete orthonormal system in $L^2_W(\mathbb{R}^d)$, (6.20) implies (6.18), and Parseval's theorem implies (6.19).

As a corollary, we have

Corollary 7. *For every $\varphi(\omega) \in L^2_W(\mathbb{R}^d)$ and every $u \in \mathbb{R}^d$, equality (6.18) holds and*

$$\sum_{n=0}^{\infty} \sum_{|\alpha|=n} |K^n_\alpha[f_\varphi](u)|^2 = \|\varphi(\omega)\|^2_W. \tag{6.21}$$

Thus, the sum $\sum_{n=0}^{\infty}\sum_{|\alpha|=n} |K^n_\alpha[f_\varphi](u)|^2$ is independent of $u \in \mathbb{R}^d$.

Proof. If $u \in \mathbb{R}^d$, then $\varphi(\omega)e^{i\langle\omega,u\rangle} \in L^2_W(\mathbb{R}^d)$ and $\|\varphi(\omega)e^{i\langle\omega,u\rangle}\|^2_W = \|\varphi(\omega)\|^2_W$; thus, Proposition 6 applies.

Definition 8. Let $z \in S_d(c/2)$ and define

$$\psi(z) = \int_{\mathbb{R}^d} e^{i\langle\omega,z\rangle} W(\omega)\mathrm{d}\omega, \tag{6.22}$$

and more generally

$$K^n_\alpha(\psi(z)) = \psi^n_\alpha(z) = \int_{\mathbb{R}^d} P^n_\alpha(\omega) e^{i\langle\omega,z\rangle} W(\omega)\mathrm{d}\omega; \tag{6.23}$$

we may also use the vector notation $\boldsymbol{\Psi}_n(z) = (\psi^n_\alpha(z))_{|\alpha|=n}$.

Note that Proposition 5 implies that the integrals in the definitions of $\psi(z)$ and $\psi^n_\alpha(z)$ are finite.

Corollary 9. *Let $\varepsilon > 0$, then for all $z \in S(\frac{c}{2} - \varepsilon)$*

$$\sum_{n=0}^{\infty} \sum_{|\alpha|=n} |\psi^n_\alpha(z)|^2 < \left\| e^{(\frac{c}{2}-\varepsilon)\|\omega\|} \right\|^2_W < \infty. \tag{6.24}$$

Proof. We apply Proposition 6 with $\varphi(\omega) = 1$, in which case $f_\varphi(z) = \psi(z)$, and, using (6.19), obtain

$$\sum_{n=0}^{\infty} \sum_{|\alpha|=n} |\psi^n_\alpha(z)|^2 = \left\| e^{i\langle\omega,z\rangle} \right\|^2_W = \left\| e^{|\omega.\Im(z)|} \right\|^2_W < \left\| e^{(\frac{c}{2}-\varepsilon)\|\omega\|} \right\|^2_W < \infty.$$

The last inequality follows from (6.13).

Definition 10. We denote by Λ^2_W the vector space of functions $f : S_d(c/2) \to \mathbb{C}$ which are analytic on $S_d(c/2)$ and satisfy $\sum_{n=0}^{\infty}\sum_{|\alpha|=n} |K^n_\alpha[f_\varphi](0)|^2 < \infty$. The chromatic series expansion of a function $f \in \Lambda^2_W$ is given by the following formal series:

$$f(z) \sim \sum_{n=0}^{\infty} [\mathbb{K}^n(f)(u)]^T \boldsymbol{\Psi}_n(z-u). \tag{6.25}$$

Proposition 11. *The mapping*

$$f(z) \mapsto \varphi_f(\omega) = \sum_{n=0}^{\infty} [\mathbb{K}^n[f](0)]^T \, \mathbb{P}_n(\omega) \tag{6.26}$$

is an isomorphism between the vector spaces Λ_W^2 *and* $L_W^2(\mathbb{R}^d)$*, and its inverse is given by* (6.16)*.*

Definition 12. For $f \in \Lambda_W^2$, we denote the corresponding $\varphi_f(\omega)$ given by (6.26) by $\mathcal{F}_W[f](\omega)$; thus, for $z \in S_d(c/2)$,

$$f(z) = \int_{\mathbb{R}^d} \mathcal{F}_W[f](\omega) e^{i\langle \omega, z\rangle} W(\omega) \mathrm{d}\omega. \tag{6.27}$$

Proposition 11 and Corollary 7 imply the following corollary.

Corollary 14. *For all* $f \in \Lambda_W^2$ *and all* $x \in \mathbb{R}^d$*, the sum* $\sum_{n=0}^{\infty} \sum_{|\alpha|=n} |K_\alpha^n[f_\varphi](x)|^2$ *converges and is independent of x. Moreover,*

$$f(z) = \sum_{n=0}^{\infty} [\mathbb{K}^n(f)(u)]^T \, \Psi_n(z-u). \tag{6.28}$$

Example. The most efficient system of orthogonal polynomials to use to generate chromatic derivatives will depend on the shape of the support of the Fourier transform $\mathcal{F}_W[f]$ of f. Below we give some examples:

1. If the support of $\mathcal{F}_W[f]$ is the half space $\{x \in \mathbb{R}^d : x_1, \geq 0, \ldots, x_d \geq 0\}$, we use the normalized Laguerre polynomials

$$L_\alpha(x) = L_{\alpha_1}^{k_1}(x) \cdots L_{\alpha_d}^{k_d}(x), \quad |\alpha| = n, x \in \mathbb{R}_+^d, k_i \geq -1,$$

   with weight function $W(x) = x^k e^{-|x|_1}$, where $|x|_1 = x_1 + \cdots + x_d$, $k = (k_1, \ldots, k_d)$ and $L_n^\alpha(x)$ is the normalized Laguerre polynomial of degree n and parameter α.
2. If the support of $\mathcal{F}_W[f]$ is the hypercube $D = [-1,1]^d$, we may use the d-dimensional Jacobi polynomials

$$P_\alpha^{(a,b)}(x) = P_{\alpha_1}^{(a_1,b_1)}(x_1) \cdots P_{\alpha_d}^{(a_d,b_d)}(x_d), \quad |\alpha| = n,$$

   with weight function

$$W^{(a,b)}(x) = \prod_{i=1}^{d} (1-x_i)^{a_i} (1+x_i)^{b_i},$$

   where $a = (a_1, a_2, \ldots, a_d), b = (b_1, b_2, \ldots, b_d)$, and $P_n^{(\alpha,\beta)}(x)$ is the normalized Jacobi polynomial of degree n with parameters α, β. See [23] for the definition of the Laguerre and Jacobi polynomials.

6.6 The Bargmann–Segal–Foch Space

In this section, we introduce the Bargmann–Segal–Foch space and some of its properties that will be used later.

Definition 14. Let $d\mu_n(z) = \rho_n d^n x d^n y$, $\rho_n = (\pi)^{-n}\exp(-\|z\|^2)$. The Bargmann–Segal–Foch space $\mathfrak{F}$ consists of all entire functions $F(z)$ in $\mathbb{C}^n$ such that

$$\|F\|_{\mathfrak{F}}^2 = \int_{\mathbb{C}^n} |F(z)|^2 \mathrm{d}\mu_{\mathrm{n}}(\mathrm{z}) < \infty.$$

It is a Hilbert space with inner product defined by

$$\langle F, G\rangle_{\mathfrak{F}} = \int_{\mathbb{C}^n} F(z)\overline{G}(z)\mathrm{d}\mu_{\mathrm{n}}(\mathrm{z}) < \infty$$

and hence with norm

$$\|F\|_{\mathfrak{F}}^2 = \int |F(z)|^2 \mathrm{d}\mu(\mathrm{z}).$$

The inner product can also be defined in terms of the Taylor series coefficients of F and G. For, if $k = (k_1, \ldots, k_n), m = (m_1, \ldots, m_n)$ are multi-indices with $k_i, m_i \in \mathbb{N}$ and

$$F(z) = \sum_{|k|=0} a_k z^k \quad \text{and} \quad G(z) = \sum_{|m|=0}^{\infty} b_m z^m, \tag{6.29}$$

$z^k = z_1^{k_1} z_2^{k_2} \cdots z_n^{k_n}$ and $z_k = r_k e^{i\theta_k}$, $k = 1, 2, \ldots, n$, then

$$\langle F, G\rangle_{\mathfrak{F}} = \pi^{-n} \sum_{k,m} a_k \overline{b}_m B_{k,m},$$

where

$$B_{k,m} = \prod_{j=1}^{n} \left\{ \left(\int_0^{2\pi} e^{i(k_j - m_j)\theta_j} \mathrm{d}\theta_{\mathrm{j}} \right) \left(\int_0^{\infty} r_j^{k_j + m_j + 1} e^{-r_j^2} \mathrm{d}r_{\mathrm{j}} \right) \right\},$$

and since

$$\int_0^{\infty} r^{2m+1} e^{-r^2} \mathrm{d}r = \frac{m!}{2},$$

it is easily seen that

$$B_{k,m} = \begin{cases} 0 & \text{if } k \neq m \\ \pi^n m! & \text{if } k = m \end{cases}.$$

Hence,

$$\langle F, G\rangle_{\mathfrak{F}} = \sum_{|m|=0}^{\infty} a_m \overline{b}_m m!\,, \tag{6.30}$$

and

$$\|F\|_{\mathfrak{F}}^2 = \sum |a_m|^2 m!. \tag{6.31}$$

Interchanging the integration and the summation signs is possible because of the uniform convergence.

Lemma 15. *The set* $\left\{u_m(z) = z^m/\sqrt{m!}\right\}_{|m|=0}^{\infty}$ *is an orthonormal basis of* $\mathfrak{F}$.

Proof. The set $\left\{u_m(z) = z^m/\sqrt{m!}\right\}_{|m|=0}^{\infty}$ is an orthonormal family, i.e.,

$$\langle u_k, u_m\rangle_{\mathfrak{F}} = \delta_{k,m}.$$

This follows from (6.30).

Moreover, it is complete because in view of (6.30), if $F \in \mathfrak{F}$, then $\sqrt{m!}\, a_m = \langle F, u_m\rangle_{\mathfrak{F}}$ for all multi-indices m. Hence, if $\langle F, u_m\rangle_{\mathfrak{F}} = 0$ for all m, then $F = 0$, and it follows that $\{u_m(z)\}_{|m|=0}^{\infty}$ is complete, and consequently it is an orthonormal basis of $\mathfrak{F}$.

By applying the Cauchy–Schwarz inequality to the Taylor series of $F = \sum_m a_m z^m$, we obtain

$$|F(z)|^2 \le \left(\sum m!|a_m|^2\right)\left(\sum \frac{|z|^{2m}}{m!}\right) = e^{|z|^2}\|F\|_{\mathfrak{F}}^2\,, \tag{6.32}$$

which shows that convergence in $\mathfrak{F}$ implies pointwise convergence and uniform convergence on compact sets. Another consequence of (6.32) is that the evaluation map $F \to F(a)$ is continuous; hence $\mathfrak{F}$ is a reproducing kernel Hilbert space. Since $\{u_m(z)\}_{|m|=0}^{\infty}$ is an orthonormal basis of $\mathfrak{F}$, the reproducing kernel can be found explicitly. In fact, the reproducing kernel is readily seen to be

$$K(z, w) = \sum_m \frac{z^m \overline{w}^m}{m!} = e^{\langle z, w\rangle}\,.$$

Thus, for any $F \in \mathfrak{F}$, we have by (6.30)

$$\langle F(z), K(z, w)\rangle_{\mathfrak{F}} = \int F(z) e^{\langle w, z\rangle}\, \mathrm{d}\mu(z) = \sum_{n=0}^{\infty} n! a_n \frac{w^n}{n!} = F(w). \tag{6.33}$$

6.7 The Bargmann Transform

The Bargmann transform was introduced in [3] to establish the mathematical foundation for some of Foch's work on quantum field theory. It has also appeared in the area of quantum optics [18, 19]. Recently, it has appeared in some important

applications, one of which is its use in the solution of the Gabor representation problem and in the proof of the completeness of the canonical coherent states in quantum mechanics and the Gabor frames in signal analysis [1, 16, 17].

Definition 16. The Bargmann transform $\mathcal{A}[f]$ of a function $f : \mathbb{R}^n \to \mathbb{C}$ is defined by

$$\mathcal{A}[f](z) = F(z) = (\pi)^{-n/4} \int f(q) \exp\left\{-(\|z\|^2 + \|q\|^2)/2 + \sqrt{2}\, z \cdot q\right\} \mathrm{d}^n q \quad (6.34)$$

which can be written as

$$F(z) = \langle k(z,q), \overline{f}(q) \rangle_{L^2(\mathbb{R}^n)}, \qquad \text{whenever } f \in L^2(\mathbb{R}^n),$$

where $z = x + iy$, and

$$k(z,q) = (\pi)^{-n/4} \exp\left\{-(\|z\|^2 + \|q\|^2)/2 + \sqrt{2}\, z \cdot q\right\}, \quad (6.35)$$

$q \in \mathbb{R}^n$ real and $z \in \mathbb{C}^n$ complex.

If f is a locally integrable function such that $f(q) = O(\exp(a\|q\|^2))$ for sufficiently large $\|q\|$, and some $a < 1/2$, in particular, if $f \in L^2(\mathbb{R}^n)$, then its Bargmann transform exists almost everywhere.

Let $\tilde{H}_m(x)$ be the Hermite polynomial of degree m defined by

$$\tilde{H}_m(x) = 2x\tilde{H}_{m-1}(x) - 2(m-1)\tilde{H}_{m-2}(x), \quad \tilde{H}_0(x) = 1, \quad \tilde{H}_1(x) = 2x.$$

We define the normalized Hermite polynomials by

$$H_m(x) = \frac{1}{\sqrt[4]{\pi} 2^{m/2} \sqrt{m!}} \tilde{H}_m(x)$$

so that

$$\int_{\mathbb{R}} H_k(x) H_m(x) e^{-x^2} dx = \delta_{k,m}$$

or

$$\int_{\mathbb{R}} h_k(x) h_m(x) dx = \delta_{k,m},$$

where $h_k(x) = H_k(x) e^{-x^2/2}$ are the normalized Hermite functions, which are an orthonormal basis of $L^2(\mathbb{R})$.

Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a multi-index with α_i being a nonnegative integer and $|\alpha| = \alpha_1 + \cdots + \alpha_n$. We define the n-dimensional normalized Hermite polynomial of degree m by

$$H_\alpha(x) = H_{\alpha_1}(x_1) \cdots H_{\alpha_n}(x_n), \quad |\alpha| = m,\ x = (x_1, \ldots, x_n),$$

with weight function $W(x) = e^{-\|x\|^2}$, $x \in \mathbb{R}^n$. The n-dimensional normalized Hermite functions are defined by

$$h_\alpha(x) = H_\alpha(x) e^{-\|x\|^2/2}$$

so that

$$\int_{\mathbb{R}^n} h_\alpha(x) h_\beta(x) d^n x = \delta_{\alpha,\beta}.$$

It then follows from the relation [8, P. 837, Formula 7.374-6]

$$\int_{\mathbb{R}} e^{-(x-y)^2} \tilde{H}_n(x) dx = \sqrt{\pi}\, 2^n y^n\,, \quad n = 0, 1, \ldots \tag{6.36}$$

that the Bargmann transform of the normalized Hermite function $h_\alpha(q)$ is $u_\alpha(z) = z^\alpha/\sqrt{\alpha!}\,, |\alpha| = 0, 1, 2, \ldots$.

Since the Hermite functions $\{h_\alpha(q)\}$ are an orthonormal basis of $L^2(\mathbb{R}^n)$, we have for any $f \in L^2(\mathbb{R}^n)$,

$$f(q) = \sum_\alpha \langle f, h_\alpha \rangle h_\alpha(q) \quad \text{with } \sum_\alpha |\langle f, h_\alpha \rangle|^2 < \infty.$$

Therefore, the Bargmann transform, $F(z) = \mathcal{A}[f](z)$, of f is given by

$$F(z) = \sum_\alpha \langle f, h_\alpha \rangle z^\alpha/\sqrt{\alpha!}$$

which is in $\mathfrak{F}$ since by (6.31)

$$\|F\|_{\mathfrak{F}} = \sum_\alpha |\langle f, h_\alpha \rangle|^2 < \infty.$$

There are several ways to prove Parseval's relation for the Bargmann transform. The one we shall use is based on formula (6.30). Let F and G be the Bargmann transforms of $f, g \in L^2(\mathbb{R}^n)$, respectively. If we denote $\langle f, h_\alpha \rangle$ and $\langle g, h_\alpha \rangle$ by $\hat{f}_\alpha$ and $\hat{g}_\alpha$, then

$$\begin{aligned} \langle f\,, g \rangle_{L^2(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} \left(\sum_{|\alpha|=0}^{\infty} \hat{f}_\alpha h_\alpha(q) \right) \left(\sum_{|\beta|=0}^{\infty} \overline{\hat{g}}_\beta \overline{h}_\beta(q) \right) \mathrm{d^n q} \\ &= \sum_{|\alpha|=0}^{\infty} \hat{f}_\alpha \overline{\hat{g}}_\alpha\,. \end{aligned}$$

On the other hand, by Lemma 15, one can show

$$\langle F\,, G \rangle_{\mathfrak{F}} = \sum_{|\alpha|=0}^{\infty} \hat{f}_n \overline{\hat{g}}_n. \tag{6.37}$$

Therefore, we have the following Parseval's relation for the Bargmann transform

$$\langle F\,,\,G\rangle_{\mathfrak{F}} = \langle f\,,\,g\rangle_{L^2(\mathbb{R}^n)}. \tag{6.38}$$

Theorem 17. *The Bargmann transformation is a unitary transformation from* $L^2(\mathbb{R}^n)$ *onto* $\mathfrak{F}$ *that maps the normalized Hermite functions* $h_\alpha(x)$ *into* $u_\alpha(z) = z^\alpha/\sqrt{\alpha!}\,, |\alpha| = 0,1,2,\ldots$.

6.8 Chromatic Expansions in the Bargmann–Segal–Foch Space

In this section we show that functions in the Bargmann–Segal–Foch space can be expanded in chromatic series. To this end we use the Bargmann transform and chromatic derivatives to show that there exists an entire function ψ in the space $\mathfrak{F}$ whose chromatic derivatives $\mathbb{K}^m[\psi](z)$ are an orthogonal basis for $\mathfrak{F}$, i.e., the basis is generated from one single function by applying successively chromatic differentiations to it.

Let

$$L_i = \frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial z_i} + z_i\right).$$

It is easy to see that

$$L_i F(z) = \int_{\mathbb{R}^n} f(q) q_i k(z,q) \mathrm{d}^{\mathrm{n}} q,$$

hence

$$H_\alpha(L) F(z) = \int_{\mathbb{R}^n} f(q) H_\alpha(q) k(z,q) \mathrm{d}^{\mathrm{n}} q,$$

where $H_\alpha(L) = H_{\alpha_1}(L_1)\cdots H_{\alpha_n}(L_n)$.

Following the results of Sect. 6.3, we have the following definition.

Definition 18. We define the αth chromatic derivative of $F(z)$ with respect to the operator L and the Hermite polynomials as

$$K^\alpha F(z) = H_\alpha(L) F(z).$$

Let $\mathcal{S}(\mathbb{R}^n)$ be the Schwartz space of rapidly decreasing functions consisting of all $\phi \in C^\infty(\mathbb{R}^n)$ such that

$$\gamma_{l,m}(\phi) = \sup_{q\in\mathbb{R}^n, |\beta|\le l, |\alpha|\le m} \left| q^\beta \frac{\partial^{|\alpha|}\phi(q)}{\partial q_1^{\alpha_1}\cdots q_n^{\alpha_n}} \right| < \infty,$$

where $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $|\beta| = \beta_1 + \cdots + \beta_n$, and $\alpha_i, \beta_i \in \mathbb{N}$.

We now state the following theorem whose proof will be published somewhere else.

Theorem 19. *There exists a function $\phi \in \mathcal{S}(\mathbb{R}^n)$ whose Bergmann transform $\mathcal{A}(\phi) = \psi(z) \in \mathfrak{F}$ has the property that its chromatic derivatives $\{K^\alpha \psi(z)\}$ are an orthogonal basis of $\mathfrak{F}$. Hence, any $F \in \mathfrak{F}$ can be written in the form*

$$F(z) = \sum_\alpha K^\alpha F(0) K^\alpha \psi(z). \tag{6.39}$$

References

1. Bargmann V, Butera P, Girardello L, Klauder JR (1971) On the completeness of the coherent states. Rep Math Phys 2:221–228
2. Bargmann V (1967) On a Hilbert space of analytic functions and an associated integral transform, Part II, A family of related function spaces application to distribution theory. Comm Pure Appl Math 20:1–101
3. Bargmann V (1961) On a Hilbert space of analytic functions and an associated integral transform, Part I. Comm Pure Appl Math 14:187–214
4. Cushman M, Herron T (2001) The general theory of chromatic derivatives. Kromos technology technical report, Los Altos
5. Dunkl C, Xu Y (2001) Orthogonal polynomials of several variables. Encyclopedia of mathematics, Cambridge University Press, Cambridge
6. Herron T, Byrnes J (2001) Families of orthogonal differential operators for signal processing. Kromos technology technical report, Los Altos
7. Foch V (1928) Verallgemeinerung und Lösung der Diracschen statistischen Gleichung. Z Phys 49:339–357
8. Gradshteyn I, Ryzhik I (1965) Tables of integrals, series, and products. Academic, New York
9. Ignjatovic A, Zayed A (2011) Multidimensional chromatic derivatives and series expansions. Proc Amer Math Soc 139(10):3513–3525
10. Ignjatovic A (2009) Chromatic derivatives, chromatic expansions and associated function spaces. East J Approx 15(2):263–302
11. Ignjatovic A (2007) Local approximations based on orthogonal differential operators. J Fourier Anal & Applns 13:309–330
12. Ignjatovic A (2001) Numerical differentiation and signal processing. Proc International Conference Information, Communications and Signal Processing (ICICS), Singapore
13. Ignjatovic A (2001) Local approximations and signal processing. Kromos technology technical report, Los Altos
14. Ignjatovic A (2001) Numerical differentiation and signal processing. Kromos technology technical report, Los Altos
15. Ignjatovic A, Carlin N (1999) Signal processing with local behavior, Provisional Patent Application, (60)-143,074 , US Patent Office, Patent issued as US 6313778, June 2001
16. Janssen AJ (1982) Bargmann transform, zak transform, and coherent states. J Math Phys 23:720–731
17. Janssen AJEM 1981) Gabor representation of generalized functions. J Math Anal Appl 83: 377–394
18. Klauder J, Skagerstam BS (1985) Coherent states. World scientific, Singapore
19. Klauder J, Sudarshan EC (1968) Fundamentals of quantum optics. W. A. Benjamin, New York
20. Marchenko V (1986) Sturm-Liouville operators and their applications. Birkhauser, New York
21. Naimark MA (1967) Linear differential operators I. George Harrap, London

22. Narasimha M, Ignjatovic A, Vaidyanathan P (2002) Chromatic derivative filter banks. IEEE Signal Processing Lett 9(7):215–216
23. Szegö G (1975) Orthogonal polynomials. Amer Math Soc, Providence, RI
24. Titchmarsh E (1962) Eigenfunction expansion I. Oxford University Press, London
25. Vaidyanathan P, Ignjatovic A, Narasimha M (2001) New sampling expansions of bandlimited signals based on chromatic derivatives. Proc 35th Asilomar Conf Signals, Systems and Computers, Monterey, pp 558–562
26. Walter G, Shen X (2005) A sampling expansion for non-bandlimited signals in chromatic derivatives. IEEE Trans Signal Process 53: 1291–1298
27. Zayed AI (2011) Chromatic expansions of generalized functions. Integral Transforms and Special Functions 22(4–5):383–390
28. Zayed AI (2010) Generalizations of chromatic derivatives and series expansions. IEEE Trans Signal Process 58(3):1638–1647
29. Zayed AI (1993) Advances in Shannon's sampling theory. CRC Press, Boca Raton

Chapter 7
Representation Formulas for Hardy Space Functions Through the Cuntz Relations and New Interpolation Problems

Daniel Alpay, Palle Jorgensen, Izchak Lewkowicz, and Itzik Marziano

Abstract We introduce connections between the Cuntz relations and the Hardy space $\mathbf{H}_2$ of the open unit disk $\mathbb{D}$. We then use them to solve a new kind of multipoint interpolation problem in $\mathbf{H}_2$, where, for instance, only a linear combination of the values of a function at given points is preassigned, rather than the values at the points themselves.

7.1 Introduction

One motivation for studying representations of the Cuntz relations comes from signal processing, subband filters, and their applications to wavelets. This falls within a larger context of multiscale problems, see, for example, [17]. In this work we study the Cuntz relations in a different context and introduce connections between them and the Hardy space $\mathbf{H}_2$ of the open unit disk $\mathbb{D}$. We prove in particular the following results: Let b be a finite Blaschke product of degree M, and let $e_1, \ldots, e_M$ be an orthonormal basis of $\mathbf{H}_2 \ominus b\mathbf{H}_2$. A function f belongs to $\mathbf{H}_2$ and has norm less or equal to 1 if and only if it can be written as

D. Alpay (✉) • I. Marziano
Department of Mathematics, Ben Gurion University of the Negev, Be'er Sheva 84105, Israel
e-mail: dany@math.bgu.ac.il; martzian@bgu.ac.il

P. Jorgensen
Department of Mathematics, The University of Iowa, 14 MLH, Iowa City, IA 52242-1419, USA
e-mail: jorgen@math.uiowa.edu

I. Lewkowicz
Department of Electrical Engineering, Ben Gurion University of the Negev,
Be'er Sheva 84105, Israel
e-mail: izchak@ee.bgu.ac.il

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,

DOI 10.1007/978-1-4614-4145-8_7,

$$f(z) = \sum_{j=1}^{M} e_j(z) f_j(b(z)), \tag{7.1}$$

where $f_1, \ldots, f_M \in \mathbf{H}_2$, are uniquely defined and are such that

$$\|f\|_{\mathbf{H}_2}^2 = \sum_{j=1}^{M} \|f_j\|_{\mathbf{H}_2}^2. \tag{7.2}$$

From now on, we denote by $\|f\|_2$ the norm of an element of $\mathbf{H}_2$. Using Leech's factorization theorem (see Sect. 7.2), we prove that, equivalently, f belongs to $\mathbf{H}_2$ and has norm less or equal to 1 if and only if it can be written as

$$f(z) = \frac{\sum_{j=1}^{M} e_j(z)\sigma_{1j}(b(z))}{1 - b(z)\sigma_2(b(z))}, \tag{7.3}$$

where

$$\sigma = \begin{pmatrix} \sigma_{11} \\ \vdots \\ \sigma_{1M} \\ \sigma_2 \end{pmatrix} \tag{7.4}$$

is a Schur function, that is, analytic and contractive in $\mathbb{D}$.

Representation (7.1) allows us to solve various interpolation problems in $\mathbf{H}_2$ by translating them into tangential interpolation problems at one point (in fact at the origin) in $\mathbf{H}_2^M$. The solution of this latter problem or, more generally, of the bitangential interpolation problem in $\mathbf{H}_2^{p\times q}$ is well known. See, for instance, [1, 2].

Similarly, the representation (7.3) allows us to solve various interpolation problems in $\mathbf{H}_2$ by translating them into tangential interpolation problems at *one* point (here too, in fact at the origin), for $\mathbb{C}^{M+1}$-valued Schur functions, whose solution is well known. See, for instance, [13] for the general bitangential interpolation problem for matrix-valued Schur functions.

We now illustrate these points. First note that, for b the Blaschke product with zeroes the points $a_1, \ldots, a_M$, (7.1) leads to

$$f(a_\ell) = \sum_{j=1}^{M} e_j(a_\ell) f_j(0), \quad \ell = 1, \ldots, M. \tag{7.5}$$

For preassigned values of $f(a_\ell)$, $\ell = 1, \ldots, M$, this reduces the Nevanlinna–Pick interpolation problem for M points in $\mathbf{H}_2$ to a tangential interpolation problem at the origin for functions in $\mathbf{H}_2^M$, whose solution, as already mentioned, is well known. The novelty in this chapter is by exploiting the above reduction scheme to solve multipoint interpolation problems in $\mathbf{H}_2$. For example, consider the following problem:

Problem. Given M points $a_1,\dots,a_M$ in $\mathbb{D}$ and $u=\begin{pmatrix}u_1 & u_2 & \cdots & u_M\end{pmatrix}\in\mathbb{C}^{1\times M}$ and $\gamma\in\mathbb{C}$, find all $f\in\mathbf{H}_2$ such that

$$\sum_{\ell=1}^{M} u_\ell f(a_\ell)=\gamma. \tag{7.6}$$

Solution using (7.1): It follows from (7.5) that

$$\begin{aligned}\sum_{\ell=1}^{M} u_\ell f(a_\ell) &= \sum_{j=1}^{M}\left(\sum_{\ell=1}^{M} u_\ell e_j(a_\ell)\right) f_j(0)\\ &= \sum_{j=1}^{M} v_j f_j(0),\end{aligned} \tag{7.7}$$

with

$$v_j=\sum_{\ell=1}^{M} u_\ell e_j(a_\ell),\quad j=1,\dots,M. \tag{7.8}$$

For preassigned value of the left side of (7.7) this is a classical tangential interpolation problem for $\mathbb{C}^M$-valued functions with entries in the Hardy space. Let $v=\begin{pmatrix}v_1 & v_2 & \cdots & v_M\end{pmatrix}\in\mathbb{C}^{1\times M}$. Assuming $vv^*\neq 0$ we have that the set of solutions is given by

$$\begin{pmatrix}f_1(z)\\ f_2(z)\\ \vdots\\ f_M(z)\end{pmatrix}=\gamma\frac{v^*}{vv^*}+\left(I_M+(z-1)\frac{v^*v}{vv^*}\right)\begin{pmatrix}g_1(z)\\ g_2(z)\\ \vdots\\ g_M(z)\end{pmatrix},$$

where $g_1,\dots,g_M\in\mathbf{H}_2$ and

$$\sum_{\ell=1}^{M}\|f_j\|_2^2=\frac{|\gamma|^2}{vv^*}+\sum_{\ell=1}^{M}\|g_j\|_2^2.$$

It follows from (7.1) that a function $f\in\mathbf{H}_2$ satisfies (7.6) if and only if it can be written as

$$f(z)=\frac{\gamma}{vv^*}\sum_{j=1}^{M} e_j(z)v_j^*+\begin{pmatrix}e_1(z) & e_2(z) & \cdots & e_M(z)\end{pmatrix}B(z)\begin{pmatrix}g_1(z)\\ g_2(z)\\ \vdots\\ g_M(z)\end{pmatrix},$$

where we have denoted $B(z)=\left(I_M+(z-1)\frac{v^*v}{vv^*}\right)$. Note that B is an elementary Blaschke factor, with zero at the origin.

Solution using (7.3)**:** In the case of representation (7.3), we have similarly

$$f(a_\ell) = \sum_{j=1}^{M} e_j(a_\ell)\sigma_{1j}(0), \quad \ell = 1,\dots,M. \tag{7.9}$$

For preassigned values of $f(a_\ell)$, $\ell = 1,\dots,M$, this reduces the Nevanlinna–Pick interpolation problem for M points in $\mathbf{H}_2$ to a tangential interpolation problem at the origin for matrix-valued Schur functions (7.4). As in the previous discussion we exploit the above reduction scheme to interpolation problem in the Schur class for multipoint interpolation problems. For example, in the case of the interpolation constraint (7.6), it follows from (7.9) that

$$\begin{aligned}\sum_{\ell=1}^{M} u_\ell f(a_\ell) &= \sum_{j=1}^{M} u_\ell \left(\sum_{\ell=1}^{M} e_j(a_\ell) \right) \sigma_{1j}(0) \\ &= \sum_{j=1}^{M} v_j \sigma_{1j}(0),\end{aligned} \tag{7.10}$$

with $v_1,\dots,v_M$ as in (7.8). For preassigned value of the left side of (7.10) this is a classical tangential interpolation problem for $\mathbb{C}^{M+1}$-valued Schur functions.

Problems of the form (7.9) have been studied for $M = 2$, under the name multipoint interpolation problem, in [6]. In that paper, an involution φ of the open unit disk which maps a_1 into a_2 is used. Then, one notes that the function

$$F(z) = \begin{pmatrix} f(z) \\ f(\varphi(z)) \end{pmatrix}$$

satisfies the symmetry

$$F(\varphi(z)) = JF(z), \quad \text{where} \quad J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

This reduces the interpolation problem in $\mathbf{H}_2$ to an interpolation problem with symmetries in $\mathbf{H}_2^2$. Unfortunately this method does not extend to the case $M > 2$. For a related interpolation problem (for Nevanlinna functions), see also [14], where the nth composition of the map φ is equal to the identity map: $\varphi^{\circ n}(z) = z$.

In the this chapter we use a decomposition of elements in $\mathbf{H}_2$ associated with isometries defined from b and which satisfy the Cuntz relation. The representation of Hardy functions, proved in [5], plays a major role in the reduction to interpolation problems in the setting of Schur functions. To ease the notation, we set the discussion in the framework of scalar-valued functions, but the paper itself (as well as [5]) is developed for matrix-valued functions. Besides being a key player in complex analysis, the Hardy space $\mathbf{H}_2$ of the open unit disk plays an important role in signal processing and in the theory of linear dynamical systems. An element f in $\mathbf{H}_2$ can be described in (at least) three different ways: in terms of (1) power

series, (2) integral conditions, or (3) a positive definite kernel. More precisely, in case (1), one sees f as the z-transform of a discrete signal with finite energy, that is, the z-transform of a sequence $(f_n)_{n\in\mathbb{N}_0} \in \ell_2$:

$$f(z) = \sum_{n=0}^{\infty} f_n z^n, \quad \|f\|_{\mathbf{H}_2}^2 \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} |f_n|^2 < \infty,$$

In case (2), one expresses the norm (in the equivalent way) as

$$\|f\|_2^2 = \frac{1}{2\pi} \sup_{r\in(0,1)} \int_0^{2\pi} |f(re^{it})|^2 dt < \infty,$$

and sees f as the transfer function (filter) of a ℓ_1-ℓ_2 stable linear system. See [20]. In case (3), we use the fact that $\mathbf{H}_2$ is the reproducing kernel Hilbert space with reproducing kernel $\frac{1}{1-zw^*}$. From the characterization of elements in a reproducing kernel Hilbert space, a function f defined in $\mathbb{D}$ belongs to $\mathbf{H}_2$ if and only if for some $M > 0$, the kernel

$$\frac{1}{1-zw^*} - \frac{f(z)f(w)^*}{M} \tag{7.11}$$

is positive definite there. The smallest such M is $\|f\|_2^2$. For $M = 1$, rewriting (7.11) as

$$\frac{a(z)a(w)^* - h(z)h(w)^*}{1-zw^*}, \quad \text{with} \quad a(z) = \begin{pmatrix} 1 & -zf(z) \end{pmatrix},$$

and using Leech's factorization theorem (see next section), it was proved in [5] that f admits a (in general not unique) representation of the form

$$f(z) = \frac{\sigma_1(z)}{1 - z\sigma_2(z)}, \tag{7.12}$$

where $\sigma(z) = \begin{pmatrix} \sigma_1(z) \\ \sigma_2(z) \end{pmatrix}$ is analytic and contractive in the open unit disk.

Let now $a \in \mathbb{D}$, and

$$b_a(z) = \frac{z-a}{1-za^*}.$$

In [10] it was proved that the map

$$T_a f(z) = \frac{\sqrt{1-|a|^2}}{1-za^*} f(b_a(z)) \tag{7.13}$$

is from $\mathbf{H}_2$ onto itself and unitary. In this chapter, we replace b_a by an arbitrary finite Blaschke product and define a counterpart of the operator T_a. If $M = \deg b$, we now have, instead of the unitary map, T_a a set of isometries $S_1, \ldots, S_M$ in $\mathbf{H}_2$, defined as follows: Take $e_1, \ldots, e_M$ be an orthonormal basis of the space $\mathbf{H}_2 \ominus b\mathbf{H}_2$. Then,

$$(S_j h)(z) = e_j(z)h(b(z)), \quad h \in \mathbf{H}_2, \tag{7.14}$$

with $S_1, \ldots, S_M$ satisfying the Cuntz relations:

$$\sum_{j=1}^{M} S_j S_j^* = I_{\mathbf{H}_2}, \tag{7.15}$$

$$S_j^* S_k = \begin{cases} I_{\mathbf{H}_2}, & \text{if} \quad j = k, \\ 0, & \text{otherwise.} \end{cases} \tag{7.16}$$

It follows from these relations that every element $f \in \mathbf{H}_2$ can be written in a unique way as (7.1):

$$f(z) = \sum_{j=1}^{M} e_j(z) f_j(b(z)),$$

where the $f_j \in \mathbf{H}_2$ and satisfy (7.2)

$$\|f\|_2^2 = \sum_{j=1}^{M} \|f_j\|_2^2.$$

We note that $S_1, \ldots, S_M$ in (7.14) form a finite system of M isometries with orthogonal ranges in $\mathbf{H}_2$, with the sum of the ranges equal to all of $\mathbf{H}_2$. Thus they define a representation of the Cuntz relations. This is a special case of a result of Courtney, Muhly, and Schmidt, see [18, Theorem 3.3]. We send the reader to [18] for a survey of the relevant literature and in particular for a discussion of the related papers [27, 28]. For completeness, we provide a proof, in the matrix-valued case, using reproducing kernel spaces techniques (see Sect. 7.4). As already mentioned, one motivation for studying representations of the Cuntz relations comes from signal processing. Our present application of the Cuntz relations to Leech's problem from harmonic analysis is entirely new. The immediate relevance to subband filters is a careful selecting of the Cuntz isometries, one for each frequency subband, see [26, Chapter 9]. In the case of wavelet applications, the number M is the scaling number characterizing the particular family of wavelets under discussion.

We note that relations with the Cuntz relations in the indefinite inner product case have been considered in [9], in the setting of de Branges–Rovnyak spaces; see [15, 16]. This suggests connections with interpolation in these spaces (see [2, Section 11], [12]), which will be considered elsewhere. We briefly discuss some of these aspects in Sect. 7.7.

Our paper is interdisciplinary, a mix of pure and applied, and we are motivated by several prior developments and work by other authors. This we discuss in Sects. 7.2–7.4. For the readers' convenience, we mention here briefly some of these connections. One motivation comes from earlier work [1] on two-sided, and tangential, interpolation for matrix functions, see also [2] through [5], and [14]. In addition we make a connection to interpolation in de Branges–Rovnyak spaces [12], to wavelet filters, see, e.g., [17], and to iterated function systems, see [18] by Courtney, Muhly, and Schmidt, and [28] by Rochberg, to Hardy classes [29,30] and to classical harmonic analysis, see, e.g., [32–35].

This chapter consists of six sections besides the introduction. Sects. 7.2 and 7.3 are of a review nature: in the second section we discuss Leech's theorem, and in the third section we discuss the realization result of [5]. In Sect. 7.4 we consider, in the matrix-valued case, a set of operators which satisfy the Cuntz relations and were considered earlier in [18] in the scalar case. Section 7.5 is devoted to the proof of the matrix version of (7.3). We use the representation theorem of Sect. 7.5 to solve in Sect. 7.6 new types of multipoint interpolation problems. Finally we outline in the last section how some of the results extend to the case of de Branges–Rovnyak spaces.

7.2 Leech's Theorem

As already mentioned in the introduction, we set this chapter in the framework of matrix-valued functions. When the Taylor coefficients f_n are $\mathbb{C}^{p\times q}$-valued, one defines a $\mathbb{C}^{q\times q}$-valued quadratic form by

$$[f,f] \stackrel{\text{def}}{=} \frac{1}{2\pi}\int_0^{2\pi} (f(\mathrm{e}^{it}))^* f(\mathrm{e}^{it})\mathrm{d}t = \sum_{n=0}^{\infty} f_n^* f_n.$$

The space $\mathbf{H}^{p\times q}$ consists of the functions for which Tr $[f,f] < \infty$. In [5] a representation theorem for elements $f \in \mathbf{H}_2^{p\times q}$ such that $[f,f] \le I_q$ in terms of Schur functions was presented. See Theorem 7.3.1. Recall first that a $\mathbb{C}^{p\times q}$-valued function σ *defined* in the open unit disk is analytic and contractive in the open unit disk if and only if the kernel

$$K_\sigma(z,w) = \frac{I_p - \sigma(z)\sigma(w)^*}{1 - zw^*} \tag{7.17}$$

is positive definite in the open unit disk. Such functions are called Schur functions, denoted by $\mathscr{S}^{p\times q}$. Given $\sigma \in \mathscr{S}^{p\times q}$ and a $\mathbb{C}^{k\times p}$-valued function A analytic in the open unit disk, the kernel

$$A(z)K_\sigma(z,w)A(w)^* = \frac{A(z)A(w)^* - B(z)B(w)^*}{1 - zw^*}$$

where $B = A\sigma$, is positive in $\mathbb{D}$. Leech's theorem asserts that the converse holds: If A and B are, respectively, $\mathbb{C}^{k\times p}$-valued and $\mathbb{C}^{k\times q}$-valued functions *defined* in $\mathbb{D}$ and such that the kernel

$$\frac{A(z)A(w)^* - B(z)B(w)^*}{1 - zw^*} \tag{7.18}$$

is positive definite in $\mathbb{D}$, then there exists $\sigma \in \mathscr{S}^{p\times q}$ such that $B = A\sigma$. Two proofs of this theorem hold. The first assumes that A and B are bounded in the open unit disk, and uses a commutant lifting result of M. Rosenblum. See [29] for Rosenblum's result and [30, Example 1, p 107], [7] for Leech's theorem. The other proof requires only analyticity of A and B in $\mathbb{D}$, and uses tangential interpolation theory for Schur functions, together with the normal family theorem. One can extend these arguments to functions of bounded type in $\mathbb{D}$, or even further weaken these hypothesis. For completeness, we now outline a proof of Leech's theorem for continuous functions A and B. We first recall the following: Let $N \in \mathbb{N}$ and let $w_1, \ldots, w_N \in \mathbb{D}$, $\xi_1, \ldots, \xi_N \in \mathbb{C}^p$, and $\eta_1, \ldots, \eta_N \in \mathbb{C}^q$. The tangential Nevanlinna-Pick interpolation problem consists in finding all Schur functions $\sigma \in \mathscr{S}^{p\times q}$ such that

$$\sigma(w_j)^*\xi_j = \eta_j, \quad j = 1, \ldots, N.$$

The fact that the function $K_\sigma(z,w)$ defined by (7.17) is positive definite in $\mathbb{D}$ implies that a necessary condition for the tangential Nevanlinna–Pick interpolation problem to have a solution is that the $N \times N$ Hermitian matrix P (known as the Pick matrix) with ℓj entry

$$P_{\ell,j} = \frac{\xi_\ell^*\xi_j - \eta_\ell^*\eta_j}{1 - w_\ell w_j^*} \tag{7.19}$$

is nonnegative. This condition is in fact also sufficient, and there are various methods to describe all solutions in terms of a linear fractional transformation. See, for instance [13, 19, 21, 22]. With this result at hand we can outline a proof of Leech's theorem as follows: We assume given two functions A and B, respectively, $\mathbb{C}^{k\times q}$-valued and $\mathbb{C}^{k\times q}$-valued, continuous in $\mathbb{D}$ and such that the kernel

$$\frac{A(z)A(w)^* - B(z)B(w)^*}{1 - zw^*}$$

is positive definite there. Consider $w_1, w_2 \ldots$ a countable set of points dense in the open unit disk. The Hermitian block matrix with ℓj entry

$$\frac{A(w_\ell)A(w_j)^* - B(w_\ell)B(w_j)^*}{1 - w_\ell w_j^*}, \quad \ell, j = 1, \ldots, N,$$

is nonnegative, and therefore, by the above-mentioned result on Nevanlinna–Pick interpolation, there exists a Schur function $\sigma_N \in \mathscr{S}^{p\times q}$ such that

$$A(w_\ell)\sigma_N(w_\ell) = B(w_\ell), \quad \ell = 1, \ldots N.$$

To conclude the proof, one uses the normal family theorem to find a function $\sigma \in \mathscr{S}^{p\times q}$ such that $A(w_\ell)\sigma(w_\ell) = B(w_\ell)$ for $\ell \in \mathbb{N}$. By continuity, this equality extends then to all of $\mathbb{D}$.

7.3 A Representation of $\mathbf{H}_2^{p\times 2}$ Functions

Leech's theorem can be used to find a representation of elements of $\mathbf{H}_2$ in terms of Schur functions, as we now recall. See [4, 5]. The following result is proved in [5]. In the scalar case, it was proved earlier by Sarason using different methods. See [35, p 50], [32–34]. In the discussion we shall find it convenient to partition $\sigma \in \mathscr{S}^{(p+q)\times q}$ as

$$\sigma = \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix}, \tag{7.20}$$

with σ_1 being $\mathbb{C}^{p\times q}$-valued and σ_2 being $\mathbb{C}^{q\times q}$-valued.

Theorem 7.3.1 *Let $H \in \mathbf{H}_2^{p\times q}$. Then, the following are equivalent:*

(1) It holds that

$$[H,H] \le I_q. \tag{7.21}$$

(2) The kernel

$$\frac{I_p}{1-zw^*} - H(z)H(w)^* \tag{7.22}$$

is positive definite in $\mathbb{D}$.

(3) There is a Schur function $\sigma \in \mathscr{S}^{(p+q)\times q}$ (see (7.20)) so that

$$H(z) = \sigma_1(z)(I_q - z\sigma_2(z))^{-1}. \tag{7.23}$$

The key to proof of this theorem is to note that the kernel (7.22) can be rewritten in the form (7.18) with

$$A(z) = \begin{pmatrix} I_p & zH(z) \end{pmatrix} \quad \text{and} \quad B(z) = H(z),$$

and apply Leech's theorem: There exists $\sigma \in \mathscr{S}^{(p+q)\times q}$ as in (7.20) such that $A(z)\sigma(z) = B(z)$, that is,

$$\sigma_1(z) + zH(z)\sigma_2(z) = H(z).$$

Equation (7.3.1) follows.

We note that an extension of the previous theorem to elements in the Arveson space was given in [3, Theorem 10.3, p 182].

7.4 The Cuntz Relations in $\mathbf{H}_2^p$

Let b be a finite Blaschke product of degree M, and let

$$\mathscr{H}(b) = \mathbf{H}_2 \ominus b\mathbf{H}_2.$$

It is well known that this space is finite dimensional and R_0-invariant, where

$$R_0 f(z) = \frac{f(z) - f(0)}{z}, \quad f \in \mathscr{H}(b).$$

Let

$$\big(f_1(z) \cdots f_M(z)\big) = C(I_M - zA)^{-1}$$

denote a basis of $\mathscr{H}(b)$, where $(C,A) \in \mathbb{C}^{1\times M} \times \mathbb{C}^{M\times M}$ is an observable pair, namely

$$\cap_{n=0}^{\infty} \ker CA^n = \{0\}.$$

Since the spectrum of A is inside the open unit disk, the series

$$P = \sum_{\ell=0}^{\infty} A^{\ell *} C^* C A^{\ell}$$

converges and $P > 0$. The matrix P is the Gram matrix (observability Gramian in control terminology) of the basis $f_1, \ldots, f_M$ and satisfies

$$P = \frac{1}{2\pi} \int_0^{2\pi} \big(f_1(\mathrm{e}^{it}) \cdots f_M(\mathrm{e}^{it})\big)^* \big(f_1(\mathrm{e}^{it}) \cdots f_M(\mathrm{e}^{it})\big)\, \mathrm{d}t.$$

This matrix turns to be identical to the Pick matrix defined in (7.19). We denote by $\mathbf{H}_2^p$ the Hilbert space of $\mathbb{C}^p$-valued functions with entries in $\mathbf{H}_2$ and with norm

$$\left\| \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix} \right\|_{\mathbf{H}_2^p}^2 = \sum_{n=1}^{p} \|h_n\|_{\mathbf{H}_2}^2.$$

We note that $\mathbf{H}_2^p$ is the reproducing kernel Hilbert space with reproducing kernel $\frac{I_p}{1-zw^*}$.

We now introduce the Cuntz relations into this framework. This was treated earlier in [18] using different methods.

Theorem 7.4.1 *Let $e_1, \ldots, e_M$ be an orthonormal basis of $\mathscr{H}(b)$ and, for $j = 1, \ldots M$*

$$(S_j h)(z) = e_j(z) h(b(z)), \quad h \in \mathbf{H}_2^p. \tag{7.24}$$

Then, the S_j satisfy the Cuntz relations:

$$\sum_{j=1}^{M} S_j S_j^* = I_{\mathbf{H}_2^p}, \tag{7.25}$$

$$S_j^* S_k = \begin{cases} I_{\mathbf{H}_2^p}, & \text{if} \quad j = k, \\ 0, & \text{otherwise.} \end{cases} \tag{7.26}$$

Proof. We proceed in a number of steps. The proof of the Cuntz identities (7.25)-(7.26) is given in Steps 4 and 5, respectively.

Step 1: *The set $\mathbf{H}_2^p(b)$ of functions of the form*

$$F(z) = f(b(z)), \quad f \in \mathbf{H}_2^p,$$

with norm

$$\|F\|_{\mathbf{H}_2^p(b)} = \|f\|_{\mathbf{H}_2^p}$$

is the reproducing kernel Hilbert space with reproducing kernel $\frac{I_p}{1-b(z)b(w)^*}$.

This can be checked directly but is also a special case of [8, Theorem 3.1, p 109].

Step 2: *The operator M_{e_j} of multiplication by e_j is an isometry from $\mathbf{H}_2^p(b)$ into $\mathbf{H}_2^p$. Furthermore, the range of M_{e_j} and M_{e_k} are orthogonal for $j \neq k$.*
Indeed, let $u, v \in \mathbb{C}^p$. It holds that

$$\langle e_j b^n u, e_k b^m v \rangle_{\mathbf{H}_2^p} = \begin{cases} v^* u & j = k \quad \text{and} \quad m = n, \\ 0, & \text{otherwise.} \end{cases}$$

We use that multiplication by b is an isometry from $\mathbf{H}_2$ into itself. If $n = m$ and $j = k$, the claim is clear. If $n = m$ and $j \neq k$, this is just the orthogonality of e_j and e_k. If $n > m$, we have

$$\langle e_j b^n u, e_k b^m v \rangle_{\mathbf{H}_2^p} = \langle e_j b^{n-m} u, e_k v \rangle_{\mathbf{H}_2^p} = 0,$$

since $e_j b^{n-m} u \in b\mathbf{H}_2^p$ is orthogonal to $e_k v$ whose components belong to $\mathscr{H}(b) = \mathbf{H}_2 \ominus b\mathbf{H}_2$. The case $n < m$ is obtained by interchanging the role of j and k.

Step 3: *Let $e_1, \ldots, e_M$ denote an orthonormal basis of $\mathscr{H}(b)$. Then,*

$$\mathbf{H}_2^p = \oplus_{j=1}^M e_j \mathbf{H}_2^p(b). \tag{7.27}$$

Indeed, the reproducing kernel is written in terms of the orthonormal basis as (see, for instance, [11, (6) p 346], [31])

$$K_b(z,w) = \sum_{j=1}^M e_j(z)(e_j(w))^*. \tag{7.28}$$

Thus

$$\frac{I_p}{1-zw^*} = \frac{I_p}{1-b(z)b(w)^*}\frac{1-b(z)b(w)^*}{1-zw^*} = \sum_{j=1}^M k_j(z,w), \tag{7.29}$$

with

$$k_j(z,w) = \frac{e_j(z)(e_j(w))^* I_p}{1-b(z)b(w)^*}.$$

Equality (7.29) expresses the positive definite kernel $\frac{I_p}{1-zw^*}$ as a sum of positive definite kernels. The reproducing kernel space associated to k_j is $e_j\mathbf{H}_2(b)$. Therefore, (7.27) holds as a sum of vector spaces, see [11, p 352]. Since, by Step 2, $e_j\mathbf{H}_2^p(b)$ is isometrically included into $\mathbf{H}_2^p$, the sum is orthogonal.

Step 4: *S_j and S_k are isometries, with orthogonal ranges when $j \neq k$.*

The fact that S_j is an isometry follows from Steps 1 and 2. Indeed the range of S_j is in $\mathbf{H}_2^p$ by Step 2 and

$$\begin{aligned}\|S_j h\|_{\mathbf{H}_2^p}^2 &= \|h(b)\|_{\mathbf{H}_2^p(b)}^2 \quad \text{(by Step 2)}\\ &= \|h\|_{\mathbf{H}_2^p}^2 \quad \text{(by Step 1)}.\end{aligned}$$

Furthermore, for $f, g \in \mathbf{H}_2^p$ and $j \neq k$,

$$\langle S_j f, S_k g\rangle_{\mathbf{H}_2^p} = \langle M_{e_j} f(b), M_{e_k} g(b)\rangle_{\mathbf{H}_2^p} = 0$$

by Step 2.

Step 5: *It holds that $\sum_{j=1}^M S_j S_j^* = I_{\mathbf{H}_2^p}$.*

Indeed, by the properties of multiplication and composition operators in reproducing kernel Hilbert spaces, we have that, with $\rho_w(z) = 1 - zw^*$, and $u \in \mathbb{C}^p$:

$$\left(S_j^* \frac{u}{\rho_w}\right)(z) = \frac{u}{\rho_{b(w)}(z)}(e_j(w))^*.$$

Thus

$$\left(\sum_{j=1}^{M} S_j S_j^*\right)\left(\frac{u}{\rho_w}\right)(z) = \sum_{j=1}^{M} \frac{1}{1-b(z)b(w)^*} e_j(z)(e_j(w))^* = \frac{u}{\rho_w(z)},$$

where we have used (7.28) and (7.29). This ends the proof since the closed linear span of the functions $\frac{1}{\rho_w}$ is all of $\mathbf{H}_2$. □

Thus we have the following decomposition result for elements in $\mathbf{H}_2$. When $M = 1$, this result reduces to the fact that the operator T_a defined in (7.13) is a unitary map from $\mathbf{H}_2$ into itself. To avoid confusion recall that $\mathscr{H}(b)$ denotes $H_2 \ominus bH_2$ while $H_2^p(b)$ has been defined in Step 1 in the proof of Theorem 7.4.1.

Theorem 7.4.2 *Let b be a finite Blaschke product of degree M, and let $e_1, \ldots, e_M$ be an orthonormal basis of $\mathscr{H}(b)$. Then, every element $H \in \mathbf{H}_2^{p\times q}$ can be written in a unique way as*

$$H(z) = \sum_{j=1}^{M} e_j(z) H_j(b(z)), \tag{7.30}$$

where the $H_j \in \mathbf{H}_2^{p\times q}$ and

$$[H,H] = \sum_{j=1}^{M} [H_j, H_j]. \tag{7.31}$$

Proof. We define operators $S_1, \ldots, S_M$ as in (7.24). Let $H \in \mathbf{H}_2^{p\times q}$ and $\xi \in \mathbb{C}^q$. It follows from the definition (7.24) of the S_j and from (7.25) that

$$H(z)\xi = \sum_{j=1}^{M} e_j(z) H_j \xi(z),$$

where $H_j \in \mathbf{H}_2^{p\times q}$ is defined by $H_j\xi = S_j^*(H\xi)$. Taking now into account (7.26) we have

$$\begin{aligned}
\xi^*[H,H]\xi &= \langle H\xi, H\xi\rangle_{\mathbf{H}_2^p} \\
&= \sum_{\ell,j=1}^{M} \langle S_\ell S_\ell^*(H\xi), S_j S_j^*(H\xi)\rangle_{\mathbf{H}_2^p} \\
&= \sum_{j=1}^{M} \langle S_j^*(H\xi), S_j^*(H\xi)\rangle_{\mathbf{H}_2^p} \\
&= \sum_{j=1}^{M} \langle H_j\xi, H_j\xi\rangle_{\mathbf{H}_2^p} \\
&= \xi^* \sum_{j=1}^{M} [H_j, H_j]\xi,
\end{aligned}$$

and hence (7.31) holds. □

7.5 Representation of Elements of $\mathbf{H}_2^{p\times q}$

We now present a generalization of Theorem 7.3.1. To this end, we generalize (7.4) to a partitioning of a matrix-valued functions $\sigma \in \mathscr{S}^{(Mp+q)\times q}$ as

$$\sigma = \begin{pmatrix} \sigma_{11} \\ \vdots \\ \sigma_{1M} \\ \sigma_2 \end{pmatrix}, \tag{7.32}$$

where $\sigma_{11}, \ldots, \sigma_{1M}$ and σ_2 are $\mathbb{C}^{p\times q}$-valued and $\mathbb{C}^{q\times q}$-valued, respectively.

Theorem 7.5.1 *Let b be a preassigned finite Blaschke product, and let $e_1, \ldots, e_M$ be an orthonormal basis of $\mathscr{H}(b)$. Let $H \in \mathbf{H}_2^{p\times q}$. Then, the following are equivalent:*

(1) Condition (7.21) holds $[H,H] \leq I_q$.
(2) There exists $\sigma \in \mathscr{S}^{(Mp+q)\times q}$ such that

$$H(z) = \left(\sum_{j=1}^{M} e_j(z)\sigma_{1j}(b(z)) \right) (I_q - b(z)\sigma_2(b(z)))^{-1}, \tag{7.33}$$

where $\sigma \in \mathscr{S}^{(Mp+q)\times q}$ is as in (7.32).

Proof. By Theorem 7.4.2, H subject to (7.21) can be written in a unique way as (7.30), and it follows from (7.31) that

$$\sum_{j=1}^{M} [H_j, H_j] \leq I_q.$$

Using Theorem 7.3.1 with the function

$$G = \begin{pmatrix} H_1 \\ H_2 \\ \vdots \\ H_M \end{pmatrix} \in \mathbf{H}_2^{Mp\times q},$$

we see that there exists $\sigma \in \mathscr{S}^{(Mp+q)\times q}$ (see (7.32)) such that

$$\begin{pmatrix} H_1(z) \\ H_2(z) \\ \vdots \\ H_M(z) \end{pmatrix} = \begin{pmatrix} \sigma_{11}(z) \\ \sigma_{12}(z) \\ \vdots \\ \sigma_{1M}(z) \end{pmatrix} (I_q - z\sigma_2(z))^{-1}.$$

The result follows using (7.30). □

The results in [5] are a special case of a family of interpolation problems with relaxed constraints. See [24, 25]. We plan in a future publication to consider these results in our new extended setting.

7.6 New Interpolation Problems

We have outlined in the introduction the connections between multipoint interpolations and representations (7.1) and (7.3). We now add some details. Interpolation problems whose solutions are outlined in this section will be considered in full details in a future publication.

The case of (7.1): We consider the following problem: Find all functions $H \in \mathbf{H}_2^{p\times q}$ such that

$$\sum_{j=1}^{M} \xi_j H^{(j-1)}(a) = \gamma, \tag{7.34}$$

for some preassigned matrices $\xi_1, \ldots, \xi_M \in \mathbb{C}^{r\times p}$ and $\gamma \in \mathbb{C}^{r\times q}$. To solve this problem we use (7.30) with

$$b(z) = \left(\frac{z-a}{1-za^*}\right)^M.$$

A basis of $\mathscr{H}(b)$ is given by

$$\frac{1}{1-za^*}, \frac{z}{(1-za^*)^2}, \ldots, \frac{z^{M-1}}{(1-za^*)^M}.$$

(see, for instance, [21]). Set

$$E(z) = \left(\tfrac{1}{1-za^*} I_p \ \tfrac{z}{(1-za^*)^2} I_p \ \cdots \ \tfrac{z^{M-1}}{(1-za^*)^M} I_p\right). \tag{7.35}$$

Since

$$b(a) = b'(a) = \cdots = b^{(M-1)}(a) = 0, \tag{7.36}$$

and with

$$\mathscr{H}(z) = \begin{pmatrix} H_1(b(z)) \\ H_2(b(z)) \\ \vdots \\ H_M(b(z)) \end{pmatrix} \in \mathbf{H}_2^{Mp\times q},$$

we have that

$$H(a) = E(a)\mathscr{H}(0)$$
$$H'(a) = E'(a)\mathscr{H}(0),$$
$$\vdots$$
$$H^{(M-1)}(0) = E^{(M-1)}(a)\mathscr{H}(0),$$

Therefore, the interpolation problem (7.34) is equivalent to

$$C\mathscr{H}(0) = \gamma,$$

with $C \in \mathbb{C}^{r\times Mp}$ given by

$$C = \sum_{j=0}^{M-1} \xi_j E^{(j)}(a).$$

When $CC^* > 0$, this in turn can be solved using [2, Section 7], or directly, as

$$\mathscr{H}(z) = C^*(CC^*)^{-1}\gamma + \big(I_{Mp} + (z-1)C^*(CC^*)^{-1}C\big)\mathscr{G}(z),$$

where $\mathscr{G} \in \mathbf{H}_2^{Mp\times q}$. The formula for H follows. We note that

$$[\mathscr{H},\mathscr{H}] = \gamma^*(CC^*)^{-1}\gamma + [\mathscr{G},\mathscr{G}].$$

The case where CC^* is not invertible is solved using pseudo-inverses.

The case of (7.3): We here assume first that $p = q = 1$ and

$$b(z) = \prod_{\ell=1}^{M} \frac{z - a_\ell}{1 - z a_\ell^*},$$

where the a_ℓ are distinct points in $\mathbb{D}$. We have now

$$C = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \quad \text{and} \quad A = \operatorname{diag}(a_1^*, a_2^*, \ldots, a_M^*).$$

Note that the pair (C,A) is observable. Define

$$P_{\ell j} = \frac{1}{1 - a_j a_\ell^*}, \quad \ell, j = 1, \ldots, M. \tag{7.37}$$

Namely we are in the case (7.19) with the $\xi_j = 1$ and the $\eta_j = 0$. In other words, P is the Pick matrix obtained while interpolating all the points a_ℓ to the origin. We mention the papers [23, 36] for a related discussion.

Proposition 7.6.1 *Let* $u = \begin{pmatrix} u_1 & u_2 & \cdots & u_M \end{pmatrix} \in \mathbb{C}^{1\times M}$ *and* $\gamma \in \mathbb{C}$ *be preassigned. Then, the following are equivalent:*

1. *It holds that*

$$\sum_{\ell=1}^{M} u_\ell h(a_\ell) = \gamma \quad \text{and} \quad \|h\|_{\mathbf{H}_2} \leq 1.$$

2. *h is of the form*

$$C(I - zA)^{-1}P^{-1/2}\sigma_1(b(z))(1 - b(z)\sigma_2(b(z))^{-1},$$

where $\sigma = \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} \in \mathscr{S}^{M+1}$ *is such that*

$$\begin{pmatrix} uP^{1/2} & 0 \end{pmatrix} \sigma(0) = \gamma, \tag{7.38}$$

where P is defined by (7.37).

When $b(z) = (\frac{z-a}{1-za^*})^M$ for some $M \in \mathbb{N}$ and $a \in \mathbb{D}$, one obtains a different kind of interpolation problem, as we now explain. Rewriting (7.33) as

$$H(z)(I_q - b(z)\sigma_2(b(z))) = E(z)\sigma(b(z)), \tag{7.39}$$

where $E(z)$ is given by (7.35) with $p = 1$, and

$$\sigma_1(z) = \begin{pmatrix} \sigma_{11}(z) \\ \sigma_{12}(z) \\ \vdots \\ \sigma_{1M}(z) \end{pmatrix}.$$

Differentiating (7.39) $M-1$ times and taking into account (7.36), we obtain that

$$\begin{pmatrix} H(a) \\ H'(a) \\ \vdots \\ H^{(M-1)}(a) \end{pmatrix} = \begin{pmatrix} E(a) & 0_{p\times p} \\ E'(a) & 0_{p\times p} \\ & \vdots \\ E^{(M-1)}(a) & 0_{p\times p} \end{pmatrix} \begin{pmatrix} \sigma_1(0) \\ \sigma_2(0) \end{pmatrix}. \tag{7.40}$$

This allows to reduce (7.34) to a standard tangential interpolation problem for Schur functions.

7.7 The Case of De Branges Rovnyak Spaces

Let s be a Schur function. The kernel $k_s(z,w) = \frac{1-s(z)s(w)^*}{1-zw^*}$ is positive definite in the open unit disk, and the associated reproducing kernel Hilbert space will be denoted by $\mathscr{H}(s)$. Such spaces were introduced and studied in depth by de Branges and Rovnyak in [16]. When s is an inner function (and in particular when s is finite Blaschke product), we have

$$\mathscr{H}(s) = \mathbf{H}_2 \ominus s\mathbf{H}_2.$$

In general, $\mathscr{H}(s)$ is only contractively included in $\mathbf{H}_2$. Let moreover b be a finite Blaschke product. We have

$$\begin{aligned}\frac{1-s(b(z))s(b(w))^*}{1-zw^*} &= \frac{1-s(b(z))s(b(w))^*}{1-b(z)b(w)^*}\frac{1-b(z)b(w)^*}{1-zw^*}\\ &= \sum_{j=1}^{M} e_j(z)e_j(w)^*\frac{1-s(b(z))s(b(w))^*}{1-b(z)b(w)^*},\end{aligned} \tag{7.41}$$

with $e_1,\ldots,e_M$ an orthonormal basis of $\mathscr{H}(b)$. This decomposition allows us to define the operators $S_1,\ldots,S_M$ as in (7.24), so that the following holds:

Theorem 7.7.1 *The operators $S_1,\ldots,S_M$ are continuous from $\mathscr{H}(s)$ into $\mathscr{H}(s(b))$ and satisfy the Cuntz relations:*

$$\sum_{j=1}^{M} S_jS_j^* = I_{\mathscr{H}(s(b))}, \tag{7.42}$$

$$S_j^*S_k = \begin{cases} I_{\mathscr{H}(s)}, & \text{if} \quad j=k,\\ 0, & \text{otherwise.}\end{cases} \tag{7.43}$$

Proof. We proceed in a number of steps.

Step 1: *The reproducing kernel Hilbert space $\mathscr{M}(s,b)$ with reproducing kernel $\frac{1-s(b(z))s(b(w))^*}{1-b(z)b(w)^*}$ consists of the functions of the form $F(z) = f(b(z))$, with $f \in \mathscr{H}(s)$ and norm*

$$\|F\|_{\mathscr{M}(s,b)} = \|f\|_{\mathscr{H}(s)}.$$

This follows from a direct computation.

Step 2: *The formula*

$$\left(T_j(k_{s(b)}(\cdot,w))\right)(z) = k_s(z,b(w))e_j(w)^*, \quad w \in \mathbb{D}$$

defines a bounded densely defined operator, which has an extension to all of $\mathscr{H}(s(b))$*, and whose adjoint is* S_j.

This follows from the decomposition (7.41).

Step 3: (7.42) *holds.*

Indeed,

$$\left(\left(\sum_{j=1}^{M} S_j S_j^*\right) k_{s(b)}(\cdot, w)\right)(z) = \sum_{j=1}^{M} e_j(z) k_s(b(z), b(w)) e_j(w)^* = k_{s(b)}(z, w),$$

and hence, by continuity, equality (7.42) holds in $\mathscr{H}(s(b))$.

Step 4: *For* $j \neq k$ *we have*

$$e_j \mathscr{M}(s,b) \cap e_k \mathscr{M}(s,b) = \{0\} \tag{7.44}$$

Indeed, let $\mathbf{H}_2(b)$ be as in Step 7.7 in the proof of Theorem 7.4.1. We have $\mathscr{H}(s) \subset \mathbf{H}_2$ and hence

$$\mathscr{M}(s,b) \subset \mathbf{H}_2(b).$$

Thus (7.44) follows from Step 7.7 of that same theorem.

Step 5: *Let* M_{e_j} *denote the operator of multiplication by* e_j*. It holds that*

$$\mathscr{H}(s(b)) = \oplus_{j=1}^{M} M_{e_j} \mathscr{M}(s,b). \tag{7.45}$$

This follows from the decomposition (7.41), which implies that the sum

$$\mathscr{H}(s(b)) = \sum_{j=1}^{M} M_{e_j} \mathscr{M}(s,b)$$

holds, and from Step 7.7, which insures that the sum is direct.

Step 5: (7.43) *holds.*

Indeed, from Step 7.7, the range of the operators M_{e_j} and M_{e_k} is orthogonal for $j \neq k$, and M_{e_j} is an isometry. □

Finally, we remark that (7.42) leads to decompositions of elements of the space $\mathscr{H}(s(b))$ in terms of elements of the space $\mathscr{H}(s)$ similar to (7.1): Every element $f \in \mathscr{H}(s(b))$ can be written in a unique way as

$$f(z) = \sum_{j=1}^{M} e_j(z) f_j(b(z)),$$

where $f_1, \ldots, f_M \in \mathscr{H}(s)$. Furthermore

$$\|f\|^2_{\mathscr{H}(s(b))} = \sum_{j=1}^{M} \|f_j\|^2_{\mathscr{H}(s)}.$$

Multipoint interpolation problems can be also considered in this setting, building in particular on the recent work of Ball, Bolotnikov, and ter Horst [12] on interpolation in de Branges–Rovnyak spaces. This will be developed in a separate publication.

Our paper is meant as an interdisciplinary contribution, and it involves an approach to filters and to operators having its genesis in many different fields, both within mathematics and within engineering. We hope that we have succeeded at least partially in communicating across traditional lines of division separating these fields. As a result our listed references included below are likely to be incomplete. We thank in particular Professor Paul Muhly for improving our reference list.

With apologies to Goethe and to Frenchmen:

Mathematicians are like Frenchmen: whatever you say to them they translate into their own language and forthwith it is something entirely different

Johann Wolfgang von Goethe.

Acknowledgements The work of the first three authors was supported in part by the Binational Science Foundation grant 2010117. D. Alpay thanks the Earl Katz family for endowing the chair which supported his research. The work was done in part while the second named author visited Department of Mathematics, Ben Gurion University of the Negev, supported by a BGU distinguished visiting scientist program. Support and hospitality is much appreciated. We acknowledge discussions with colleagues there, and in the US, Dorin Dutkay, Myung-Sin Song, and Erin Pearse.

References

1. Alpay D, Bolotnikov V (1995) Two-sided interpolation for matrix functions with entries in the Hardy space. Linear Algebra Appl 223/224:31–56 (Special issue honoring Miroslav Fiedler and Vlastimil Pták.)
2. Alpay D, Bolotnikov V (1997) On tangential interpolation in reproducing kernel Hilbert space modules and applications. In: Dym H, Fritzsche B, Katsnelson V, Kirstein B (eds.) Topics in interpolation theory, vol 95 of operator theory: advances and applications. Birkhäuser Verlag, Basel, pp 37–68
3. Alpay D, Bolotnikov V, Kaptanoğlu HT (2002) The Schur algorithm and reproducing kernel Hilbert spaces in the ball. Linear Algebra Appl 342:163–186
4. Alpay D, Bolotnikov V, Loubaton Ph (1996) On tangential H_2 interpolation with second order norm constraints. Integral Equat Oper Theor 24:156–178
5. Alpay D, Bolotnikov V, Peretz Y (1995) On the tangential interpolation problem for H_2 functions. Trans Am Math Soc 347:675–686
6. Alpay D, Bolotnikov V, Rodman L (1998) Tangential interpolation with symmetries and two-point interpolation problem for matrix-valued H_2-functions. Integral Equat Oper Theor 32(1):1–28
7. Alpay D, Dewilde P, Dym H (1989) On the existence and construction of solutions to the partial lossless inverse scattering problem with applications to estimation theory. IEEE Trans Inform Theor 35:1184–1205
8. Alpay D, Dym H (1993) On a new class of reproducing kernel spaces and a new generalization of the Iohvidov laws. Linear Algebra Appl 178:109–183

9. Alpay D, Jorgensen P, Lewkowicz I (2011). Extending wavelet filters. Infinite dimensions, the non-rational case, and indefinite-inner product space. To appear in: Proceedings of the February Fourier Talks (FFT 2006 -2011). Springer-Birkhauser Applied and Numerical Harmonic Analysis (ANHA) Book Series.
10. Alpay D, Mboup M (2010) Discrete-time multi-scale systems. Integral Equat Oper Theor 68:163–191
11. Aronszajn N (1950) Theory of reproducing kernels. Trans Am Math Soc 68:337–404
12. Ball Joseph A, Bolotnikov V, ter Horst S (2011) Interpolation in de Branges-Rovnyak spaces. Proc Am Math Soc 139(2):609–618
13. Ball J, Gohberg I, Rodman L (1990) Interpolation of rational matrix functions, vol 45 of operator theory: advances and applications. Birkhäuser Verlag, Basel
14. Bekker MB, Nudel′man AA (2005) Tangential interpolation problems for a class of automorphic matrix-functions. Linear Algebra Appl 397:171–184
15. Branges L. de, Rovnyak J (1966) Canonical models in quantum scattering theory. In: Wilcox C (ed) Perturbation theory and its applications in quantum mechanics. Wiley, New York, pp 295–392
16. Branges L. de, Rovnyak J (1966) Square summable power series. Holt, Rinehart and Winston, New York
17. Bratteli O, Jorgensen P (2002) Wavelets through a looking glass. Applied and Numerical Harmonic Analysis. Birkhäuser Boston Inc., Boston, MA The world of the spectrum.
18. Courtney D, Muhly Paul S, Schmidt Samuel W (2012) Composition operators and endomorphisms. Complex Anal Oper Theor (To appear) 6(1):163–188
19. Derkach V, Dym H (2010) Bitangential interpolation in generalized Schur classes. Complex Anal Oper Theor 4(4):701–765
20. Doyle J, Francis B, Tannenbaum A (1992) Feedback control theory. Macmillan Publishing Company, New York
21. Dym H (1989) *J*–contractive matrix functions, reproducing kernel Hilbert spaces and interpolation. In: Published for the conference board of the mathematical sciences. Washington, DC
22. Dym H (2003) Linear fractional transformations, Riccati equations and bitangential interpolation, revisited. In: Alpay D (ed) Reproducing kernel spaces and applications, vol 143 of Oper Theory Adv Appl. Birkhäuser, Basel, pp 171–212
23. FitzGerald Carl H, Horn Roger A (1982) On quadratic and bilinear forms in function theory. Proc London Math Soc (3) 44(3):554–576
24. Frazho AE, Horst S ter, Kaashoek MA (2006) All solutions to the relaxed commutant lifting problem. Acta Sci Math (Szeged) 72(1–2):299–318
25. Frazho AE, Horst S ter, Kaashoek MA (2006) Coupling and relaxed commutant lifting. Integral Equat Oper Theor 54(1):33–67
26. Jorgensen Palle ET (2006) Analysis and probability: wavelets, signals, fractals, vol 234 of graduate texts in mathematics. Springer, New York
27. McDonald John N (2003) Adjoints of a class of composition operators. Proc Am Math Soc 131(2):601–606 (electronic)
28. Rochberg R (1973) Linear maps of the disk algebra. Pacific J Math 44:337–354
29. Rosenblum M (1980) A corona theorem for countably many functions. Integral Equat Oper Theor 3:125–137
30. Rosenblum M, Rovnyak J (1985) Hardy classes and operator theory. Birkhäuser Verlag, Basel
31. Saitoh S (1988) Theory of reproducing kernels and its applications, vol 189. Longman scientific and technical
32. Sarason D (1986) Shift–invariant spaces from the brangesian point of view. In: Proceedings of the symposium on the occasion of the proof of the Bieberbach conjecture, math. survey monographs, vol 21. American Mathematical Society, Providence, RI
33. Sarason D (1989) Exposed points in H^1, I, vol 41 of operator theory: advances and applications . Birkhäuser Verlag, Basel, pp 485–496
34. Sarason D (1990) Exposed points in H^1, II, vol 48 of operator theory: advances and applications. Birkhäuser Verlag, Basel, pp 333–347

35. Sarason D (1990) Function theory and de Branges spaces, vol 51 of Proceedings of symposia in pure mathematics. American Mathematical Society, pp 495–501
36. Szafraniec FH (1986) On bounded holomorphic interpolation in several variables. Monatshefte für. Mathematik 101:59–66

Chapter 8
Constructions and a Generalization of Perfect Autocorrelation Sequences on $\mathbb{Z}$

John J. Benedetto and Somantika Datta

Abstract Low autocorrelation signals have fundamental applications in radar and communications. We construct constant amplitude zero autocorrelation (CAZAC) sequences x on the integers $\mathbb{Z}$ by means of Hadamard matrices. We then generalize this approach to construct unimodular sequences x on $\mathbb{Z}$ whose autocorrelations A_x are building blocks for all functions on $\mathbb{Z}$. As such, algebraic relations between A_x and A_y become relevant. We provide conditions for the validity of the formulas $A_{x+y} = A_x + A_y$.

8.1 Introduction

8.1.1 Background

Let $\mathbb{R}$ be the real numbers, let $\mathbb{Z}$ be the integers, and set $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. A general problem is to characterize the family of positive bounded Radon measures F, whose inverse Fourier transforms are the autocorrelations of bounded sequences x. A special case is when $F \equiv 1$ on $\mathbb{T}$ and x is unimodular on $\mathbb{Z}$. The statement that $F \equiv 1$ is the same as saying that the autocorrelation of x vanishes except at 0, where it takes the value 1. We shall construct such unimodular sequences x based on the analysis of Hadamard matrices.

J.J. Benedetto (✉)
Department of Mathematics, Norbert Wiener Center, University of Maryland, College Park, MD 20742, USA
e-mail: jjb@math.umd.edu

S. Datta
Department of Mathematics, University of Idaho, Moscow, ID 83844, USA
e-mail: sdatta@uidaho.edu

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_8,

The problem of constructing unimodular sequences with zero autocorrelation, which our constructions address, is central in the general area of waveform design, and it is particularly relevant in several applications in the areas of radar and communications and in the general area of constructing phase coded waveforms on $\mathbb{R}$ with optimal narrow band ambiguity function behavior. In radar, the sequences x can play a role in effective target recognition, see, e.g., [1, 9, 15, 20–23, 28]; and in communications they can be used to address synchronization issues in cellular (phone) access technologies, especially code division multiple access (CDMA), e.g., [30–32]. With regard to the narrow band ambiguity function we refer to [5, 6, 20, 25], which in turn refer to the vast literature in this subject.

In radar there are two main reasons that the sequences x should be unimodular, that is, have constant amplitude. First, a transmitter can operate at peak power if x has constant peak amplitude—the system does not have to deal with the surprise of greater than expected amplitudes. Second, amplitude variations during transmission due to additive noise can be theoretically eliminated. The zero autocorrelation property ensures minimum interference between signals sharing the same channel.

8.1.2 Autocorrelation

We shall use the standard notation from harmonic analysis, e.g., [4, 27]. $\mathbb{N}$ is the set of natural numbers and $\mathbb{C}$ is the set of complex numbers. $C(\mathbb{T}^d)$ is the space of $\mathbb{C}$-valued continuous functions on $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$, and $A(\mathbb{T}^d)$ is the subspace of absolutely convergent Fourier series. $M(\mathbb{T}^d)$ is the space of bounded Radon measures on $\mathbb{T}^d$, i.e., $M(\mathbb{T}^d)$ is the dual space of the Banach space $C(\mathbb{T}^d)$ taken with the sup norm. $L^1(\mathbb{T})$ and $L^2(\mathbb{T})$ are the spaces of integrable and square integrable functions on $\mathbb{T}$, respectively. For a given $\lambda > 0$, the L^1-dilation of f, f_λ, is defined as $f_\lambda(t) = \lambda f(\lambda t)$. Let $\triangle(t) = \max(1-|t|, 0)$ on $\mathbb{R}$. Let $\omega(\gamma) = \frac{1}{2\pi}\left(\frac{\sin \gamma/2}{\gamma/2}\right)^2$; ω is called the *Fejér function* [4]. The Fourier transform of $f \in L^1(\mathbb{R})$ is the function $\widehat{f}$ defined by

$$\widehat{f}(\gamma) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i t\gamma} dt, \ \gamma \in \widehat{\mathbb{R}} \ (= \mathbb{R}).$$

$A(\widehat{\mathbb{R}})$ denotes the space of such absolutely convergent Fourier transforms on $\widehat{\mathbb{R}}$, with an analogous definition for $A(\widehat{\mathbb{R}}^d)$. We write the pairing between the function f and $\widehat{f}$ as $f \leftrightarrow \widehat{f}$. The Fourier transform of $\triangle$ is $\omega_{2\pi}$. The complex conjugate of a function f at a point t is denoted by $\overline{f(t)}$. For a set E, the measure of E is denoted by $|E|$. Given two sets A and B, the set $A \setminus B$ consists of all elements in A that are not in B.

Definition 1. The *autocorrelation* $A_x : \mathbb{Z} \to \mathbb{C}$ of $x : \mathbb{Z} \to \mathbb{C}$ is formally defined as

$$\forall k \in \mathbb{Z}, \quad A_x[k] = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} x[k+m]\overline{x[m]}.$$

(Lower case Roman letters, such as x, are often used in some applied communities to denote functions $\mathbb{Z} \to \mathbb{C}$.) There is an analogous definition of autocorrelation for functions $f : \mathbb{R}^d \to \mathbb{C}$, e.g., see Theorem 1.

If $F \in A(\mathbb{T}^d)$ we write $\check{F} = f = \{f_k\}$, i.e., $\check{F}[k] = f_k$, where, for all $k \in \mathbb{Z}^d$, $f_k = \int_{\mathbb{T}^d} F(\gamma) \mathrm{e}^{2\pi i k\cdot\gamma} \mathrm{d}\gamma$. There is a similar definition for $\check{\mu}$ where $\mu \in M(\mathbb{T}^d)$, e.g., see Theorem 1.

In the setting of $\mathbb{R}$, we have the following theorem due to Wiener and Wintner [36], which was later extended to $\mathbb{R}^d$ in [3, 18].

Theorem 1. *Let μ be a bounded positive Radon measure on $\mathbb{R}$. There is a constructible function $f \in L^\infty_{loc}(\mathbb{R})$ whose autocorrelation A_f exists for all $t \in \mathbb{R}$, and $A_f = \check{\mu}$ on $\mathbb{R}$, i.e.,*

$$\forall t \in \mathbb{R}, \quad \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t+x)\overline{f(x)} \mathrm{d}x = \int_{\mathbb{R}} \mathrm{e}^{2\pi i t x} \mathrm{d}\mu(x).$$

For any positive integer N, we denote the d-dimensional square in $\mathbb{Z}^d$ by $S(N)$, i.e.,

$$S(N) = \{\mathbf{m} = (m_1, m_2, \ldots, m_d) \in \mathbb{Z}^d : -N \le m_i \le N, i = 1, \ldots, d\}.$$

On $\mathbb{Z}^d$ the following version of the Wiener–Wintner theorem can be obtained [12].

Theorem 2. *Let $\mu \in A(\mathbb{T}^d)$ be positive on $\mathbb{T}^d$. There is a constructible function $x : \mathbb{Z}^d \to \mathbb{C}$ such that*

$$\begin{aligned} \forall \mathbf{k} \in \mathbb{Z}^d, A_x[\mathbf{k}] &= \lim_{N\to\infty} \frac{1}{(2N+1)^d} \sum_{\mathbf{m}\in S(N)} x[\mathbf{k}+\mathbf{m}]\overline{x[\mathbf{m}]} \\ &= \check{\mu}[\mathbf{k}]. \end{aligned} \tag{8.1}$$

Although the Wiener–Wintner theorem gives the construction of the function x, it does not ensure boundedness of x. In fact, x need not be an element of $\ell^\infty(\mathbb{Z})$ [19]. Our desire is to construct sequences x that have constant amplitude.

Let $\lambda \in (0, 1)$ have the binary expansion $0.\alpha_1\alpha_2\alpha_3\cdots$, where each α_i is either 0 or 1. It has been shown in [34, 35] that if we consider the Lebesgue measure on $(0, 1)$ and if we define the unimodular (in fact, ± 1-valued) function y by

$$y[k] = \begin{cases} 2\alpha_{2n+1} - 1 & \text{if } k = n+1,\ n \in \mathbb{N}\cup\{0\}, \\ 2\alpha_{2n} - 1 & \text{if } k = 1-n,\ n \in \mathbb{N}, \end{cases} \tag{8.2}$$

then, for *almost all* values of λ, the autocorrelation of y, A_y, is

$$A_y[k] = \begin{cases} 0 & \text{if } k \neq 0, \\ 1 & \text{if } k = 0. \end{cases} \tag{8.3}$$

Thus, A_y is the inverse Fourier transform of $F \equiv 1$ on $\mathbb{T}$. Here, Lebesgue measure on (0, 1) is the probability measure ([12], p 77).

The expression (8.3) defines a sequence y having *perfect autocorrelation*. An explicit or deterministic construction of such a unimodular sequence on $\mathbb{Z}$ is given in [34], where the sequence consists of ± 1 s. Inspired by that we propose a different class of deterministic unimodular sequences with perfect autocorrelation that are constructed from real Hadamard matrices. In fact, an extensive generalization of such constructions can be found in [8].

Definition 2. (a) Let $\mathbb{Z}/N\mathbb{Z}$ be the finite group $\{0,1,\ldots,N-1\}$ with addition modulo N. We say that $x : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ is a *constant amplitude zero autocorrelation* (CAZAC) sequence if $|x[k]| = 1$ for each $k \in \mathbb{Z}/N\mathbb{Z}$ and if

$$\forall k = 1,\ldots,N-1, \quad \frac{1}{N}\sum_{m=0}^{N-1} x[m+k]\overline{x[m]} = 0.$$

(b) Given $x : \mathbb{Z} \to \mathbb{C}$. The sequence x is a CAZAC sequence on $\mathbb{Z}$ if $|x[k]| = 1$ for each $k \in \mathbb{Z}$ and if $A_x[k] = 0$ for each $k \in \mathbb{Z}\backslash\{0\}$.

8.1.3 Outline

In Sect. 8.2.1, we review properties and problems related to Hadamard matrices. This serves as background for Sect. 8.2.2, where we establish the relation between CAZAC sequences on $\mathbb{Z}/N\mathbb{Z}$, Hadamard matrices, and the discrete Fourier transform. Then, in Sect. 8.2.3, we construct CAZAC sequences on $\mathbb{Z}$ by means of Hadamard matrices. Sect. 8.3 is devoted to extending the material of Sect. 8.2 in the following way. In Sect. 8.3.1 we construct unimodular functions on $\mathbb{Z}$ whose autocorrelations are triangles, and we view this as a generalization of the construction of CAZACs on $\mathbb{Z}$. It is natural to think of such triangles as building blocks of the functions on $\mathbb{Z}$. As such, Sect. 8.3.2 is devoted to the formula $A_{x+y} = A_x + A_y$, and we prove its validity a.e.

8.2 Hadamard Matrices and CAZAC Sequences

8.2.1 Hadamard Matrices

Definition 3. A real *Hadamard matrix* is a square matrix whose entries are either $+1$ or -1 and whose rows are mutually orthogonal.

Let H be a Hadamard matrix of order n. Then, the matrix

$$\begin{bmatrix} H & H \\ H & -H \end{bmatrix}$$

is a Hadamard matrix of order $2n$. This observation can be applied repeatedly (as Kronecker products) to obtain the following sequence of Hadamard matrices:

$$H_1 = \begin{bmatrix} 1 \end{bmatrix},$$

$$H_2 = \begin{bmatrix} H_1 & H_1 \\ H_1 & -H_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

$$H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}, \cdots.$$

Thus,

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}$$

$$= \begin{bmatrix} H_{2^{k-2}} & H_{2^{k-2}} & H_{2^{k-2}} & H_{2^{k-2}} \\ H_{2^{k-2}} & -H_{2^{k-2}} & H_{2^{k-2}} & -H_{2^{k-2}} \\ H_{2^{k-2}} & H_{2^{k-2}} & -H_{2^{k-2}} & -H_{2^{k-2}} \\ H_{2^{k-2}} & H_{2^{k-2}} & -H_{2^{k-2}} & H_{2^{k-2}} \end{bmatrix}. \tag{8.4}$$

This method of constructing Hadamard matrices is due to Sylvester (1867) [29]. In this manner, he constructed Hadamard matrices of order 2^k for every nonnegative integer k.

The most important open question in the theory of Hadamard matrices is that of existence. The *Hadamard conjecture* asserts that a Hadamard matrix of order $4N$ exists for every positive integer N [16]. Hadamard matrices of orders 12 and 20 were constructed by Hadamard in 1893 [14]. He also proved that if U is a unimodular matrix of order N, then $|\det(U)| \leq N^{N/2}$, with equality in the case U is real if and only if U is Hadamard [14]. In 1933, Paley discovered a construction that produces a Hadamard matrix of order $q+1$ when q is any prime power that is congruent to 3 modulo 4, and that produces a Hadamard matrix of order $2(q+1)$ when q is a prime power that is congruent to 1 modulo 4 [24]. His method uses finite fields. The Hadamard conjecture should probably be attributed to Paley. The smallest order that cannot be constructed by a combination of Sylvester's and Paley's methods is 92. A Hadamard matrix of this order was found using a computer by Baumert, Golomb, and Hall in 1962. They used a construction, due to Williamson, that has yielded many additional orders. In 2004, Hadi Kharaghani and Behruz Tayfeh-Rezaie announced that they constructed a Hadamard matrix of order 428. As a result, the smallest order for which no Hadamard matrix is presently known is 668.

Hadamard matrices are closely connected with Walsh functions [2, 26]. The Walsh functions, constructed by Walsh [33], are an orthonormal basis for $L^2(\mathbb{T})$. Every Walsh function is constant over each of a finite number of subintervals

of $(0,1)$. A set of Walsh functions written down in appropriate order as rows of a matrix will give a Hadamard matrix of order 2^N as obtained by Sylvester's method. The Walsh functions defined on $\mathbb{R}$ correspond to the wavelet packets associated with the Haar multiresolution analysis.

8.2.2 CAZACs and Circulant Hadamard Matrices

An $N \times N$ matrix A of the form

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_N \\ a_N & a_1 & a_2 & \cdots & a_{N-1} \\ a_{N-1} & a_N & a_1 & \cdots & a_{N-2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_2 & a_3 & \cdots & a_N & a_1 \end{bmatrix}$$

is called a circulant matrix [17]. Each row is just the previous row cycled forward by one step, so that the entries in each row are just a cyclic permutation of those in the first. There is a characterization of CAZAC sequences in terms of circulant Hadamard matrices with complex entries, see Theorem 4, e.g., [10]. For any finite sequence $x = (x[0], x[1], \ldots, x[N-1])$ of N complex numbers $(N \geq 1)$, its *normalized discrete Fourier transform* $\hat{x} = (\hat{x}[0], \hat{x}[1], \ldots, \hat{x}[N-1])$ is defined by

$$\hat{x}[j] = N^{-\frac{1}{2}} \sum_{k=0}^{N-1} x[k] \mathrm{e}^{-2\pi i k j/N} \quad (j = 0, 1, \ldots, N-1).$$

By Parseval's relation,

$$\sum_{k=0}^{N-1} |x[k]|^2 = \sum_{j=0}^{N-1} |\hat{x}[j]|^2.$$

It is easy to see that x is CAZAC if and only if x and $\hat{x}$ are unimodular (Corollary 1). This fact is a consequence of the following result.

Theorem 3. *Let $x : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ be the sequence $x = (x[0], x[1], \ldots, x[N-1])$. The condition,*

$$\forall m = 1, \ldots, N-1, \quad \frac{1}{N} \sum_{k=0}^{N-1} x[m+k]\overline{x[k]} = 0, \tag{8.5}$$

is valid if and only if there is a constant c such that $|\hat{x}| = c$ on $\mathbb{Z}/N\mathbb{Z}$.

Proof. (i) Suppose that $|\hat{x}| = c$ on $\mathbb{Z}/N\mathbb{Z}$. Then, for each $j \in \mathbb{Z}/N\mathbb{Z}$,

$$|\hat{x}[j]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |x[k]|^2 + \frac{1}{N} \sum_{k \neq \ell} x[k]\overline{x[\ell]} \mathrm{e}^{-2\pi i (k-\ell) j/N},$$

and so

$$N|\hat{x}[j]|^2 = \sum_{k=0}^{N-1} |x[k]|^2 + \sum_{k\neq\ell} x[k]\overline{x[\ell]}\mathrm{e}^{-2\pi i(k-\ell)j/N}.$$

Thus, by hypothesis, we have $N|\hat{x}[j]|^2 = \sum_{n=0}^{N-1} |\hat{x}[n]|^2$ $(= Nc^2)$, and so

$$\sum_{n=0}^{N-1} |\hat{x}[n]|^2 = \sum_{k=0}^{N-1} |x[k]|^2 + \sum_{k\neq\ell} x[k]\overline{x[\ell]}\mathrm{e}^{-2\pi i(k-\ell)j/N}.$$

Hence, by Parseval's identity, we have

$$\forall j \in \mathbb{Z}/N\mathbb{Z}, \quad \sum_{k\neq\ell} x[k]\overline{x[\ell]}\mathrm{e}^{-2\pi i(k-\ell)j/N} = 0. \tag{8.6}$$

Fix $k \in \{0,1,\ldots,N-1\}$ and let $m = k-\ell \pmod N$. Then, (8.6) becomes

$$\sum_{m=1}^{N-1}\sum_{\ell=0}^{N-1} x[\ell+m]\overline{x[\ell]}\mathrm{e}^{-2\pi imj/N} = 0. \tag{8.7}$$

In particular, there are N^2-N terms in the sum of (8.6) since we exclude the diagonal of an $N\times N$ array. For compatibility, for each m, there are N terms in (8.7), and since there are $N-1$ values of m, we see that there are N^2-N terms in the sum of (8.7). Now let $f[m] = \sum_{\ell=0}^{N-1} x[\ell+m]\overline{x[\ell]}$. Then (8.7) becomes

$$\forall j \in \mathbb{Z}/N\mathbb{Z}, \quad \sum_{m=1}^{N-1} f[m]\mathrm{e}^{-2\pi imj/N} = 0. \tag{8.8}$$

Multiplying both sides of (8.8) by $\mathrm{e}^{2\pi ikj/N}$, for a fixed $k \in \{0,1,\ldots,N-1\}$, we have

$$\forall j \in \mathbb{Z}/N\mathbb{Z}, \quad \sum_{m=1}^{N-1} f[m]\mathrm{e}^{-2\pi i(m-k)j/N} = 0,$$

and so

$$\sum_{m=1}^{N-1} f[m]\left(\sum_{j=0}^{N-1} \mathrm{e}^{-2\pi i(m-k)j/N}\right) = 0 \tag{8.9}$$

for every fixed $k \in \{0,1,\ldots,N-1\}$. Since

$$\sum_{j=0}^{N-1} \mathrm{e}^{-2\pi i(m-k)j/N} = \begin{cases} N, & k=m, \\ \frac{\mathrm{e}^{-2\pi i(m-k)}-1}{\mathrm{e}^{-2\pi i(m-k)/N}-1} = 0, & k\neq m, \end{cases}$$

and since $m \in \{1,\dots,N-1\}$, (8.9) allows us to assert that $f[m] = 0$ for each $m \in \{1,\dots,N-1\}$. In fact, for any fixed $k \in \{1,\dots,N-1\}$, the left side of (8.9) becomes $Nf[k]$, and so $f[k] = 0$ by the right side of (8.9).

(ii) The converse is proved by retracing the steps of (i).

□

Corollary 1. *Let $x : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ be the unimodular sequences $x = (x[0], x[1], \dots, x[N-1])$. The sequence x is a CAZAC sequence if and only if $\hat{x}$ is a unimodular sequence.*

Proof. If x is a CAZAC sequence, then (8.5) is valid, and so $|\hat{x}| = c$ by Theorem 3. By Parseval's relation,

$$\sum_{j=0}^{N-1} |\hat{x}[j]|^2 = \sum_{k=0}^{N-1} |x[k]|^2$$

$$\text{or,} \quad Nc^2 = N,$$

where in the last step we use the fact that x is unimodular. Thus, the constant c is equal to 1 and $\hat{x}$ is a unimodular sequence. The converse follows by retracing this proof. □

Definition 4. A complex Hadamard matrix is a square matrix whose entries are unimodular and whose rows are mutually orthogonal.

We have the following characterization of CAZAC sequences in terms of circulant Hadamard matrices with complex entries.

Theorem 4. *Given a sequence $x : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, and let H_x be a circulant matrix with first row $x = (x[0], x[1], \dots, x[N-1])$. Then x is a CAZAC sequence if and only if H_x is a Hadamard matrix.*

Proof.

$$H_x = \begin{bmatrix} x[0] & x[1] & \cdots & x[N-1] \\ x[N-1] & x[0] & \cdots & x[N-2] \\ \vdots & \vdots & \cdots & \vdots \\ x[1] & x[2] & \cdots & x[0] \end{bmatrix}, \quad H_x^* = \begin{bmatrix} \overline{x[0]} & \overline{x[N-1]} & \cdots & \overline{x[1]} \\ \overline{x[1]} & \overline{x[0]} & \cdots & \overline{x[2]} \\ \vdots & \vdots & \cdots & \vdots \\ \overline{x[N-1]} & \overline{x[N-2]} & \cdots & \overline{x[0]} \end{bmatrix}.$$

(i) Assume that H_x is a complex Hadamard matrix. Hence, all of the entries of H_x are unimodular and

$$H_x H_x^* = NI_N, \tag{8.10}$$

where I_N is the $N \times N$ identity matrix. As a consequence of (8.10) one has for $m = 1,\dots,N-1$,

$$\sum_{\ell=0}^{N-1} x[\ell+m]\overline{x[\ell]} = 0,$$

which means that x has zero autocorrelation and is thus a CAZAC.

(ii) Conversely, suppose that x is a CAZAC. We want to show that H_x is a Hadamard matrix. We already know that all the entries of H_x are unimodular since x is unimodular and the entries of H_x are the elements of x. We want to show that $H_x H_x^* = NI_N$. Due to unimodularity,

$$\sum_{\ell=0}^{N-1} |x[\ell]|^2 = N, \tag{8.11}$$

and so the diagonal entries of $H_x H_x^*$ equal N as required. Since x is CAZAC,

$$\sum_{\ell=0}^{N-1} x[\ell+m]\overline{x[\ell]} = 0$$

for $m \neq 0$, which means that every off-diagonal entry of $H_x H_x^*$ equals zero, and this together with (8.11) implies that $H_x H_x^*$ is a Hadamard matrix. □

Due to this characterization of CAZACs there is a basic relation between CAZACs and *finite unit normed tight frames* (FUNTFs) in $\mathbb{C}^d$. We shall say that $x : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}^d$ is a CAZAC sequence in $\mathbb{C}^d$ if each $\|x[k]\| = 1$ and

$$\forall k = 1,\ldots,N-1, \quad \frac{1}{N}\sum_{m=0}^{N-1}\langle x[m+k],x[m]\rangle = 0.$$

Each $x[m] = (x_1[m],\ldots,x_d[m])$, where $x_j[m] \in \mathbb{C}$, $m \in \mathbb{Z}/N\mathbb{Z}$, and $j = 1,\ldots,d$, and the inner product is

$$\langle x[k],x[m]\rangle = \sum_{j=1}^{d} x_j[k]\overline{x_j[m]}.$$

The norm of each $x[k]$ is then $\|x[k]\| = \langle x[k],x[k]\rangle^{1/2}$. For fundamentals on frame theory we refer to [11] or [13]. The following has been shown in [9].

Theorem 5. *Let $x = \{x[n]\}_{n=1}^N$ be a CAZAC sequence in $\mathbb{C}$. Define*

$$\forall k = 1,\ldots,N, \quad v(k) = \frac{1}{\sqrt{d}}(x[k],x[k+1],\ldots,x[k+d-1]).$$

Then $v = \{v(k)\}_{k=1}^N$ is a CAZAC sequence in $\mathbb{C}^d$ and $\{v(k)\}_{k=1}^N$ is a FUNTF for $\mathbb{C}^d$ with frame constant $\frac{N}{d}$.

8.2.3 CAZACs and Hadamard Sequences

In this section we construct infinite CAZAC sequences, i.e., CAZAC sequences on $\mathbb{Z}$, from real Hadamard matrices. Two different constructions are given. For the proofs of Theorems 6 and 7 we refer the readers to [8].

Example 1. To construct a unimodular sequence x, let H_1 be repeated once ($2^0 = 1$), H_2 be repeated twice (2^1), H_4 be repeated 2^2 times, H_8 be repeated 2^3 times, and, in general, let H_{2^n} be repeated 2^n times. For the positive integers, let x take values row by row from the elements of the sequence of matrices

$$H_1, H_2, H_2, H_4, H_4, H_4, H_4, H_8, \ldots. \tag{8.12}$$

Set $x[0] = 1$, and for any $k \in \mathbb{N}$, define $x[-k] = x[k]$. The sequence x is called the *exponential Hadamard sequence.*

Theorem 6. *Let x be the exponential Hadamard sequence. Then,*

$$A_x[k] = \begin{cases} 1 & \text{if} \quad k = 0, \\ 0 & \text{if} \quad k \neq 0. \end{cases}$$

Instead of having the Hadamard matrices repeat exponentially as described in Example 1, we can construct unimodular sequences, whose autocorrelations vanish everywhere except at the origin, by letting the Hadamard matrices repeat linearly.

Example 2. To construct the linear Hadamard sequence x, let H_1 be repeated zero times, H_2 be repeated once, H_4 be repeated twice, H_8 be repeated thrice, and, in general, let H_{2^n} be repeated n times. For the positive integers, let x take values row by row from the elements of the sequence of matrices

$$H_2, H_4, H_4, H_8, H_8, H_8, H_{16}, H_{16}, H_{16}, H_{16}, H_{32}, \ldots.$$

Set $x[0] = 1$, and, for any $k \in \mathbb{N}$, define $x[-k] = x[k]$. The sequence x is called the *linear Hadamard sequence.*

The proof of the following result is similar to that of Theorem 6.

Theorem 7. *Let x be the linear Hadamard sequence. Then,*

$$A_x[k] = \begin{cases} 1 \text{ if } k = 0, \\ 0 \text{ if } k \neq 0. \end{cases}$$

These two constructions are more general than they appear. For example, instead of $H_1 = [1]$, one could start with $H_1 = [-1]$ and obtain the following sequence of Hadamard matrices:

$$H_1 = [-1],$$
$$H_2 = \begin{bmatrix} H_1 & H_1 \\ H_1 & -H_1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix},$$
$$H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix} = \begin{bmatrix} -1 & -1 & -1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}, \ldots.$$

Using this sequence of Hadamard matrices in Example 1 or Example 2 would give a different sequence x but one which would still have perfect autocorrelation.

Example 3. In practice, for applications, we cannot use an infinite sequence and we would like to estimate the number of elements of the sequences in Examples 1 and 2 that can be used to make the corresponding autocorrelation reasonably small. In other words, we would like to solve the following problem: given $\varepsilon > 0$, find $N \in \mathbb{N}$ such that

$$\forall k \in \mathbb{Z}, \quad \left| \frac{1}{N} \sum_{m=1}^{N} x[m+k]x[m] \right| < \varepsilon.$$

Let x be the exponential Hadamard sequence of Example 1. Let $\varepsilon > 0$ and $K \in \mathbb{N}$. The smallest N such that

$$\forall\, 0 < |k| \leq K, \quad \left| \frac{1}{N} \sum_{m=1}^{N} x[m+k]x[m] \right| < \varepsilon$$

satisfies the inequality

$$\frac{1}{N} \frac{8^{\lceil \log_2(K) \rceil + 1} - 1}{7} + 7 \frac{1}{2^{M+1}} < \varepsilon, \tag{8.13}$$

where M is a function of N. For more information about the relationship between M and N we refer to [8, 12].

(8.14) gives the values of N obtained via (8.13) for $K = 16$ and several values of ε.

ε	1	0.5	0.25	0.1
K	16	16	16	16
M	14	15	16	17
N	$O(8^{15})$	$O(8^{16})$	$O(8^{17})$	$O(8^{18})$

(8.14)

The actual error estimate for the exponential Hadamard sequence is illustrated in Fig. 8.1. This estimate is significantly better than that obtained in (8.13). The

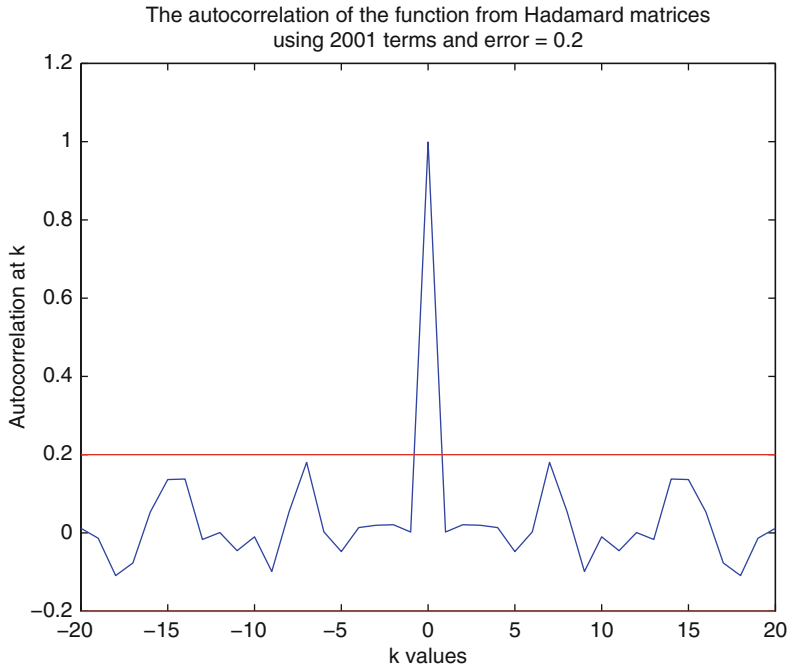

Fig. 8.1 Error estimates of the exponential Hadamard sequence; $\varepsilon = 0.2$

disparity is a consequence of the difficult counting problems inherent in dealing with Hadamard matrices. However, Fig. 8.1 does imply a valid use of these sequences in applications.

Next, let x be the linear Hadamard sequence of Example 2. Given $\varepsilon > 0$ and $K \in \mathbb{N}$. The smallest N such that

$$\forall\, 0 < |k| \leq K, \quad \left| \frac{1}{N} \sum_{m=1}^{N} x[m+k]x[m] \right| < \varepsilon$$

satisfies the inequality

$$\frac{(3\lceil \log_2(K) \rceil - 1)4^{\lceil \log_2(K) \rceil + 1} + 4 + 9 \cdot 4^{M+1}}{3M4^{M+1} - 4(4^M - 1)} < \varepsilon, \tag{8.15}$$

where M is a function of N.

(8.16) gives the values of N obtained from (8.15) for $K = 16$ and several values of ε. Once again, Fig. 8.2 illustrates that the actual error estimates are much better than that obtained in (8.15).

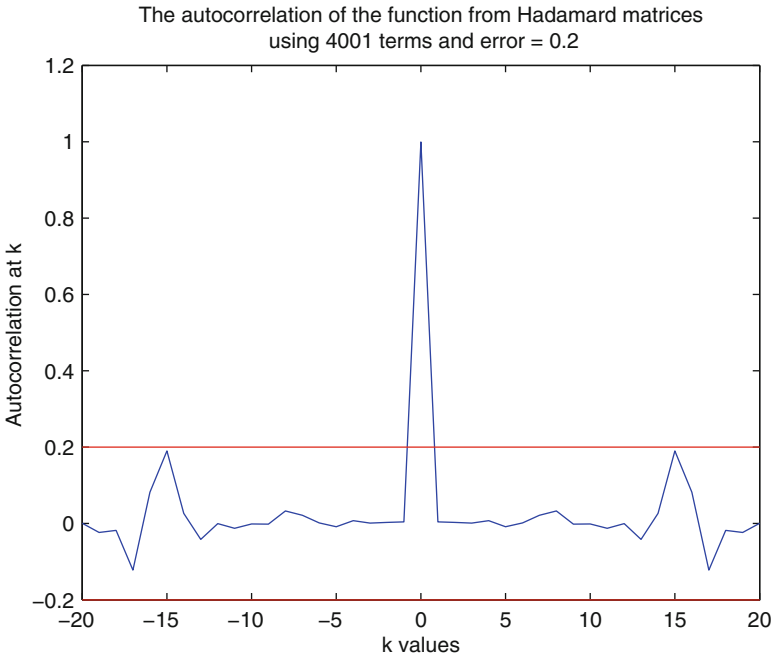

Fig. 8.2 Error estimates of the linear Hadamard sequence; $\varepsilon = 0.2$

ε	1	0.5	0.25	0.1
K	16	16	16	16
M	5	7	13	31
N	35048	735464	5.16×10^9	7.97×10^{20}

(8.16)

8.3 Autocorrelations as Sums of Triangles

8.3.1 The Construction of Sequences with Triangular Autocorrelation

In this section a generalization of (8.3), the autocorrelation function of the sequence given by (8.2), and of those constructed from Hadamard matrices in Sect. 8.2.3 and also in [8] is given.

Theorem 8. *Given $M \in \mathbb{N}$ and $K > 0$. Let $A : \mathbb{Z} \to \mathbb{R}$ be defined by*

$$A[k] = \begin{cases} K\left(1 - \frac{|k|}{M}\right) & \text{if } 0 \le |k| \le M, \\ 0 & \text{otherwise.} \end{cases} \tag{8.17}$$

Then there exists a constructible sequence $x : \mathbb{Z} \to \mathbb{R}$ with constant amplitude $\sqrt{K}$ whose autocorrelation, A_x, is A.

Proof. (i) As mentioned in Sect. 8.1.2 one can deterministically construct a unimodular sequence y on $\mathbb{Z}$ whose autocorrelation is

$$A_y[k] = \begin{cases} 0 & \text{if } k \neq 0, \\ 1 & \text{if } k = 0, \end{cases} \tag{8.18}$$

and we use (8.18) at the end of the proof. Wieners construction [34] of y is as follows. On the positive integers, let y take values in the following order:
$[1,-1]$ (this row has 1×2^1 elements and is repeated $2^0 = 1$ time);
$[1,1;1,-1;-1,1;-1,-1]$ (this row has 2×2^2 elements and is repeated $2^1 = 2$ times);
$[1,1,1;1,1,-1;1,-1,1;1,-1,-1;-1,1,1;-1,1,-1;$
$-1,-1,1;-1,-1,-1]$ (this row has 3×2^3 elements and is repeated $2^2 = 4$ times); etc. Thus, $y[1] = 1, y[2] = -1, y[3] = 1, y[4] = 1, \ldots$. In addition, let $y[0] = 1$, and, for $k \in \mathbb{N}$, let $y[-k] = y[k]$.

(ii) We define the function $x : \mathbb{Z} \to \mathbb{C}$ by $x[k] = \sqrt{K} y[\lceil \frac{k}{M} \rceil]$, where $\lceil . \rceil$ denotes the next largest integer. Note that $|x| = \sqrt{K}$.

We show that the autocorrelation A_x of x is A as defined in (8.17). Since x is a real sequence, the autocorrelation function is even, and so it is enough to prove the result for $k > 0$. Let $0 \le Mp \le k \le M(p+1)$ for some $p \in \mathbb{N} \cup \{0\}$. For any given integer N, let n_N be the smallest integer such that $N < M(n_N + 1)$. Then we have

$$\begin{aligned} A_x[k] &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} x[m+k]\overline{x[m]} \\ &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-Mn_N}^{Mn_N} x[k+m]x[m] \\ &\quad + \lim_{N\to\infty} \frac{1}{2N+1} \sum_{Mn_N < |m| \le N} x[m+k]x[m] \\ &= \lim_{N\to\infty} (S_{1,N}(k) + S_{2,N}(k)) = S_1(k) + S_2(k). \end{aligned} \tag{8.19}$$

First, we calculate bounds on $S_{2,N}(k)$.

$$
\begin{aligned}
|S_{2,N}(k)| &= \left| \frac{1}{2N+1} \sum_{Mn_N < |m| \le N} x[m+k]x[m] \right| \\
&\le \frac{1}{2N+1} \sum_{Mn_N < |m| \le N} |x[m+k]x[m]| \\
&= \frac{K}{2N+1} \sum_{Mn_N < |m| \le N} 1 = \frac{2K(N - Mn_N)}{2N+1}.
\end{aligned}
$$

We know from the definition of n_N that $N - Mn_N < M$. Therefore, $S_2(k) = 0$. Consequently, $A_x[k] = \lim_{N\to\infty} S_{1,N}(k) = S_1(k)$. Next, we write

$$
\begin{aligned}
S_1(k) &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-Mn_N}^{Mn_N} x[k+m]x[m] \\
&= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-n_N}^{n_N-1} \sum_{m=Mn+1}^{M(n+1)} x[k+m]x[m] \\
&\quad + \lim_{N\to\infty} \frac{1}{2N+1} x[-n_N+k]x[-n_N]. \qquad (8.20)
\end{aligned}
$$

Since x has the same value $\sqrt{K}y[n+1]$ for all the integers $m \in [Mn+1, M(n+1)]$, one can replace the $x[m]$ in the first term of the right side of (8.20) by $\sqrt{K}y[n+1]$. Since the second term of the right side of (8.20) is 0 this implies

$$
\begin{aligned}
S_1(k) &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-n_N}^{n_N-1} \sum_{m=Mn+1}^{M(n+1)} x[m+k]\sqrt{K}y[n+1] \\
&= \lim_{N\to\infty} \frac{1}{2N+1} \left(\sum_{n=-n_N}^{n_N-1} \sum_{m=Mn+1}^{Mn+M(p+1)-k} x[m+k]\sqrt{K}y[n+1] \right. \\
&\quad \left. + \sum_{n=-n_N}^{n_N-1} \sum_{Mn+M(p+1)-k+1}^{M(n+1)} x[m+k]\sqrt{K}y[n+1] \right) \\
&= \lim_{N\to\infty} \frac{K}{2N+1} \sum_{n=-n_N}^{n_N-1} \left(\sum_{m=Mn+1}^{Mn+M(p+1)-k} y[n+p+1]y[n+1] \right. \\
&\quad \left. + \sum_{m=Mn+M(p+1)-k+1}^{M(n+1)} y[n+p+2]y[n+1] \right) \\
&= \lim_{N\to\infty} \frac{K}{2N+1} \sum_{n=-n_N}^{n_N-1} ((M(p+1)-k)y[n+p+1]y[n+1]
\end{aligned}
$$

$$+ (k - Mp)y[n+p+2]y[n+1])$$

$$= \lim_{N\to\infty} \frac{M(p+1)-k}{2N+1} \frac{2n_N K}{2n_N} \sum_{n=-n_N}^{n_N-1} y[n+p+1]y[n+1] +$$

$$+ \lim_{N\to\infty} \frac{(k-Mp)}{2N+1} \frac{2n_N K}{2n_N} \sum_{n=-n_N}^{n_N-1} y[n+p+2]y[n+1].$$

Since $n_N \to \infty$ as $N \to \infty$, we have

$$\begin{aligned} \lim_{N\to\infty} S_{1,N}(k) &= \lim_{N\to\infty} \frac{M(p+1)-k}{2N+1}\, 2n_N K A_y[p] \\ &\quad + \lim_{N\to\infty} \frac{k-Mp}{2N+1}\, 2n_N K A_y[p+1] \\ &= \lim_{N\to\infty} \left(p+1-\frac{k}{M}\right) \frac{2n_N M}{2N+1} K A_y[p] \\ &\quad + \lim_{N\to\infty} \left(\frac{k}{M}-p\right) \frac{2n_N M}{2N+1} K A_y[p+1]. \end{aligned} \tag{8.21}$$

Note that

$$\lim_{N\to\infty} \frac{2n_N M}{2N+1} = 1. \tag{8.22}$$

In fact, from the choice of n_N, we have $Mn_N \le N < M(n_N+1)$ so that $2Mn_N + 1 \le 2N+1 < 2M(n_N+1)+1$, and hence

$$\frac{2Mn_N}{2M(n_N+1)+1} < \frac{2Mn_N}{2N+1} \le \frac{2Mn_N}{2Mn_N+1}.$$

n_N goes to infinity as N goes to infinity and so taking limits throughout as N goes to infinity, we obtain (8.22).

Substituting (8.22) in (8.21) and using the fact that $S_2(k) = 0$, we obtain from (8.19) that

$$A_x[k] = S_1(k) = K\left(p+1-\frac{k}{M}\right) A_y[p] + K\left(\frac{k}{M}-p\right) A_y[p+1].$$

If $0 \le k \le M$, then $p = 0$. For every other range of k, p is nonzero. Using the values of $A_y[p]$ as given by (8.18) and the fact that A_x is an even function one obtains (8.17). □

Remark 1. The function A defined in Theorem 8 is the triangle $\triangle_{K,M}(t) = K\max(1-\frac{|t|}{M}, 0)$ on $\mathbb{R}$ with height K and base length $2M$ restricted to the integers.

The Fourier transform of $\triangle_{K,M}(t)$ is $KM\left(\frac{\sin \pi M\gamma}{\pi M\gamma}\right)^2$. Thus, in Theorem 8, we have constructed a sequence x of constant amplitude whose autocorrelation is the inverse Fourier transform of the dilated Fejér function $K\omega_{2\pi M}$.

8.3.2 *The Additive Property of Triangular Autocorrelation a.e.*

As mentioned in Sect. 8.1.2, and repeated in the proof of Theorem 8, it has been shown in [34,35] that if $\lambda \in (0,1)$ has binary expansion $0.\alpha_1\alpha_2\alpha_3\cdots$, if we consider the Lebesgue measure on $(0,1)$, and if we define the unimodular (in fact, ± 1-valued) function y by

$$y[k] = \begin{cases} 2\alpha_{2n+1} - 1 & \text{if } k = n+1, n \in \mathbb{N} \cup \{0\}, \\ 2\alpha_{2n} - 1 & \text{if } k = 1-n, n \in \mathbb{N}, \end{cases}$$

then, for *almost all* values of λ, the autocorrelation of y, A_y, is

$$A_y[k] = \begin{cases} 0 & \text{if } k \neq 0, \\ 1 & \text{if } k = 0. \end{cases}$$

In Theorem 8 it was shown that given $M \in \mathbb{N}$ this y can be used to construct x such that x has constant amplitude and

$$A_x[k] = \begin{cases} 1 - \frac{|k|}{M}, & \text{if } 0 \leq |k| \leq M, \\ 0, & \text{otherwise.} \end{cases}$$

In this case, x is unimodular. We shall now show that the autocorrelation of the sum of two such functions is the sum of the respective autocorrelations for almost all x.

We begin with the following calculation.

Example 4. Let X be the set of unimodular functions $x : \mathbb{Z} \to \mathbb{C}$ for which there exists a positive integer M with the property,

$$A_x[k] = \begin{cases} 1 - \frac{|k|}{M}, & \text{if } 0 \leq |k| \leq M, \\ 0, & \text{otherwise.} \end{cases}$$

For given $M \in \mathbb{N}$, let Ω be the set of all possibilities of any $2M$ consecutive values of $x \in X$. Then $\text{card}(\Omega) = 2^{2M}$. Let E be the subset of Ω such that given ε, the sum of the $2M$ consecutive values of x exceeds $M\varepsilon$ in absolute value. Among the $2M$ values, suppose that there are $(M-j)$ $+1$s and $(M+j)$ -1s, where $-M \leq j \leq M$. So the absolute value of the sum of $2M$ consecutive values would be $|M+j-(M-j)| = 2|j|$. The sum of these values exceeds $M\varepsilon$ in absolute value if $[M\varepsilon] \leq 2|j| \leq 2M$.

The number of ways of having $(M-j)$ $+1$s and $(M+j)$ -1s is $\binom{2M}{M-j} = \binom{2M}{M+j}$. The total number of possible values for which the sum exceeds $M\varepsilon$ is

$$\text{card}(E) = \sum_{|j|=[\frac{M\varepsilon}{2}]}^{M} \binom{2M}{M-j} = \sum_{j=[\frac{M\varepsilon}{2}]}^{M} \binom{2M}{M-j} + \sum_{j=[\frac{M\varepsilon}{2}]}^{M} \binom{2M}{M+j} = 2\sum_{j=[\frac{M\varepsilon}{2}]}^{M} \binom{2M}{M-j}.$$

Consequently,

$$\frac{\text{card}(E)}{\text{card}(\Omega)} = 2^{-2M} 2 \sum_{j=[\frac{M\varepsilon}{2}]}^{M} \binom{2M}{M-j} = 2^{-2M+1} \sum_{j=[\frac{M\varepsilon}{2}]}^{M} \binom{2M}{M-j}.$$

Theorem 9. *(a) Let X be the set of unimodular functions $x : \mathbb{Z} \to \mathbb{C}$ for which there exists a positive integer M with the property,*

$$A_x[k] = \begin{cases} 1 - \frac{|k|}{M}, & \text{if } 0 \le |k| \le M, \\ 0, & \text{otherwise.} \end{cases}$$

Then there is a well-defined finite Borel measure p on X induced from Lebesgue measure [1] *on $(0,1)$, in a manner described in the proof.*

(b) For almost all $x, y \in X$, with respect to p, we have

$$A_{x+y} = A_x + A_y,$$

noting that $x+y$ does not necessarily have constant amplitude and that A_{x+y} is not generally a triangle.

Proof. (a) We know from (8.2) and (8.3) that there is $S_0 \subseteq [0,1]$ defined by the properties: $|S_0| = 1$ and

$$\forall \lambda \in S_0, \quad \exists \mu_\lambda : \mathbb{Z} \to \mathbb{C} \text{ such that } |\mu_\lambda| = 1 \text{ and } A_{\mu_\lambda}[k] = \delta_{0,k} \text{ on } \mathbb{Z}.$$

From Theorem 8 we know that for each $M \in \mathbb{N}$, there is $S_M \subseteq [0,1]$ defined by the properties: $|S_M| = 1$ and

$$\forall \lambda \in S_M, \ \exists \mu_\lambda : \mathbb{Z} \to \mathbb{C} \text{ such that } |\mu_\lambda| = 1 \text{ and } A_{\mu_\lambda}[k] = \max\left(0, 1 - \frac{|k|}{M}\right) \text{ on } \mathbb{Z}.$$

In fact, by the way we defined μ_λ in Theorem 8, we could take $S_M = S_0$. However, we can equally well choose $\{S_M : S_M \subseteq S_0, |S_M| = 1\}$ to be a disjoint collection whose union is S_0. In this case we define the functions $f_M : S_M \to X$,

[1] For the necessary measure theory and definitions of Borel and Lebesgue measure we refer to [7].

$\lambda \mapsto \mu_\lambda$, where $A_{\mu_\lambda}[k] = \max\left(0, 1 - \frac{|k|}{M}\right)$ on $\mathbb{Z}$, and $f : S_0 \to X$, $\lambda \mapsto f_M(\lambda)$ when $\lambda \in S_M$. In this way we use f to define a compact topology on X induced from $S_0 \subseteq [0,1]$, and to define a bounded Borel measure p on X induced from Lebesgue measure on $[0,1]$.

We provide the technical properties of p in part (b) of the proof.

(b) We have already seen the construction of such x and y in Theorem 8. Formally,

$$\begin{aligned}
A_{x+y}[k] &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} (x+y)[m+k]\overline{(x+y)[m]} \\
&= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} (x[m+k]+y[m+k])(\overline{x[m]}+\overline{y[m]}) \\
&= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} x[m+k]\overline{x[m]} + \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} y[m+k]\overline{y[m]} \\
&\quad + \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} x[m+k]\overline{y[m]} + \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} y[m+k]\overline{x[m]} \\
&= A_x(k) + A_y(k) + \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} x[m+k]\overline{y[m]} \\
&\quad + \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} y[m+k]\overline{x[m]}. \qquad (8.23)
\end{aligned}$$

Let us denote the last two terms on the right side of (8.23) by S_3 and S_4, respectively. We want to show that $S_3 = 0$ and $S_4 = 0$.

$$S_3 = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} x[m+k]\overline{y[m]}. \qquad (8.24)$$

Without loss of generality we take y to be real-valued, and so (8.24) becomes

$$S_3 = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} x[m+k]y[m]. \qquad (8.25)$$

Suppose that

$$A_x[k] = \begin{cases} 1 - \frac{|k|}{M_1}, & \text{if } 0 \le |k| \le M_1, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$A_y[k] = \begin{cases} 1 - \frac{|k|}{M_2}, & \text{if } 0 \le |k| \le M_2, \\ 0, & \text{otherwise.} \end{cases}$$

Let P_N be the largest integer so that

$$M_2 P_N \le N \le M_2(P_N + 1). \tag{8.26}$$

Then S_3 can be written as

$$\begin{aligned} S_3 = & \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{-M_2P_N-1} x[m+k]y[m] \\ & + \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=M_2P_N+1}^{N} x[m+k]y[m] \\ & + \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-M_2P_N}^{M_2P_N} x[m+k]y[m]. \end{aligned} \tag{8.27}$$

Let us denote the first two terms of (8.27) by s_1 and s_2, respectively. Now,

$$|s_1| \le \sum_{m=-N}^{-M_2P_N-1} 1 = N - M_2P_N$$

and

$$|s_2| \le \sum_{m=M_2P_N+1}^{N} 1 = N - M_2P_N.$$

From (8.26),

$$N - M_2P_N \le M_2(P_N + 1) - M_2P_N = M_2$$

which means $|s_1| \le M_2$ and $|s_2| \le M_2$. Therefore,

$$\lim_{N\to\infty} \frac{|s_1|}{2N+1} \le \lim_{N\to\infty} \frac{M_2}{2N+1} = 0$$

and also

$$\lim_{N\to\infty} \frac{|s_2|}{2N+1} \le \lim_{N\to\infty} \frac{M_2}{2N+1} = 0.$$

Thus,

$$\begin{aligned} S_3 &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-M_2P_N}^{M_2P_N} x[m+k]y[m] \qquad (8.28) \\ &= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-P_N}^{P_N-1} \sum_{m=M_2n+1}^{M_2(n+1)} x[m+k]y[m] \end{aligned}$$

$$+\lim_{N\to\infty}\frac{1}{2N+1}x[-M_2P_N+k]y[-M_2P_N]$$

$$=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-P_N}^{P_N-1}\sum_{m=M_2n+1}^{M_2(n+1)}x[m+k]y[M_2(n+1)]. \tag{8.29}$$

The last step (8.29) follows due to the fact that by construction, y is constant and equal to either $+1$ or -1 in the interval $[M_2n+1, M_2(n+1)]$. So $y[M_2(n+1)]$ is either $+1$ or -1. Between (M_2n+1) and $M_2(n+1)$ there are M_2 terms. So there are M_2 values of x. Suppose that of these M_2 values there are j that have the value $+1$ and (M_2-j) that have the value -1. Upon multiplication by $y(M_2(n+1))$ we have either j values that are -1 and (M_2-j) values that are $+1$ or vice versa. In the sum on the right side of (8.29) there are $2P_N$ blocks of length M_2. Let us say that the first block has j_1 terms equal to $+1$ and (M_2-j_1) terms equal to -1, the second block has j_2 terms equal to $+1$ and (M_2-j_2) terms equal to -1, and so on. Together, there are $(j_1+j_2+\cdots+j_{2P_N})$ terms equal to $+1$ and $(M_2-j_1+M_2-j_2+\cdots+M_2-j_{2P_N})=2P_NM_2-(j_1+j_2+\cdots+j_{2P_N})$ terms equal to -1. Let $P_NM_2=M$ and $j_1+j_2+\cdots+j_{2P_N}=M-j$, where $-M\le j\le M$. Note that this M is unrelated to the M that appears in Theorem 8 and part (a) of the statement of this theorem where it indicates the length of the base of a triangle. Then $2P_NM_2-(j_1+j_2+\cdots+j_{2P_N})=2M-(M-j)=M+j$. Thus, out of $2M$ consecutive values of $x[m+k]y[m]$, there are $(M-j)$ values that are $+1$ and $(M+j)$ values that are -1. So the absolute value of the sum of $2P_NM_2=2M$ and consecutive values of $x[m+k]y[m]$ would be $M+j-(M-j)=2|j|$.

Let Ω be the set of all possibilities for the $2M$ consecutive values of $x[m+k]y[m]$. From (8.2), each such x and y corresponds to some $\lambda\in(0,1)$. From Example 4 and the definition of $E\subseteq\Omega$ there, it follows that given ε the measure of the set for which the sum of $2M$ consecutive values exceeds $M\varepsilon$ in absolute value is

$$\frac{\mathrm{card}(E)}{\mathrm{card}(\Omega)}=2^{-2M+1}\sum_{j=[\frac{M\varepsilon}{2}]}^{M}\binom{2M}{M-j}.$$

This can be transported as an explicit, computable property of p.

It can be shown in a manner identical to that in [34] that

$$\lim_{M\to\infty}2^{-2M+1}\sum_{j=[\frac{M\varepsilon}{2}]}^{M}\binom{2M}{M-j}=0.$$

Thus the set of x and y for which there should fail to be an integral value of $M=P_NM_2$ such that from that value on [see (8.28)]

$$\left|\sum_{m=-M}^{M}x[m+k]y[m]\right|\le M\varepsilon+1$$

has measure zero. Therefore,

$$\overline{\lim}_{N\to\infty}\left|\frac{1}{2N+1}\sum_{m=-M}^{M}x[m+k]y[m]\right| \le \frac{M\varepsilon+1}{2N+1} = \frac{P_N M_2\varepsilon}{2N+1}+\frac{1}{2N+1}. \tag{8.30}$$

From (8.26),

$$\frac{P_N M_2}{2N+1} \le \frac{N}{2N+1} \to \frac{1}{2}$$

as N goes to infinity. So, the left side of (8.30) is less than $\frac{\varepsilon}{2}$ and for almost all x and y,

$$\lim_{N\to\infty}\frac{1}{2N+1}\sum_{m=-N}^{N}x[m+k]y[m]=0.$$

In a similar way one can show that

$$S_4=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{m=-N}^{N}y[m+k]x[m]=0$$

for almost every x and y. This concludes proving part (b). □

Remark 2. Due to Theorem 8, Theorem 9 can be trivially generalized to x and y that have constant amplitude K_1 and K_2, respectively, and have autocorrelation functions

$$A_x[k]=\begin{cases}K_1\left(1-\frac{|k|}{M_1}\right), & \text{if } 0\le|k|\le M_1,\\ 0, & \text{otherwise,}\end{cases}$$

and

$$A_y[k]=\begin{cases}K_2\left(1-\frac{|k|}{M_2}\right), & \text{if } 0\le|k|\le M_2,\\ 0, & \text{otherwise.}\end{cases}$$

Remark 3. Given $K>0$ and $M\in\mathbb{N}$, on $\mathbb{R}$, the inverse Fourier transform of $MK\left(\frac{\sin\pi M\gamma}{\pi M\gamma}\right)^2$ is $K\max\left(1-\frac{|t|}{M},0\right)$. By the additive property of Fourier transform, the inverse Fourier transform of $F(\gamma)=\sum_{n=1}^{N}nK_n\left(\frac{\sin\pi n\gamma}{\pi n\gamma}\right)^2$, restricted to $\mathbb{Z}$, is

$$\check{F}[m]=\sum_{n=1}^{N}K_n\max\left(1-\frac{|m|}{n},0\right).$$

Due to Theorem 8, one can construct functions x_n such that $A_{x_n}=K_n\max(1-\frac{|m|}{n},0)$ with $|x_n|=\sqrt{K_n}$. Theorem 9 implies that the sequence $x=x_1+\cdots+x_N$ has autocorrelation $\check{F}$. Also, $x\in\ell^\infty(\mathbb{Z})$ since $|x|$ is bounded by $\sum_{n=1}^{N}\sqrt{K_n}$. Thus, we

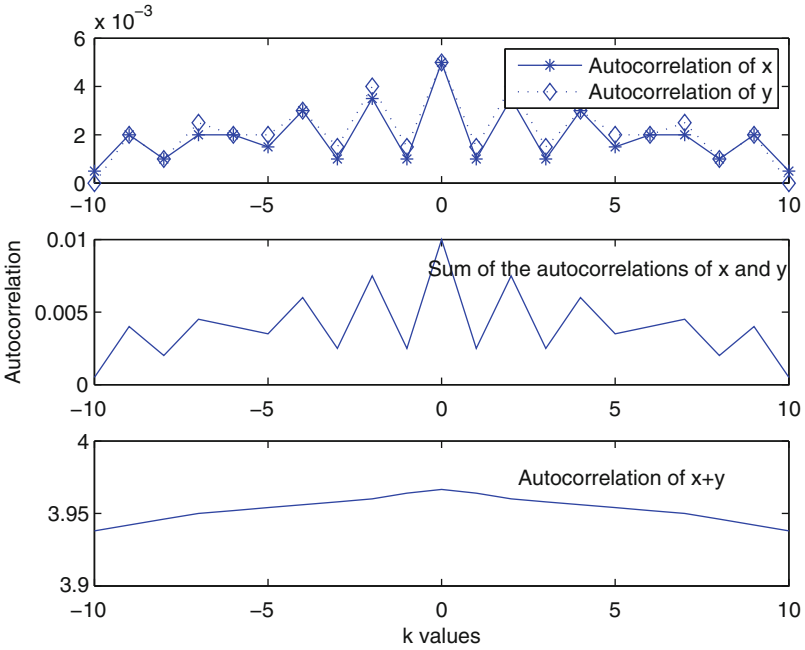

Fig. 8.3 Autocorrelations of two sequences x and y and their sum

have a function $x \in \ell^\infty(\mathbb{Z})$ whose autocorrelation is the inverse Fourier transform of dilates of Fejér functions.

Example 5. Generally, $A_{x+y}[k] \neq A_x[k] + A_y[k]$. In fact, in the case of real-valued sequences $x, y \in \ell^\infty(\mathbb{Z})$, when all limits as $N \to \infty$ exist, $A_{x+y}[k] = A_x[k] + A_y[k] + 2A_{xy}[-k]$, and there is no reason to expect $A_{xy}[-k] = 0$ for each $k \in \mathbb{Z}$. Here, A_{xy} is the cross-correlation of x and y defined by

$$\forall k \in \mathbb{Z},\ A_{xy}[k] = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} x[k+m]\overline{y[m]}.$$

As a particular example, note that the binary expansions, with a precision of 16 bit, of $\lambda_x = 0.35$ and $\lambda_y = 0.9$ are 0.01011001100110011 and 0.1110011001100110, respectively. From these one can obtain sequences x and y of ± 1s by following the definition of y in (8.2). The partial autocorrelations of x, y, and $x+y$ have been calculated by computing the sum in Definition 1 for $N = 1,000$, i.e., $2N+1 = 2,001$ terms. These partial autocorrelations at the integers between -10 and 10 are plotted in Fig. 8.3. Clearly, the sums of the autocorrelations of x and y do not match the autocorrelation of $x+y$.

8.4 Conclusions

In this chapter Hadamard matrices have been used to construct CAZAC sequences on $\mathbb{Z}$. Such sequences are important in the areas of radar and communication. This is generalized to the construction of unimodular sequences on $\mathbb{Z}$ whose autocorrelations are triangles. Finally, conditions under which the autocorrelation of the sum of two sequences is the same as the sum of the respective autocorrelations are studied.

Acknowledgment The first named author gratefully acknowledges the support of ONR Grant N00014-09-1-0144 and MURI-ARO Grant W911NF-09-1-0383. The second named author gratefully acknowledges the support of AFOSR Grant FA9550-10-1-0441.

References

1. Auslander L, Barbano PE (1998) Communication codes and Bernoulli transformations. Appl Comput Harmon Anal 5(2):109–128
2. Bell DA (1966) Walsh functions and Hadamard matrices. Electron Lett 2:340–341
3. Benedetto JJ (1991) A multidimensional Wiener–Wintner theorem and spectrum estimation. Trans Am Math Soc 327(2):833–852
4. Benedetto JJ (1997) Harmonic analysis and applications. CRC Press, Boca Raton, FL
5. Benedetto JJ, Benedetto RL, Woodworth JT (2011) Björk CAZACs: theory, geometry, implementation, and waveform ambiguity behavior. Preprint, to be submitted
6. Benedetto JJ, Benedetto RL, Woodworth JT (2012) Optimal ambiguity functions and Weil's exponential sum bounds. J. Fourier Analysis and Applications 18:471–487
7. Benedetto JJ, Czaja W (2009) Integration and modern analysis. Birkhäuser Boston Inc., Boston, MA
8. Benedetto JJ, Datta S (2010) Construction of infinite unimodular sequences with zero autocorrelation. Adv Comput Math 32(2):191–207
9. Benedetto JJ, Donatelli JJ (2007) Ambiguity function and frame theoretic properties of periodic zero autocorrelation waveforms. IEEE J Special Topics Signal Process 1:6–20
10. Björck J, Saffari B (1995) New classes of finite unimodular sequences with unimodular Fourier transforms. Circulant Hadamard matrices with complex entries. C R Acad Sci 320:319–324
11. Christensen O (2003) An introduction to frames and riesz bases. Birkhäuser, Basel
12. Datta S (2007) Wiener's generalized harmonic analysis and waveform design. Ph.D. thesis. University of Maryland, College Park, Maryland
13. Daubechies I (1992) Ten lectures on wavelets. Society for Industrial and Applied Mathematics, Philadelphia, PA
14. Hadamard J (1893) Résolution d'une question relative aux déterminants. Bulletin des Sciences Mathématiques 17:240–246
15. Helleseth T, Kumar PV (1998) Sequences with low correlation. In: Handbook of coding theory, Vol. I, II. North-Holland, Amsterdam, pp 1765–1853
16. Horadam KJ (2007) Hadamard matrices and their applications. Princeton University Press, Princeton, NJ
17. Horn RA, Johnson CR (1990) Matrix analysis. Cambridge University Press, Cambridge (Corrected reprint of the 1985 original)
18. Kerby R (1990) The correlation function and the Wiener–Wintner theorem in higher dimension. Ph.D. thesis. University of Maryland, College Park

19. Kerby R (2008) Personal communication
20. Levanon N, Mozeson E (2004) Radar signals. Wiley Interscience, IEEE Press
21. Long ML (2001) Radar reflectivity of land and sea. Artech House, Boston
22. Mow WH (1996) A new unified construction of perfect root-of-unity sequences. In: Proceedings of IEEE 4th International Symposium on Spread Spectrum Techniques and Applications (Germany), pp 955–959
23. Nathanson FE (1999) Radar design principles—signal processing and the environment. SciTech Publishing Inc., Mendham, NJ
24. Paley REAC (1933) On orthogonal matrices. J Math Phys 12:311–320
25. Richards MA, Scheer JA, Holm WA (eds.) (2010) Principles of modern radar: basic principles. SciTech Publishing Inc., Raleigh, NC
26. Schipp F, Wade WR, Simon P (1990) Walsh series, an introduction to dyadic harmonic analysis. Taylor & Francis
27. Stein EM, Weiss G (1971) Introduction to fourier analysis on euclidean spaces. Princeton University Press, Princeton, NJ
28. Stimson GW (1998) Introduction to airborne radar. SciTech Publishing Inc., Mendham, NJ
29. Sylvester JJ (1867) Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to newton's rule, ornamental tile-work, and the theory of numbers. Phil Mag 34:461–475
30. Ulukus S, Yates RD (2001) Iterative construction of optimum signature sequence sets in synchronous CDMA systems. IEEE Trans Inform Theor 47(5):1989–1998
31. Verdú S (1998) Multiuser detection. Cambridge University Press, Cambridge, UK
32. Viterbi AJ (1995) CDMA: principles of spread spectrum communication. Addison-Wesley, Pearson, Boston
33. Walsh JL (1923) A closed set of normal orthogonal functions. Am J Math 45:5–24
34. Wiener N (1930) Generalized harmonic analysis. Acta Math (55):117–258
35. Wiener N (1988) The Fourier integral and certain of its applications. Cambridge Mathematical Library (Reprint of the 1933 edition. With a foreword by Jean-Pierre Kahane). Cambridge University Press, Cambridge
36. Wiener N, Wintner A (1939) On singular distributions. J Math Phys 17:233–246

Part II
Multiscale Analysis

Chapter 9
On the Application of the SDLE to the Analysis of Complex Time Series

Jianbo Gao, Jing Hu, and Wen-wen Tung

Abstract Complex systems often generate highly nonstationary and multiscale signals because of nonlinear and stochastic interactions among their component systems and hierarchical regulations imposed by the operating environments. Rapid accumulation of such complex data in life sciences, systems biology, nano-sciences, information systems, and physical sciences has made it increasingly important to develop complexity measures that incorporate the concept of scale explicitly, so that different behaviors of signals on varying scales can be simultaneously characterized by the same scale-dependent measure. The scale-dependent Lyapunov exponent (SDLE) discussed here is such a measure and can be used as the basis for a unified theory of multiscale analysis of complex data. The SDLE can readily characterize deterministic low-dimensional chaos, noisy chaos, random $1/f^{\alpha}$ processes, random Levy processes, stochastic oscillations, and processes with multiple scaling behavior. It can also readily deal with many types of nonstationarity, detect intermittent chaos, and accurately detect epileptic seizures from EEG data and distinguish healthy subjects from patients with congestive heart failure from heart rate variability (HRV) data. More importantly, analyses of EEG and HRV data illustrate that commonly used complexity measures from information theory,

J. Gao (✉)
PMB Intelligence LLC, West Lafayette, IN 47996, USA

Mechanical and Materials Engineering, Wright State University, Dayton, Ohio 45435, USA
e-mail: jbgao.pmb@gmail.com

J. Hu
Affymetrix, Inc., 3380 Central Expressway, Santa Clara, CA 95051, USA
e-mail: jing.hu@gmail.com

W.-w. Tung
Department of Earth & Atmospheric Sciences, Purdue University,
West Lafayette, IN 47907, USA
e-mail: wwtung@purdue.edu

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,

DOI 10.1007/978-1-4614-4145-8_9,

chaos theory, and random fractal theory can be related to the values of the SDLE at specific scales, and useful information on the structured components of the data is also embodied by the SDLE.

9.1 Introduction

Complex systems are usually comprised of multiple subsystems that exhibit both highly nonlinear deterministic, as well as, stochastic characteristics, and are regulated hierarchically. These systems generate signals that exhibit complex characteristics such as sensitive dependence on small disturbances, long memory, extreme variations, and nonstationarity [1]. Examples of such signals are abundant, including biological data such as heart rate variability (HRV) and EEG data [2], highly bursty traffic on the Internet [3–5], and highly varying stock prices and foreign exchange rates in financial markets [6, 7]. For illustration, in Fig. 9.1, we have shown an example of HRV data for a normal young subject [8]. Evidently, the signal is

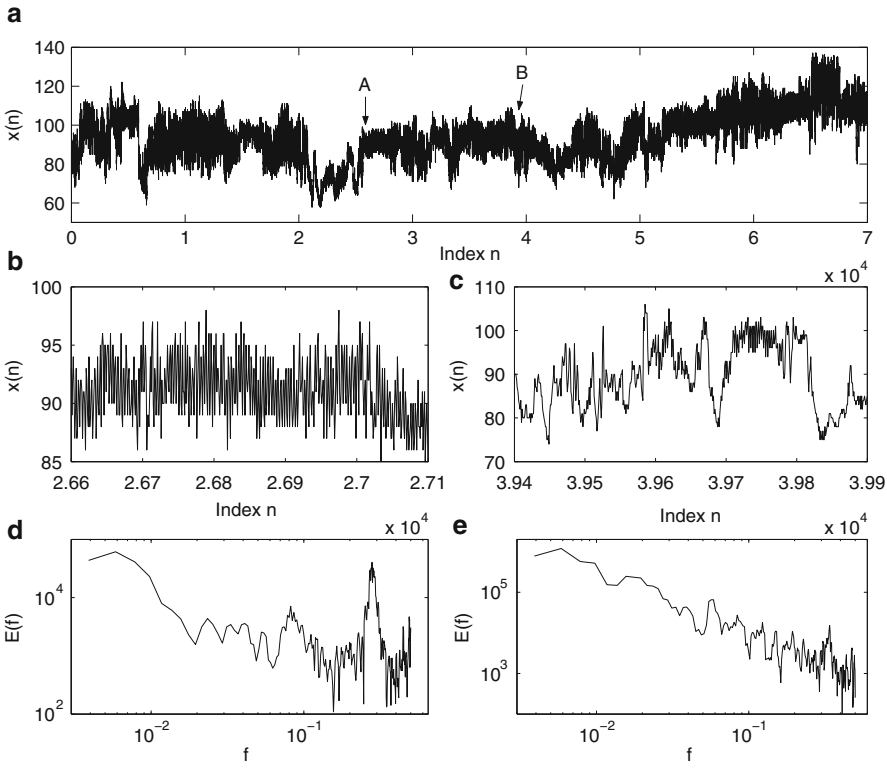

Fig. 9.1 (**a**) The HRV data for a normal subject; (**b**, **c**) the segments of signals indicated as A and B in (a); (**d**, **e**) power spectral density for the signals shown in (b,c)

highly nonstationary and multiscaled, appearing oscillatory for some period of time (Fig. 9.1b, d) and then varying as a $1/f$ process for another period of time (Fig. 9.1c, e).

While the multiscale nature of signals such as that shown in Fig. 9.1 cannot be fully characterized by existing methods, the nonstationarity of the data is even more troublesome, because it prevents direct application of spectral analysis or methods based on chaos theory and random fractal theory. For example, in order to reveal that the HRV data is of $1/f$ nature [9, 10] with anti-persistent long-range correlations [11, 12] and multifractality [13], time series such as shown in Fig. 9.1a have to be preprocessed to remove components (such as oscillatory ones) that do not conform to fractal scaling analysis. However, automated segmentation of complex biological signals to remove undesired components is a significant open problem, since it is closely related to the challenging task of accurately detecting transitions from normal to abnormal states in physiological data.

Rapid accumulation of complex data in life sciences, systems biology, nanosciences, information systems, and physical sciences has made it increasingly important to be able to analyze multiscale and nonstationary data. Since multiscale signals behave differently depending upon which scale the data are looked at, it is of fundamental importance to develop measures that explicitly incorporate the concept of scale so that different behaviors of the data on varying scales can be simultaneously characterized by the same scale-dependent measure. Here, we discuss such a measure, the scale-dependent Lyapunov exponent (SDLE), and develop a unified multiscale analysis theory of complex data.

This chapter is organized as follows. We first define the SDLE, then apply it to characterize low-dimensional chaos, noisy chaos, and random processes with power-law decaying power spectral density (so-called $1/f^{\alpha}$ processes). We then show how it can readily detect intermittent chaos, deal with nonstationarity, and apply it to characterize EEG and HRV data. Finally, we make a few concluding remarks, including a discussion of best practices for experimental data analysis.

9.2 SDLE: Definitions and Fundamental Properties

Chaos and random fractal theories have been used extensively in the analysis of complex data [2, 3, 6, 11–18]. Chaos theory shows that apparently irregular behaviors in a complex system may be generated by nonlinear deterministic interactions of only a few numbers of degrees of freedom while noise or intrinsic randomness does not play any role. Random fractal theory, on the other hand, assumes that the dynamics of the system are inherently random. One of the most important classes of random fractals is the set of $1/f^{\alpha}$ processes that display long-range correlations. Since the foundations of chaos theory and random fractal theory are entirely different, different conclusions may be drawn depending upon which theory is used to analyze the data. In fact, much of the research in the past has been devoted to determining whether a complex time series is generated

by a chaotic or a random system [19–30]. In this effort, $1/f^{\alpha}$ processes have distinguished themselves as the key counter examples invalidating commonly used tests for chaos [30–32]. Thus, successful classification of chaos and $1/f^{\alpha}$ processes based on scales may fundamentally change the practice of time series analysis—these theories will be used synergistically, instead of individually, to characterize the behaviors of signals on a wide range of scales.

SDLE is a generalization of two important concepts, the time-dependent exponent curves [24] and the finite size Lyapunov exponent (LE) [33]. It was first introduced by the authors in [34, 35], and has been further extended in [36, 37] and applied to study EEG [38], HRV [39, 40], and Earth's geodynamo [41].

We assume that all that is known is a scalar time series $x(1), x(2), \ldots, x(n)$. Regardless of whether the dynamics are chaotic or random, we use time-delay embedding [42–44] to form vectors of the form: $V_i = [x(i), x(i+L), \ldots, x(i+(m-1)L)]$, where the embedding dimension m and the delay time L are chosen according to optimization criteria [24, 45]. When the time series is random, such a procedure transforms the self-affine stochastic process to a self-similar process in phase space. In this case, however, the specific values of m and L are not important, so long as $m > 1$.

After a proper phase space is reconstructed, we consider an ensemble of trajectories. We denote the initial separation between two nearby trajectories by ε_0 and their *average separation* at time t and $t + \Delta t$ by ε_t and $\varepsilon_{t+\Delta t}$, respectively. We then examine the relation between ε_t and $\varepsilon_{t+\Delta t}$, where Δt is small. When $\Delta t \to 0$, we have

$$\varepsilon_{t+\Delta t} = \varepsilon_t e^{\lambda(\varepsilon_t)\Delta t}, \tag{9.1}$$

where $\lambda(\varepsilon_t)$ is the SDLE. It is given by

$$\lambda(\varepsilon_t) = \frac{\ln \varepsilon_{t+\Delta t} - \ln \varepsilon_t}{\Delta t}. \tag{9.2}$$

Equivalently, we have a differential equation for ε_t,

$$\frac{d\varepsilon_t}{dt} = \lambda(\varepsilon_t)\varepsilon_t. \tag{9.3}$$

Given a time series data, the smallest Δt possible is the sampling time τ.

To compute SDLE, we can start from an arbitrary number of shells,

$$\varepsilon_k \le \|V_i - V_j\| \le \varepsilon_k + \Delta\varepsilon_k, \quad k = 1, 2, 3, \ldots, \tag{9.4}$$

where V_i, V_j are reconstructed vectors and ε_k (the radius of the shell) and $\Delta\varepsilon_k$ (the width of the shell) are arbitrarily chosen small distances ($\Delta\varepsilon_k$ is not necessarily a constant). We then monitor the evolution of all of the pairs of vectors (V_i, V_j) within a shell and take the average. As we will see shortly, as far as estimation of

the parameters corresponding to exponential or power-law divergence is concerned, taking logarithm and averaging can be exchanged; (9.2) can now be written as

$$\lambda(\varepsilon_t) = \frac{\left\langle \ln \|V_{i+t+\Delta t} - V_{j+t+\Delta t}\| - \ln \|V_{i+t} - V_{j+t}\| \right\rangle}{\Delta t} \tag{9.5}$$

where t and Δt are integers in units of the sampling time and the angle brackets denote the average within a shell.

To see why taking logarithm and averaging can be exchanged for the purpose of computing $\lambda(\varepsilon_t)$, let us consider a case involving $\varepsilon_1(t) = \varepsilon_1(0)\mathrm{e}^{\lambda_1 t}, \varepsilon_2(t) = \varepsilon_2(0)\mathrm{e}^{\lambda_2 t}$, where $\lambda_1 = \lambda_2$ is a positive constant. Let $\varepsilon(t)$ be the average of $\varepsilon_1(t)$ and $\varepsilon_2(t)$. Then it is clear that SDLE is simply λ whether one takes average first or takes logarithm first. In fact, when large t is concerned, if λ_1 is slightly larger than λ_2, then taking logarithm first is beneficial, since, otherwise, the term $\mathrm{e}^{\lambda_1 t}$ will dominate, and thus the presence of λ_2 will not be captured. Clearly, similar argument applies to the situation of power-law divergence.

Note that in the above computational procedure, the initial set of shells for computing SDLE serve as initial values of the scales; through evolution of the dynamics, they will automatically converge to the range of inherent scales—which are the scales that define (9.2) and (9.3). Also note that when analyzing chaotic time series, the condition

$$|j - i| \geq (m-1)L \tag{9.6}$$

needs to be imposed when finding pairs of vectors within a shell, to eliminate the effects of tangential motions [24] and for an initial scale to converge to the inherent scales [35].

To better understand the notion of "inherent scales," it is beneficial to discuss the notion "characteristic scale" (or "limiting scale"), ε_∞, defined as the scale where SDLE is close to 0. If one starts from $\varepsilon_0 \ll \varepsilon_\infty$, then, regardless of whether the data is deterministically chaotic or simply random, ε_t will initially increase with time and gradually settle around ε_∞. Consequentially, $\lambda(\varepsilon_t)$ will be positive before ε_t reaches ε_∞. On the other hand, if one starts from $\varepsilon_0 \gg \varepsilon_\infty$, then ε_t will simply decrease, yielding negative $\lambda(\varepsilon_t)$, again regardless of whether the data are chaotic or random. When $\varepsilon_0 \sim \varepsilon_\infty$, then $\lambda(\varepsilon_t)$ will stay around 0—note, however, that ε_∞ may not be a single point but a function of time, such as a periodic function of time. These discussions make it clear that chaos can only be observed on scales much smaller than ε_∞.

To better understand SDLE, we now point out a relation between SDLE and the largest positive LE λ_1 estimated for a true chaotic signal using, say, the Wolf et al.'s algorithm [21]. It is given by [35]

$$\lambda_1 = \int_0^{\varepsilon^*} \lambda(\varepsilon) p(\varepsilon) \mathrm{d}\varepsilon, \tag{9.7}$$

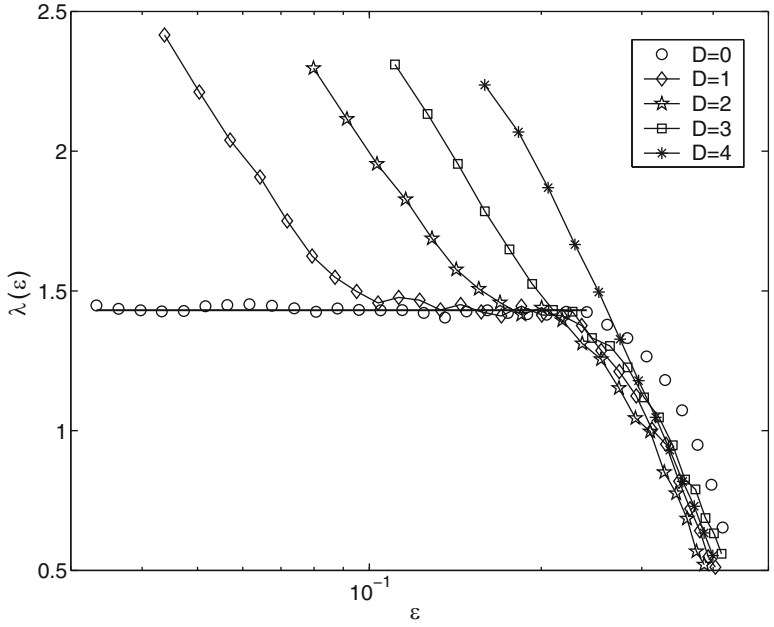

Fig. 9.2 $\lambda(\varepsilon)$ curves for clean and noisy Lorenz systems

where ε^* is a scale parameter (e.g., used for renormalization when using Wolf et al.'s algorithm [21]); $p(\varepsilon)$ is the probability density function for the scale ε given by

$$p(\varepsilon) = Z\frac{\mathrm{d}C(\varepsilon)}{\mathrm{d}\varepsilon}, \tag{9.8}$$

where Z is a normalization constant satisfying $\int_0^{\varepsilon^*} p(\varepsilon)\mathrm{d}\varepsilon = 1$ and $C(\varepsilon)$ is the well-known Grassberger–Procaccia's correlation integral [19]. Note that the lower bound for the integration is set to be zero here. In practice, on scales smaller than $\varepsilon_{\min}$, the probability $p(\varepsilon)$ will be zero. Therefore, one could replace the lower bound for the integration by $\varepsilon_{\min}$.

To understand the SDLE, it is instructive to apply it to characterize chaotic signals and $1/f^{\alpha}$ processes. First, we analyze the chaotic Lorenz system with stochastic forcing

$$\begin{aligned} \mathrm{d}x/\mathrm{d}t &= -16(x-y) + D\eta_1(t), \\ \mathrm{d}y/\mathrm{d}t &= -xz + 45.92x - y + D\eta_2(t), \\ \mathrm{d}z/\mathrm{d}t &= xy - 4z + D\eta_3(t). \end{aligned} \tag{9.9}$$

where $\eta_i(t)$, $i = 1, 2, 3$ are independent Gaussian noise forcing terms with zero mean and unit variance. When $D = 0$, the system is clean. Figure 9.2 shows five curves,

for the cases with $D = 0, 1, 2, 3, 4$. The computations are done with 10,000 points and $m = 4$, $L = 2$. We observe the following interesting features:

- For the clean chaotic signal, $\lambda(\varepsilon)$ fluctuates slightly around a constant. As is expected, this constant is the very largest positive LE, λ_1,

$$\lambda(\varepsilon) = \lambda_1. \tag{9.10}$$

  The small fluctuation in $\lambda(\varepsilon)$ is due to the fact that the divergence rate on the Lorenz attractor varies from one region to another.
- When there is stochastic forcing, $\lambda(\varepsilon)$ is no longer a constant when ε is small but diverges to infinity as $\varepsilon \to 0$ according the following scaling law,

$$\lambda(\varepsilon) \sim -\gamma \ln \varepsilon, \tag{9.11}$$

  where γ is a coefficient controlling the speed of loss of information. This feature suggests that entropy generation is infinite when the scale ε approaches zero.
- When the noise is increased, the part of the curve with $\lambda(\varepsilon) \sim -\gamma \ln \varepsilon$ shifts to the right. In fact, little chaotic signature can be identified when D is increased beyond 3.

Note that similar results to those shown in Fig. 9.2 have been observed in other model chaotic systems, including the Mackey–Glass delay differential equation with multiple positive Lyapunov exponents [46]. To simplify our discussion of HRV data analysis, we note that in order to resolve the behavior of $\lambda(\varepsilon)$ on ever smaller scales, longer and longer time series have to be used. More precisely, for a given dataset, if the smallest resolvable scale is ε_0, in order to resolve a smaller scale ε_0/r, where $r > 1$, a larger dataset has to be used—the larger the dimension of the attractor, the longer the time series has to be.

Next we consider $1/f^{\alpha}$ processes. Such type of processes is ubiquitous in science and engineering (see [47] and references therein). Two important prototypical models for such processes are fractional Brownian motion (fBm) process [48] and ON/OFF intermittency with power-law distributed ON and OFF periods [47]. For convenience, we introduce the Hurst parameter $0 < H < 1$ through a simple equation

$$\alpha = 2H + 1. \tag{9.12}$$

Depending on whether H is smaller than, equal to, or larger than $1/2$, the process is said to have anti-persistent correlation, short-range correlation, and persistent long-range correlation [47]. Note that $D = 1/H$ is the fractal dimension of such processes and Kolmogorov's 5/3 law for the energy spectrum of fully developed turbulence [49] corresponds to $H = 1/3$.

It is well known that the variance of such stochastic processes increases with t as t^{2H}. Translating this into the average distance between nearby trajectories, we immediately have

$$\varepsilon_t = \varepsilon_0 t^H. \tag{9.13}$$

Using (9.2), we then have $\lambda(\varepsilon_t) \sim H/t$. Expressing t by ε_t, we obtain

$$\lambda(\varepsilon_t) \sim H\varepsilon_t^{-1/H}. \tag{9.14}$$

Equation(9.14) can be readily verified by calculating $\lambda(\varepsilon_t)$ from such processes.

SDLE also has distinct scaling laws for random Levy processes, stochastic oscillations, and complex motions with multiple scaling laws on different scale ranges. For the details, we refer to [34, 35].

Finally, we emphasize that $\lambda_1 > 0$ (say, computed by Wolf et al.'s algorithm [21]) is not a sufficient condition for chaos. This is evident from (9.7): any non-chaotic scalings of SDLE such as (9.11) and (9.14) will yield $\lambda_1 > 0$.

9.2.1 Detecting Intermittent Chaos by SDLE

Intermittent chaos is a type of complex motion where regular (i.e., periodic) and chaotic motions alternate. Note that the stretches of the time periods for regular motions could be considerably longer than those of chaotic motions. Exactly because of this, standard methods are unable to detect chaos in such motions. This, however, is a simple task for SDLE. To illustrate the idea, we examine the logistic map

$$x_{n+1} = ax_n(1 - x_n). \tag{9.15}$$

When $a = 3.8284$, we have intermittent chaos. An example of the time series is shown in Fig. 9.3a. We observe that time intervals exhibiting chaos are very short compared with those exhibiting periodic motions. Traditional methods for computing LE, being based on global average, are unable to quantify chaos in such an intermittent situation, since the laminar phase dominates, neither can FSLE, since it requires that divergence dominates most of the time. Interestingly, the SDLE curve shown in Fig. 9.3b clearly indicates existence of chaotic motions, since the plateau region extends almost one order of magnitude.

Why can SDLE even detect chaos in such a situation? The reason is that the oscillatory part of the data only affects the scale range where $\lambda(\varepsilon) \sim 0$. It cannot affect the positive portion of $\lambda(\varepsilon)$. This means SDLE has a nice scale separation property to automatically separate the regular from chaotic motions.

9.2.2 Dealing with Nonstationarity

To facilitate our discussion of HRV data below, we now consider complicated processes generated by the following two scenarios. One is to concatenate randomly $1/f^{2H+1}$ and oscillatory components. Another is to superimpose oscillatory components on $1/f^{2H+1}$ process at randomly chosen time intervals. Either scenario generates signals that appear quite similar to that shown in Fig. 9.1a. The $\lambda(\varepsilon)$ curves for such processes are shown in Fig. 9.4, for a wide range of the H parameter.

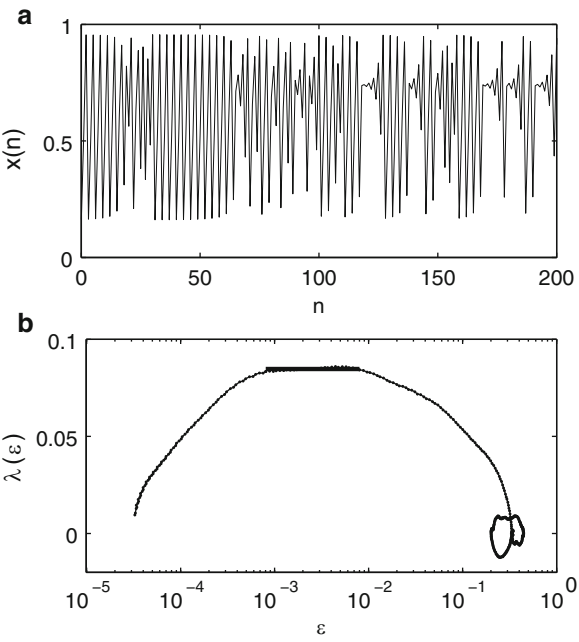

Fig. 9.3 (**a**) An intermittent time series generated by the logistic map with $a = 3.8284$ and $\sigma = 0$. (**b**) The SDLE curve for a time series of 10,000 points, with $m = 4, L = 1$, and a shell size of $(2^{-13.5}, 2^{-13})$. A plateau is clearly visible

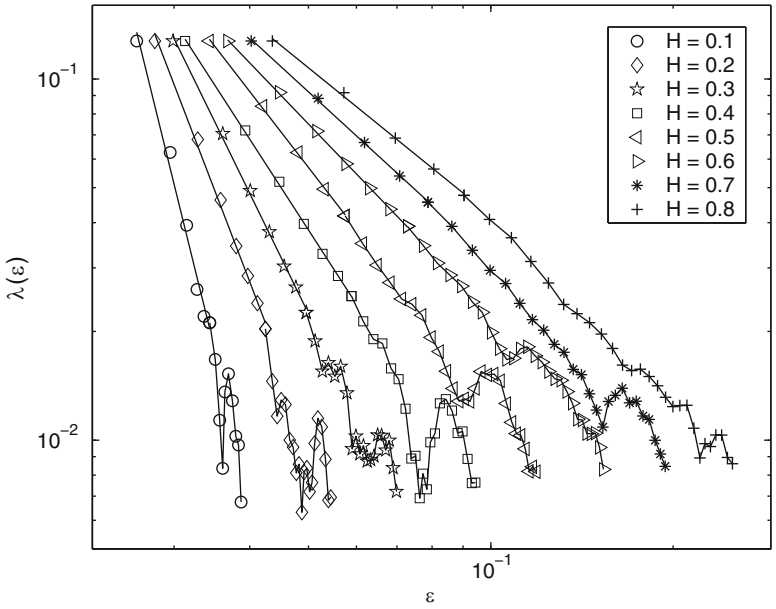

Fig. 9.4 $\lambda(\varepsilon)$ vs. ε curves for the simulation data. Eight different H are considered. To put all the curves on one plot, the curves for different H (except the smallest one considered here) are arbitrarily shifted rightward

We observe well-defined power-law relations, consistent with (9.14), when $\lambda(\varepsilon) > 0.02$. This result clearly shows that oscillatory components in the signals can only affect the SDLE where $\lambda(\varepsilon)$ is close to 0. This is an illustration of the effect of scale isolation by the SDLE.

When we perturb chaotic data by similar procedures, will we still be able to detect chaos? The answer is yes. In fact, the intermittent chaos discussed above may be viewed as an example of such a procedure.

We are now ready to fully understand why the SDLE can deal with the types of nonstationarity discussed here. One type of nonstationarity causes shifts of the trajectory in phase space—the greater the nonstationarity, the larger the shifts. SDLE, however, cannot be affected much by shifts, especially large ones, since it is based on the coevolution of pairs of vectors within chosen small shells. The other type is related to oscillatory components. They only affect SDLE near where it is close to zero; therefore, they will not alter the distinct scaling for chaos and fractal processes.

9.3 Applications: Biological Data Analysis

To better understand the SDLE and appreciate its power, we now apply it to examine two types of physiological data, HRV and EEG.

9.3.1 EEG Analysis

EEG signals provide a wealth of information about brain dynamics, especially related to cognitive processes and pathologies of the brain such as epileptic seizures. To understand the nature of brain dynamics as well as to develop novel methods for the diagnosis of brain pathologies, a number of complexity measures have been used in the analysis of EEG data. These include the Lempel–Ziv (LZ) complexity [50], the permutation entropy [51], the LE [21], the Kolmogorov entropy [20], the correlation dimension D_2 [19, 52], and the Hurst parameter [53–55]. We compare the SDLE with these complexity measures or their close relatives.

The EEG signals analyzed here were measured intracranially by the Shands hospital at the University of Florida. Such EEG data are also called depth EEG and are considered cleaner and more free of artifacts than scalp (or surface) EEG. Altogether, we have analyzed seven patients' multiple-channel EEG data, each with a duration of a few hours, with a sampling frequency of 200 Hz. When analyzing EEG for epileptic seizure prediction/detection, it is customary to partition a long EEG signal into short windows of length W points and calculate the measure of interest for each window. The criterion for choosing W is such that the EEG signal in each window is fairly stationary, is long enough to reliably estimate the measure of interest, and is short enough to accurately resolve localized activities such as

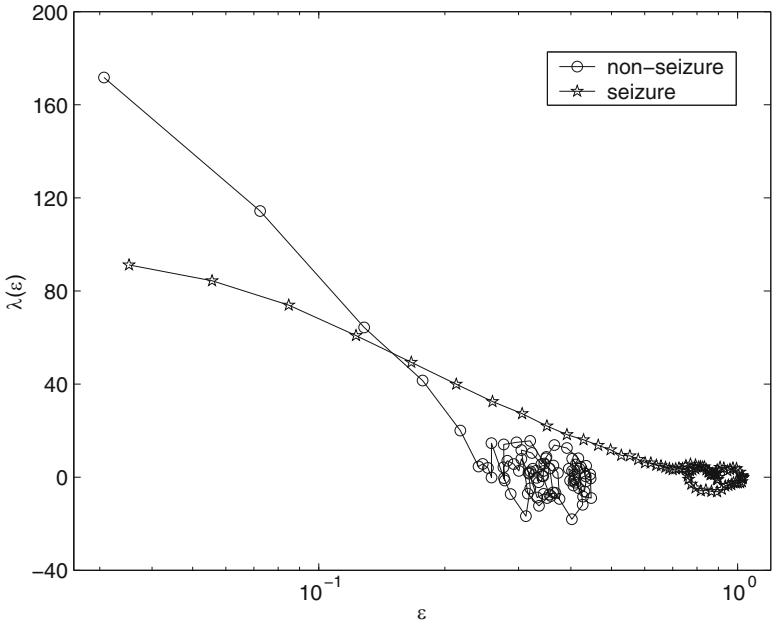

Fig. 9.5 Representative $\lambda(\varepsilon)$ (per second) vs. ε for a seizure and non-seizure EEG segment

seizures. Since seizure activities usually last about 1–2 min, in practice, one often chooses W to be about 10 s. When applying methods from random fractal theory such as detrended fluctuation analysis (DFA) [53], it is most convenient when the length of a sequence is a power of 2. Therefore, we have chosen $W = 2,048$ when calculating various measures. We have found, however, that the variations of these measures with time are largely independent of the window size W. The relations among the measures studied here are the same for all the seven patients' EEG data, so we illustrate the results based on only one patient's EEG signals.

We have examined the variation of $\lambda(\varepsilon)$ with ε for each segment of the EEG data. Two representative examples for seizure and non-seizure segments are shown in Fig. 9.5. We observe that on a specific scale ε^*, the two curves cross. Loosely, we may term any $\varepsilon < \varepsilon^*$ as small scale, while any $\varepsilon > \varepsilon^*$ as large scale. Therefore, on small scales, $\lambda(\varepsilon)$ is smaller for seizure than for non-seizure EEG, while on large scales, the opposite is true. The variations of $\lambda_{\text{small}-\varepsilon}$ and $\lambda_{\text{large}-\varepsilon}$ with time for this patient's data, where small $-\ \varepsilon$ and large $-\ \varepsilon$ stand for (more or less arbitrarily) chosen fixed small and large scales, are shown in Fig. 9.6a, b, respectively. We observe two interesting features: (1) the pattern of variation of $\lambda_{\text{small}-\varepsilon}(t)$ is reciprocal of that of $\lambda_{\text{large}-\varepsilon}(t)$. This result can be expected from Fig. 9.5. (2) The variations in $\lambda_{\text{small}-\varepsilon}(t)$ and $\lambda_{\text{large}-\varepsilon}(t)$ clearly indicate the two seizure events. Therefore, either $\lambda_{\text{small}-\varepsilon}(t)$ or $\lambda_{\text{large}-\varepsilon}(t)$ can be used to detect epileptic seizures accurately.

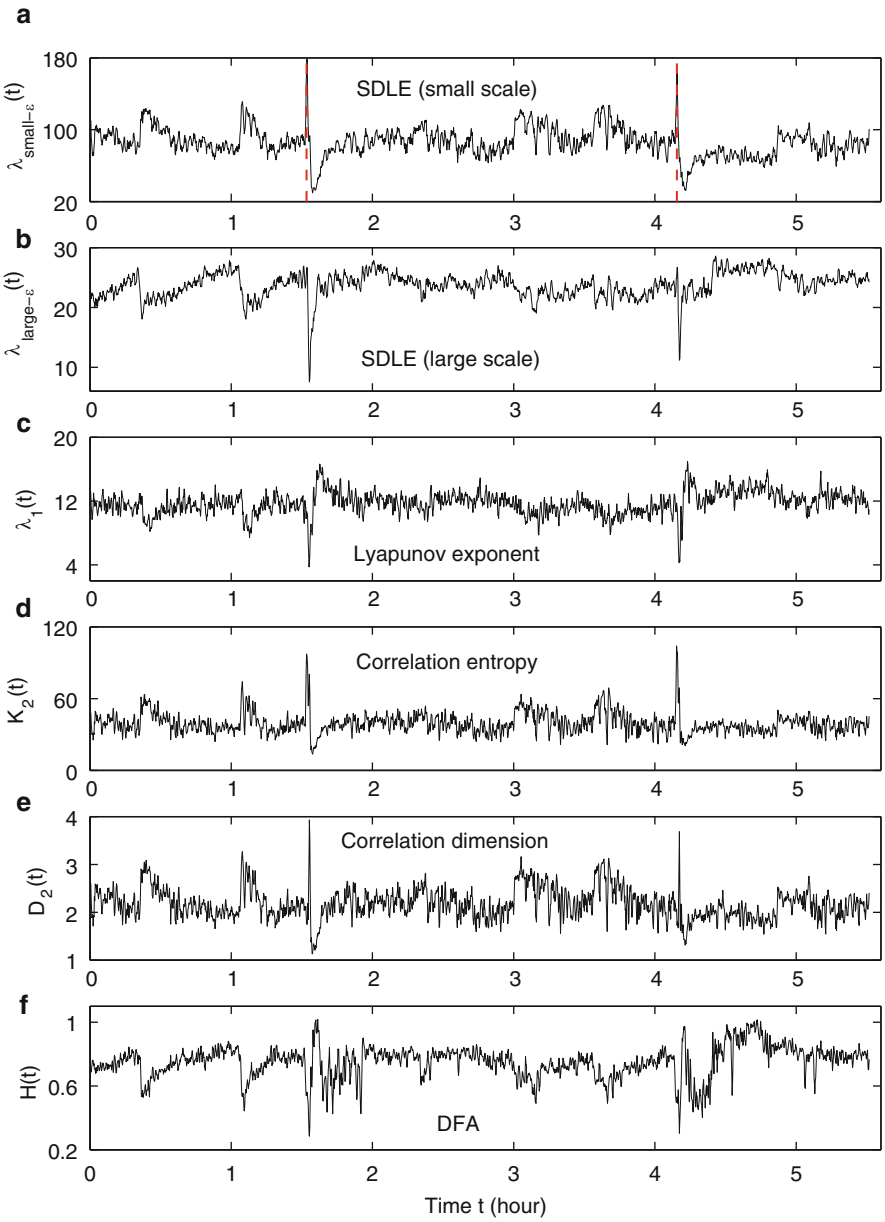

Fig. 9.6 The variation of (**a**) $\lambda_{\text{small}-\varepsilon}$, (**b**) $\lambda_{\text{large}-\varepsilon}$, (**c**) the LE, (**d**) the K_2 entropy, (**e**) the D_2, and (**f**) the Hurst parameter with time for EEG signals of a patient. The vertical *dashed lines* in (a) indicate seizure occurrence times determined by medical experts

We now compare the SDLE with three commonly used measures from chaos theory, the largest positive LE [21], the correlation entropy [20], the correlation dimension [19], and one measure from random fractal theory, the Hurst parameter. We discuss the three measures from chaos theory first.

The LE is a dynamic quantity. It characterizes the exponential growth of an infinitesimal line segment, $\varepsilon_t \sim \varepsilon_0 e^{\lambda_1 t}$, $\varepsilon_0 \to 0$. It is often computed by the algorithm of Wolf et al. [21], by monitoring the exponential divergence between a reference and a perturbed trajectory. For truly chaotic signals, $1/\lambda_1$ gives the prediction time scale of the dynamics. Also, it is well known that the sum of all the positive Lyapunov exponents in a chaotic system equals the Kolmogorov–Sinai (KS) entropy. The KS entropy characterizes the rate of creation of information in a system. It is zero, positive, and infinite for regular, chaotic, and random motions, respectively. It is difficult to compute, however. Therefore, one usually computes the correlation entropy K_2, which is a tight lower bound of the KS entropy. Similarly, the box-counting dimension, which is a geometrical quantity characterizing the minimal number of variables that are needed to fully describe the dynamics of a motion, is difficult to compute, and one often calculates the correlation dimension D_2 instead. Again, D_2 is a tight lower bound of the box-counting dimension. Both K_2 and D_2 can be readily computed from the correlation integral through the relation [19, 20].

$$C(m,\varepsilon) \sim \varepsilon^{D_2} e^{-mL\tau K_2} \tag{9.16}$$

where m and L are the embedding dimension and the delay time, τ is the sampling time, and $C(m,\varepsilon)$ is the correlation integral defined by

$$C(m,\varepsilon) = \frac{1}{N^2} \sum_{i,j=1}^{N} \theta(\varepsilon - \|V_i - V_j\|), \tag{9.17}$$

where $\theta(y)$ is the Heaviside step function taking values 1 or 0 depending on whether $y \geq 0$ or not, V_i and V_j are reconstructed vectors, N is the number of points in the time series, and ε is a prescribed small distance. Equation (9.16) means that in a plot of $\ln C(m,\varepsilon)$ vs. $\ln \varepsilon$ with m as a parameter, for truly low-dimensional chaos, one observes a series of parallel straight lines, with the slope being D_2 and the spacing between the lines estimating K_2 (where lines for larger m lie below those for smaller m). From these descriptions, one would expect that $\lambda_1(t)$ and $K_2(t)$ are similar, while $D_2(t)$ has little to do with either $\lambda_1(t)$ or $K_2(t)$. Surprisingly, from Fig. 9.6c–e, we observe that this is not the case: $\lambda_1(t)$ is similar to $D_2(t)$ but reciprocal of $K_2(t)$. In a moment, we shall explain how these puzzling relations may be understood based on $\lambda_{\text{small}-\varepsilon}(t)$ and $\lambda_{\text{large}-\varepsilon}(t)$.

Next we consider the calculation of the Hurst parameter H. As pointed out earlier, H characterizes the long-term correlations in a time series. There are many different ways to estimate H. We have chosen DFA [53], since it is more reliable [47] and has been used to study EEG [54, 55].

DFA works as follows: First, divide a given EEG data of length N, which is treated as a random walk process, into $\lfloor N/l \rfloor$ nonoverlapping segments (where $\lfloor x \rfloor$ denotes the largest integer that is not greater than x), each containing l points; then define the local trend in each segment to be the ordinate of a linear least-squares fit of the time series in that segment; finally compute the "detrended walk," denoted by $x_l(n)$, as the difference between the original "walk" $x(n)$ and the local trend. One then examines the following scaling behavior:

$$F_d(l) = \left\langle \sum_{i=1}^{l} x_l(i)^2 \right\rangle \sim l^{2H} \tag{9.18}$$

where the angle brackets denote ensemble averages of all the segments. From our EEG data, we have found that the power-law fractal scaling breaks down around $l \approx 6$. This is caused by distinct time scales defined by the α rhythm [54] or the dendritic time constants [55]. Figure 9.6f shows $H(t)$ for our EEG data . We observe that the pattern of $H(t)$ is very similar to that of $\lambda_1(t)$ but reciprocal to $K_2(t)$ and $D_2(t)$. Such relations cannot be readily understood intuitively, since the foundations for chaos theory and random fractal theory are entirely different.

Let us now resolve all of the curious relations observed between $\lambda_1(t)$, $K_2(t)$, $D_2(t)$, and $H(t)$.

- Generally, entropy measures the randomness of a dataset. This pertains to small scale. Therefore, $K_2(t)$ should be similar to $\lambda_{\text{small}-\varepsilon}(t)$. This is indeed the case. We should point out that we have also calculated other entropy-related measures, such as the LZ complexity [50], which is closely related to the Shannon entropy and permutation entropy [51], and observed similar variations. Therefore, we can conclude that the variation of the entropy is represented by $\lambda_{\text{small}-\varepsilon}(t)$, regardless of how entropy is defined.
- To understand why $\lambda_1(t)$ calculated by the algorithm of Wolf et al. [21] corresponds to $\lambda_{\text{large}-\varepsilon}(t)$, we note that the algorithm of Wolf et al. [21] involves a scale parameter that whenever the divergence between a reference and a perturbed trajectory exceeds this chosen scale, a renormalization procedure is performed. When the algorithm of Wolf et al. [21] is applied to a time series with only a few thousand points, in order to obtain a well-defined LE, a fairly large-scale parameter has to be chosen. This is the reason that the LE and $\lambda_{\text{large}-\varepsilon}$ are similar. In fact, the scale we have chosen to calculate $\lambda_1(t)$ is even larger than that for calculating $\lambda_{\text{large}-\varepsilon}(t)$. This is the reason that the value of $\lambda_1(t)$ shown in Fig. 9.6c is smaller than that of $\lambda_{\text{large}-\varepsilon}(t)$ shown in Fig. 9.6b.
- It is easy to see that if one fits the $\lambda(\varepsilon)$ curves shown in Fig. 9.5 by a straight line, then the variation of the slope with time should be similar to $\lambda_{\text{small}-\varepsilon}(t)$ but reciprocal of $\lambda_{\text{large}-\varepsilon}(t)$. Such a pattern will be preserved even if one takes the logarithm of $\lambda(\varepsilon)$ first and then does the fitting. Such a discussion makes it clear that even if EEG is not ideally of the $1/f^{2H+1}$ type, qualitatively, the relation

$\lambda(\varepsilon) \sim \varepsilon^{-1/H}$ holds. This in turn implies $D_2 \sim 1/H$. With these arguments, it is clear that the seemingly puzzling relations among the measures considered here can be readily understood by the $\lambda(\varepsilon)$ curves. Most importantly, we have established that commonly used complexity measures can be related to the values of the SDLE at specific scales.

As we have pointed out, around the characteristic scale $\overline{\varepsilon}$, $\lambda(\varepsilon)$ is always close to 0. The pattern of $\lambda(\varepsilon)$ around $\overline{\varepsilon}$ is governed by the structured components in the data, such as the α, γ, β, and δ waves. From Fig. 9.5, we observe that the patterns for seizure and non-seizure EEG segments are very different. Such information is certainly helpful in predicting/detecting seizures. Since the numerous measures considered here are already very effective for this purpose, we will not pursue this issue further here. The issue becomes more important when distinguishing healthy subjects from patients with heart disease using the HRV data, as we will soon show.

9.3.2 HRV Analysis

HRV is an important dynamical variable of the cardiovascular function. Its most salient feature is the spontaneous fluctuation, even when the environmental parameters are maintained constant and no perturbing influences can be identified. Since the observation that HRV is related to various cardiovascular disorders [56], a number of methods have been proposed to analyze HRV data. They include methods based on simple statistics from time and frequency domain analyses (see [57] and references therein), as well as those derived from chaos theory and random fractal theory [10, 14, 15, 58–61]. We shall now show that the SDLE can readily characterize the hidden differences in the HRV under healthy and diseased conditions and shed much new light on the dynamics of the cardiovascular system.

We examine two types of HRV data, one for healthy subjects and another for subjects with the congestive heart failure (CHF), a life-threatening disease. The data were downloaded from the PhysioNet [8]. There are 18 healthy subjects and 15 subjects with CHF. Part of these datasets were analyzed by random fractal theory. In particular, 12 of the 15 CHF datasets were analyzed by wavelet-based multifractal analysis [13], for the purpose of distinguishing healthy subjects from CHF patients. For ease of comparison, we take the first 3×10^4 points of both groups of HRV data for analysis. In Fig. 9.7a, b, we have shown two typical $\lambda(\varepsilon)$ vs. ε curves, one for a healthy subject and another for a patient with CHF. We observe that for the healthy subject, $\lambda(\varepsilon)$ linearly decreases with $\ln\varepsilon$ before λ reaches around 0, or, before ε settles around the characteristic scale, $\overline{\varepsilon}$. Recall that this is a characteristic of noisy dynamics (Fig. 9.2). For the CHF case plotted in Fig. 9.7b, we observe that the $\lambda(\varepsilon)$ is oscillatory, with its value always close to 0, and hence, the only scale resolvable is around $\overline{\varepsilon}$. Since the length of the time series used in our analysis for the healthy and the CHF subjects is the same, the inability of resolving the $\lambda(\varepsilon)$ behavior on scales

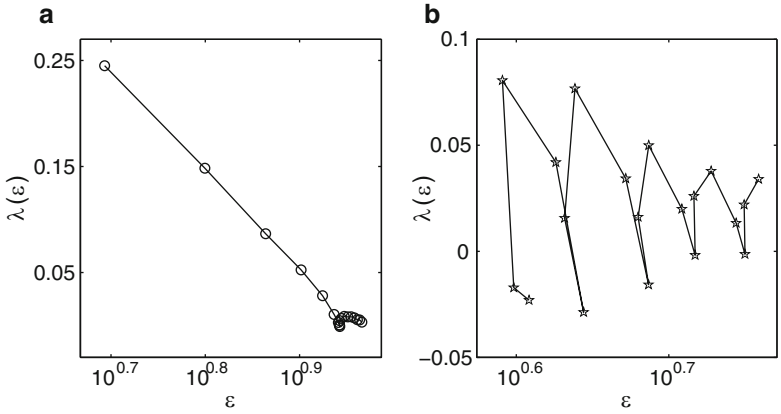

Fig. 9.7 $\lambda(\varepsilon)$ (per beat) vs. ε (in semi-log scale) for HRV data of (**a**) a healthy subject and (**b**) a subject with CHF

much smaller than $\overline{\varepsilon}$ for patients with CHF strongly suggests that the dimension of the dynamics of the cardiovascular system for CHF patients is considerably higher than that for healthy subjects.

We now discuss how to distinguish between healthy subjects and patients with CHF from HRV analysis. We have devised two simple measures, or features. One is to characterize how well the linear relation between $\lambda(\varepsilon)$ and $\ln\varepsilon$ can be defined. We have quantified this by calculating the error between a fitted straight line and the actual $\lambda(\varepsilon)$ vs. $\ln\varepsilon$ plots of Fig. 9.7a, b. The second feature is to characterize how well the characteristic scale $\overline{\varepsilon}$ is defined. This is quantified by the ratio between two scale ranges, one is from the second to the sixth point of the $\lambda(\varepsilon)$ curves and another is from the seventh to the 11th point of the $\lambda(\varepsilon)$ curves. Now, each subject's data can be represented as a point in the feature plane, as shown in Fig. 9.8. We observe that for healthy subjects, feature 1 is generally very small but feature 2 is large, indicating that the dynamics of the cardiovascular system is like a nonlinear system with stochasticity, with resolvable small-scale behaviors and well-defined characteristic scale $\overline{\varepsilon}$. The opposite is true for the patients with CHF: feature 1 is large but feature 2 is small, indicating that not only small-scale behaviors of the $\lambda(\varepsilon)$ curves cannot be resolved, but also that the characteristic scale $\overline{\varepsilon}$ is not well defined. Very interestingly, these two simple features separate completely the normal subjects from patients with CHF. In fact, each feature alone can almost perfectly separate the two groups of subjects studied here.

It is interesting to note that for the purpose of distinguishing normal HRV from CHF HRV, the features derived from SDLE are much more effective than other metrics including the Hurst parameter, the sample entropy, and multiscale entropy. For the details of the comparisons, we refer to [39].

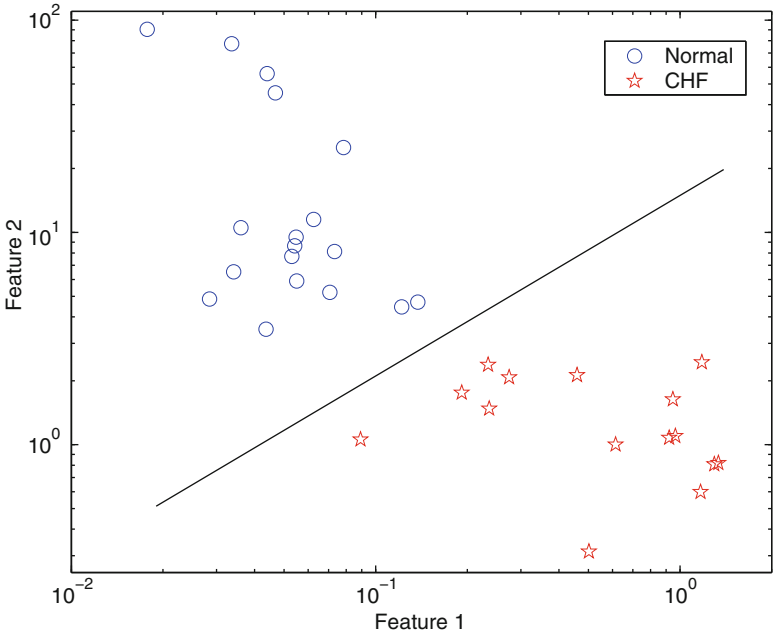

Fig. 9.8 Feature plane separating normal subjects from subjects with CHF

9.4 Concluding Remarks

In this chapter, we have discussed a multiscale complexity measure, the SDLE. We have shown that it can readily characterize low-dimensional chaos and random $1/f^{\alpha}$ processes, can readily detect intermittent chaos, can conveniently deal with nonstationarity, and can accurately detect epileptic seizures from EEG and distinguish healthy subjects from patients with CHF from HRV. More importantly, we have established that commonly used complexity measures can be related to the value of the SDLE at specific scales, and that the pattern of the SDLE around the characteristic scale $\overline{\varepsilon}$ contains a lot of useful information on the structured components of the data that may greatly help detect significant patterns. Because of the ubiquity of chaos-like motions and $1/f^{\alpha}$-type processes and the complexity of HRV and EEG data, our analyses strongly suggest that the SDLE is potentially important for clinical practice and provides a comprehensive characterization of complex data arising from a wide range of fields in science and engineering.

Our analyses have a number of important implications:

- To comprehensively characterize the complexity of complicated data such as HRV or EEG data, a wide range of scales has to be considered, since the complexity may be different on different scales. For this purpose, the entire $\lambda(\varepsilon)$ curve, where ε is such that $\lambda(\varepsilon)$ is positive, provides a good solution. This point

is particularly important when one wishes to compare the complexity between two signals—the complexity for one signal may be higher on some scales, but lower on other scales. The situation shown in Fig. 9.5 may be considered one of the simplest.

- For detecting important events such as epileptic seizures, $\lambda_{\text{small}-\varepsilon}$ and $\lambda_{\text{large}-\varepsilon}$ appear to provide better defined features than other commonly used complexity measures. This may be due to the fact that $\lambda_{\text{small}-\varepsilon}$ and $\lambda_{\text{large}-\varepsilon}$ are evaluated at fixed scales, while other measures are not. In other words, scale mixing may blur the features for events being detected, such as seizures.
- In recent years, there has been much effort in searching for cardiac chaos [14–18,58]. Due to the inability of unambiguously distinguishing deterministic chaos from noise by calculating the largest positive LE and the correlation dimension, it is still unclear whether the control mechanism of cardiovascular system is truly chaotic or not. Our analysis here highly suggests that if cardiac chaos does exist, it is more likely to be identified in healthy subjects than in pathological groups. This is because the dimension of the dynamics of the cardiovascular system appears to be lower for healthy than for pathological subjects. Intuitively, such an implication makes sense, because a healthy cardiovascular system is a tightly coupled system with coherent functions, while components in a malfunctioning cardiovascular system are somewhat loosely coupled and function incoherently.

As example applications, we have focused on the analyses of HRV and EEG data here. It is evident that SDLE will be useful for other kinds of complex data analyses, including financial time series and various kinds of physiological data. While much of the past as well as current research has been focused on determining whether some experimental data are chaotic or not, the scaling laws of SDLE suggest that it is often feasible to obtain the defining parameters of the data under study. That is, if the data is chaotic, then one should find out what kind of chaos it is; and if it is random, one can aim to find out what kind of random process that is, including its correlation structure. While in principle, SDLE is able to do so without pre-processing of the data under study, suitable detrending and denoising may help. A particularly simple and versatile procedure is the smooth adaptive filter developed by the authors, which has been successfully applied to recover chaos in heavy noise environment [62–64].

Acknowledgments This work is supported in part by NSF grants CMMI-1031958 and 0826119.

References

1. Fairley P (2004) The unruly power grid. IEEE Spectrum 41:22–27
2. Bassingthwaighte JB, Liebovitch LS, West BJ (1994) Fractal physiology. Oxford University Press, New York
3. Leland WE, Taqqu MS, Willinger W, Wilson DV (1994) On the self-similar nature of ethernet traffic (extended version). IEEE/ACM Trans Networking 2:1–15

4. Gao JB, Rubin I (2001) Multiplicative multifractal modeling of long-range-dependent network traffic. Int J Commun Syst 14:783–801
5. Gao JB, Rubin I (2001) Multifractal modeling of counting processes of long-range-dependent network traffic. Comput Commun 24:1400–1410
6. Mandelbrot BB (1997) Fractals and scaling in finance. Springer, New York
7. Gao JB, Hu J, Tung WW, Zheng Y (2011) Multiscale analysis of economic time series by scale-dependent Lyapunov exponent. Quantitative Finance DOI:10.1080/14697688.2011.580774
8. Mit-bih normal sinus rhythm database & bidmc congestive heart failure database available at http://www.physionet.org/physiobank/database/#ecg.
9. Akselrod S, Gordon D, Ubel FA, Shannon DC, Barger MA, Cohen RJ (1981) Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. Science 213:220–222
10. Kobayashi M, Musha T (1982) 1/f fluctuation of heart beat period. IEEE Trans Biomed Eng 29:456–457
11. Peng CK, Mietus J, Hausdorff JM, Havlin S, Stanley HE, Goldberger AL (1993) Long-range anticorrelations and non-gaussian behavior of the heartbeat. Phys Rev Lett 70:1343–1346
12. Ashkenazy Y, Ivanov PC, Havlin S, Peng CK, Goldberger AL, Stanley HE (2001) Magnitude and sign correlations in heartbeat fluctuations. Phys Rev Lett 86:1900–1903
13. Ivanov PC, Amaral L, Goldberger AL, Havlin S, Rosenblum MG, Struzik ZR, Stanley HE (1999) Multifractality in human heartbeat dynamics. Nature 399:461–465
14. Kaplan DT, Goldberger AL (1991) Chaos in cardiology. J Cardiovasc Electrophysiol 2: 342–354
15. Goldberger AL, West BJ (1987) Applications of nonlinear dynamics to clinical cardiology. Ann New York Acad Sci 504:155–212
16. Garfinkel A, Spano ML, Ditto WL, Weiss JN (1992) Controlling cardiac chaos. Science 257:1230–1235
17. Fortrat JO, Yamamoto Y, Hughson RL (1997) Respiratory influences on non-linear dynamics of heart rate variability in humans. Biol Cybernet 77:1–10
18. Kaneko K, Tsuda I (2000) Complex systems: chaos and beyond. Springer, Berlin
19. Grassberger P, Procaccia I (1983) Characterization of strange attractors. Phys Rev Lett 50: 346–349
20. Grassberger P, Procaccia I (1983) Estimation of the kolmogorov entropy from a chaotic signal. Phys Rev A 28:2591–2593
21. Wolf A, Swift JB, Swinney HL, Vastano JA (1985) Determining lyapunov exponents from a time series. Physica D 16:285–317
22. Sugihara G, May RM (1990) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature 344:734–741
23. Kaplan DT, Glass L (1992) Direct test for determinism in a time-series. Phys Rev Lett 68: 427–430
24. Gao JB, Zheng ZM (1994) Direct dynamical test for deterministic chaos and optimal embedding of a chaotic time series. Phys Rev E 49:3807–3814
25. Gao JB, Zheng ZM (1994) Direct dynamical test for deterministic chaos. Europhys Lett 25:485–490
26. Pei X, Moss F (1996) Characterization of low-dimensional dynamics in the crayfish caudal photoreceptor. Nature 379:618–621
27. Poon CS, Barahona M (2001) Titration of chaos with added noise. Proc Natl Acad Sci USA 98:7107–7112
28. Gaspard P, Briggs ME, Francis MK, Sengers JV, Gammons RW, Dorfman JR, Calabrese RV (1998) Experimental evidence for microscopic chaos. Nature 394:865–868
29. Dettmann CP, Cohen EGD (2000) Microscopic chaos and diffusion. J Stat Phys 101:775–817
30. Hu J, Tung WW, Gao JB, Cao YH (2005) Reliability of the 0-1 test for chaos. Phys Rev E 72:056207
31. Osborne AR, Provenzale A (1989) Finite correlation dimension for stochastic-systems with power-law spectra. Physica D 35:357–381

32. Provenzale A, Osborne AR (1991) Convergence of the k2 entropy for random noises with power law spectra. Physica D 47:361–372
33. Aurell E, Boffetta G, Crisanti A, Paladin G, Vulpiani A (1997) Predictability in the large: an extension of the concept of lyapunov exponent. Physica A 30:1–26
34. Gao JB, Hu J, Tung WW, Cao YH (2006) Distinguishing chaos from noise by scale-dependent lyapunov exponent. Phys Rev E 74:066204
35. Gao JB, Cao YH, Tung WW, Hu J (2007) Multiscale analysis of complex time series—integration of chaos and random fractal theory, and beyond. Wiley, New York
36. Gao JB, Hu J, Mao X, Tung WW (2012) Detecting low-dimensional chaos by the "noise titration" technique: possible problems and remedies. Chaos, Solitons, and Fractals 45:213–223
37. Gao JB, Tung WW, Hu J (2009) Quantifying dynamical predictability: the pseudo-ensemble approach (in honor of professor andrew majda's 60th birthday). Chinese Annal Math Ser B 30:569–588
38. Gao JB, Hu J, Tung WW (2011) Complexity measures of brain wave dynamics. Cogn Neurodyn 5:171–182
39. Hu J, Gao JB, Tung WW, Cao YH (2010) Multiscale analysis of heart rate variability: a comparison of different complexity measures. Annal Biomed Eng 38:854–864
40. Hu J, Gao JB, Tung WW (2009) Characterizing heart rate variability by scale-dependent lyapunov exponent. Chaos (special issue on Controversial Topics in Nonlinear Science: Is the Normal Heart Rate Chaotic? it is one of the most downloaded papers in that issue) 19:028506
41. Ryan DA, Sarson GR (2008) The geodynamo as a low-dimensional deterministic system at the edge of chaos. Europhys Lett 83:49001
42. Packard NH, Crutchfield JP, Farmer JD, Shaw RS (1980) Gemomtry from time-series. Phys Rev Lett 45:712–716
43. Takens F (1981) Detecting strange attractors in turbulence. In: Rand DA, Young LS (eds.) Dynamical systems and turbulence, lecture notes in mathematics, vol 898. Springer, Berlin, p 366
44. Sauer T, Yorke JA, Casdagli M (1991) Embedology. J Stat Phys 65:579–616
45. Gao JB, Zheng ZM (1993) Local exponential divergence plot and optimal embedding of a chaotic time series. Phys Lett A 181:153–158
46. Mackey MC, Glass L (1977) Oscillation and chaos in physiological control-systems. Science 197:287–288
47. Gao JB, Hu J, Tung WW, Cao YH, Sarshar N, Roychowdhury VP (2006) Assessment of long range correlation in time series: How to avoid pitfalls. Phys Rev E 73:016117
48. Mandelbrot BB (1982) The fractal geometry of nature. Freeman, San Francisco
49. Frisch U (1995) Turbulence—the legacy of A.N. Kolmogorov. Cambridge University Press
50. Lempel A, Ziv J (1976) On the complexity of finite sequences. IEEE Trans Inform Theor 22:75–81
51. Cao YH, Tung WW, Gao JB, Protopopescu VA, Hively LM (2004) Detecting dynamical changes in time series using the permutation entropy. Phys Rev E 70:046217
52. Martinerie J, Adam C, Quyen MLV, Baulac M, Clemenceau S, Renault B, Varela FJ (1998) Epileptic seizures can be anticipated by non-linear analysis. Nat Med 4:1173–1176
53. Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL (1994) On the mosaic organization of dna sequences. Phys Rev E 49:1685–1689
54. Hwa RC, Ferree TC (2002) Scaling properties of fluctuations in the human electroencephalogram. Phys Rev E 66:021901
55. Robinson PA (2003) Interpretation of scaling properties of electroencephalographic fluctuations via spectral analysis and underlying physiology. Phys Rev E 67:032902
56. Hon EH, Lee ST (1965) Electronic evaluations of the fetal heart rate patterns preceding fetal death: further observations. Am J Obstet Gynecol 87:814–826
57. Task Force of the European Society of Cardiology & the North American Society of Pacing & Electrophysiology (1996) Heart rate variability: Standards of measurement, physiological interpretation,and clinical use. Circulation 93:1043–1065

58. Babyloyantz A, Destexhe A (1988) Is the normal heart a periodic oscillator? Biol Cybern 58:203–211
59. Bigger JT, Steinman RC, Rolnitzky LM, Fleiss JL, Albrecht P, Cohen RJ (1996) Power law behavior of rr-interval variability in healthy middle-aged persons, patients with recent acute myocardial infarction, and patients with heart transplants. Circulation 93:2142–2151
60. Ho KKL, Moody GB, Peng CK, Mietus JE, Larson MG, Levy D, Goldberger AL (1997) Predicting survival in heart failure cases and controls using fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. Circulation 96:842–848
61. Pincus SM, Viscarello RR (1992) Approximate entropy: a regularity statistic for fetal heart rate analysis. Obst Gynecol 79:249–255
62. Hu J, Gao JB, Wang XS (2009) Multifractal analysis of sunspot time series: the effects of the 11-year cycle and fourier truncation. J Stat Mech. DOI: 10.1088/1742-5468/2009/02/P02066
63. Gao JB, Sultan H, Hu J, Tung WW (2010) Denoising nonlinear time series by adaptive filtering and wavelet shrinkage: a comparison. IEEE Signal Process Lett 17:237–240
64. Tung WW, Gao JB, Hu J, Yang L (2011) Recovering chaotic signals in heavy noise environments. Phys Rev E 83:046210

Chapter 10
Wavelet Analysis of ECG Signals

En-Bing Lin, Megan Haske, Marilyn Smith, and Darren Sowards

Abstract This study evaluated the effectiveness of different types of wavelets and thresholds to process electrocardiograms. An electrocardiogram, or ECG, shows the electrical activity in the heart and can be used to detect abnormalities. The first process used term-by-term thresholding to denoise ECGs. The second process denoised and compressed ECGs using global thresholding. The effectiveness was determined by using the signal-to-noise ratio (SNR) and the percentage root mean square difference (PRD).

10.1 Introduction

Electrocardiogram (ECG or EKG) signals are due to ionic current flows which cause the cardiac fibers to contract and relax, subsequently generating a time variant periodic signal. The ECG is a diagnostic tool that measures and records the electrical activity of the heart in great detail. Interpretation of these details allows diagnosis of a wide range of heart conditions which can vary from minor to life threatening. The term electrocardiogram was introduced by the Dutch physiologist, Willem Einthoven, in the 1890s and early 1900s. In 1924, Einthoven received the Nobel Prize for his life's work in developing the ECG [6]. By the 1950s, developments in ECG technology allowed medical professionals to observe electrical stimulated heart signals by placing electrodes in and around the heart muscle. More recently, the study of ECG signals in medical applications has become one of the fastest growing research topics. Proficiency in the interpretation of the ECGs is an essential skill for medical professionals. A single normal cycle of the ECG represents successive arterial depolarization and repolarization as well as ventricular depolarization and repolarization, which are due to cardiac contractions that occur with every

E.-B. Lin (✉) • M. Haske • M. Smith • D. Sowards
Department of Mathematics, Central Michigan University, Mt. Pleasant, MI 48859, USA
e-mail: lin1e@cmich.edu

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_10,

heartbeat. These are approximated by the peaks and troughs of the ECG waveform. It consists of well-defined, successive waves denoted P, Q, R, S, and T waves (see Fig. 10.4). Much attention has been paid to the adequate and accurate analysis of the ECG signal that would lead to correct diagnoses of cardiac anomalies. However, the picked-up ECG signal is corrupted by several sources of noise. This corrupted noise considerably prevents accurate analysis of the ECG signal and thereby preventing the potential for useful information extraction. Problematically, errors in reading are common, and may lead to serious consequences. In an ECG system, the potential difference between two electrodes placed on the skin surface is considered as an input to the ECG plotter. Statistical data from past research reveals that there is approximately 20–50 % discordance between the early ECG interpretation and the final interpretation by a senior cardiologist. The interpreted results from the ECG is evaluated by the doctor for the final diagnosis in deciding how to best administer urgent treatment for the ailing patients with life-threatening cardiovascular diseases. Unstable recording environment, spurious signals from nearby equipment, poor electrodes, and electromagnetic pollution are a few causes of unwanted noise contamination on the ECG signal. Results from laboratory and clinical studies suggest that the existence of abnormal ventricular conduction during sinus rhythm in regions surrounding a myocardial infraction, will generate delayed and fractionated micro potentials on the ECG signals. The morphology of ECG signal has been used for recognizing much variability in heart activity, so the establishment of parameters of an ECG signal clear of noise is of utmost importance. This gives a full picture complete with detailed information about the electrophysiology of the hearts diseases and any ischemic changes that may occur such as myocardial infarction, conduction defects, and arrhythmia. Different attempts have been made to design filtering algorithms aimed to improve the signal-to-noise ratio (SNR) values and recovering the ECG waves in noisy environments. ECG signal is considered as a non-stationary signal. An efficient technique for this nonstationary signal processing is the wavelet transform [14]. The wavelet transform can be used as a decomposition of a signal in the time-frequency scale plane. There are many applications of wavelet transform such as sub-band coding data compression, characteristic points detection, and noise reduction. In order to reduce the noise of ECG signals, many techniques including digital filters finite impulse response (FIR or IIR), adaptive method, and wavelet transform thresholding methods are available [7]. The goal of this chapter is to determine the optimal wavelet, order, level, and threshold for denoising and compressing an ECG while smoothing out and maintaining the integrity of the original signal. The wavelets used were: Daubechies, Biorthogonal Spline, Coiflet, and Symlet. Various thresholds were utilized: soft, hard, global, rigorous SURE, heuristic SURE, universal, and minimax. The SNR in combination with the percentage root mean square difference (PRD) helped determine the optimal conditions for wavelet denoising. Compression scores and L^2 norm recovery values determined the ideal conditions for wavelet compression. This report includes background information about wavelets and thresholding. The two processes and their results will be explained and analyzed.

10.2 Wavelet Analysis

Wavelets provide time and frequency analysis simultaneously and offer flexibility with many different properties and useful applications. In this section, we give a brief overview on wavelet analysis [3].

10.2.1 Convolution, Filters, and Filter Banks

Convolution is the process that if an input vector is processed through a filter, the result is an output vector. The output vector can be used for many applications, for example, to reconstruct the input vector. The output vector can be calculated as a series of shifting inner products:

$$y = h * x, \tag{10.1}$$

$$y_n = \sum_{k=-\infty}^{\infty} h_k x_{n-k}, \tag{10.2}$$

Where x is input data (signal), y is output data, h is filter, and * is convolution. A wide variety of filters can be used in the above equations. The filters used in this chapter were all FIR filters. An FIR filter is a casual filter with a finite number of elements that are nonzero [15]. FIR filters can further be classified as either high-pass filters or low-pass filters. A low-pass filter annihilates high oscillatory trends, or details, of a signal while maintaining low oscillatory trends, or approximations, of a signal. A high-pass filter maintains high oscillatory trends of a signal while eliminating the locally constant trends. The combination of a high-pass filter and low-pass filter is called a filter bank [13].

10.2.2 Multiresolution Analysis

Multiresolution analysis is based upon a sequence of approximation spaces (V_j) which must satisfy certain conditions. Let L^2 be the space of square-integrable functions. These conditions are [3]

$$\ldots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \ldots, \tag{Cond. 1}$$

$$\overline{\bigcup_{j \in Z} V_j} = L^2(R), \tag{Cond. 2}$$

$$\bigcap_{j \in Z} V_j = 0 \tag{Cond. 3}$$

Multiresolution is the idea that each subspace (V_J) is a scaled representation of V_0:

$$f \in V_j \Leftrightarrow f(2^j) \in V_0 : \qquad \text{(Cond. 4)}$$

The concept that an entire space can be represented using only scaled representations of a single function led to general expansion wavelet systems. In V_1, general expansion wavelet systems use a father function ($\phi(2t-n)$) and its scaled and dilated representations to form a basis. The relationship between V_0 and V_1 can be described by the dilation equation [3]:

$$\phi(t) = \sum_n h(n)\sqrt{2}\phi(2t-n). \qquad \text{(Cond. 5)}$$

10.2.3 Wavelet Systems

A wavelet system is comprised of a father function ϕ and a mother function ψ. In what follows, we will limit our discussion to discrete wavelet transforms and orthogonal wavelet systems with compact support. Wavelet systems, also called wavelet families, are rather unusual because their properties and conditions are not derived from an equation. Rather, a wavelet system derives its equation from a set of conditions and properties. One of the approaches in the design of a wavelet system is to determine the intended application of the system. The purpose of the system could be data compression, modeling, or denoising of signals. Once the purpose of the system has been determined, desirable conditions are set and the important properties are determined. These important properties can include:

- Compact Support
- Type of filter
- Length of filter
- Orthogonality
- Support Width
- Number of Vanishing Moments for ϕ and ψ
- Regularity

These conditions and properties lead to a set of equations. These equations are used to determine the dilation equation. The dilation equation is

$$\phi(t) = \sum_n h(n)\sqrt{2}\phi(2t-n). \qquad (10.3)$$

The dilation equation uses an FIR, low-pass filter ($h(n)$). A low-pass filter must be used in conjunction with a high-pass filter in order to have the most accurate representation of a signal. Therefore, a mother function ($\psi(t)$) must also be derived

from the dilation equation [15]. This is done by using the relationship between the low-pass filter ($h(n)$) and high-pass filter ($g(n)$) [13]:

$$g(n) = (-1)^n h(1-n). \tag{10.4}$$

This relationship leads to the mother function:

$$\psi(t) = \sum_n g(n)\sqrt{2}\phi(2t-n). \tag{10.5}$$

The dilation equation uses the father function, and the wavelet equation determines the mother function.

10.2.4 The Four Families

The Daubechies family (DBF) uses a general expansion wavelet system. A general expansion wavelet system is a system that is generated from a single scaling function $\phi(2t-n)$. This system forms an orthonormal basis with compact support. Daubechies set conditions for the number of zero moments (vanishing moments) for the mother function $\psi(t)$. The desire was to have the maximum number of vanishing moments for $\psi(t)$ in order to have increased smoothness for the mother function [3]:

$$\int dx x^l \psi(x) = 0, \quad l = 0, \ldots, L-1. \tag{10.6}$$

A high number of vanishing moments results in more accurate detail coefficients because these coefficients can now be almost zero where a function is smooth. Coiflets based on the following conditions:

$$\int dx x^l \psi(x) = 0, \quad l = 0, \ldots, L-1. \tag{10.7}$$

and

$$\int dx x^l \phi(x) = 0, \quad l = 1, \ldots, L-1. \tag{10.8}$$

One shortcoming of orthonormal wavelet bases is that the one FIR filter is used for deconstruction and its transpose is used for reconstruction. When this occurs, one wants to recover the signal after it is processed, the exact reconstruction of the original signal and symmetry of the FIR filters are impossible. The Cohen−Daubechies−Feauveau Biorthogonal Spline Family (BSF) was designed to overcome this shortcoming. If one FIR filter is used for deconstruction and a different FIR filter for reconstruction, symmetry of the filters is possible [3]. Therefore, the BSF is defined as follows.

Table 10.1 Key properties of four wavelet families

Wavelet Families Properties				
	Daubechies	**Symlet**	**Coiflet**	**Biorthogonal**
Order(N)	Any Positive Integer	2,3,…	1,2, …, 5	N_r, N_d
Orthogonal	Yes	Yes	Yes	No
Biorthogonal	Yes	Yes	Yes	Yes
Compact Support	Yes	Yes	Yes	Yes
Support Width	2N-1	2N-1	6N-1	$2N_r+1$, $2N_d+1$
Regularity	approx. 0.2N for large N			Nr-1 and Nr-2 at knots*
Vanishing Moments for ψ	N	N	2N-1	N_r(dec)

$$\psi(t) = \sum \sqrt{2} g_n \phi(2t - n), \tag{10.9}$$

$$\tilde{\psi}(t) = \sum \sqrt{2} \tilde{g}_n \phi(2t - n), \tag{10.10}$$

$$\phi(t) = \sum \sqrt{2} h_n \phi(2t - n), \tag{10.11}$$

$$\tilde{\phi}(t) = \sum \sqrt{2} \tilde{h}_n \phi(2t - n), \tag{10.12}$$

where $\phi(t)$ and $\psi(t)$ are the functions used for deconstruction and $\tilde{\phi(t)}$ and $\tilde{\psi(t)}$ are used for reconstruction. The order of deconstruction and reconstruction does not have to be the same. The relationship between the fathers' filter coefficients is

$$\left(\sum h_n\right)\left(\sum \tilde{h}_n\right) = 2. \tag{10.13}$$

The relationship between the low-pass and high-pass filters is maintained for reconstruction and deconstruction.

10.2.5 Key Properties

Table 10.1 shows some key properties of the wavelet families used in this chapter.

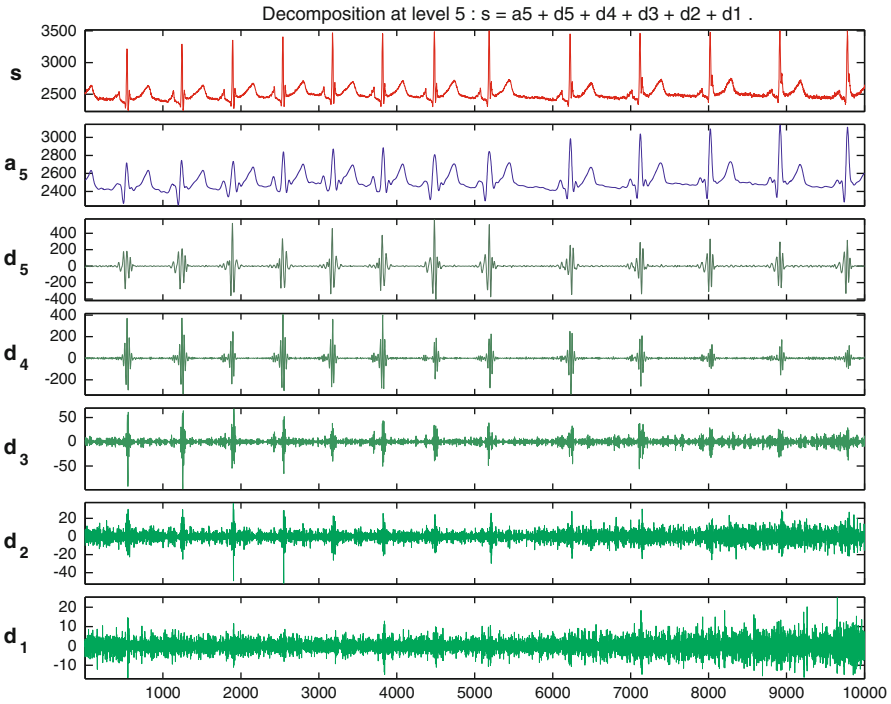

Fig. 10.1 Decomposition of a signal

10.2.6 Discrete Wavelet Transforms

10.2.6.1 Convolution Approach

Discrete wavelet transforms are defined by using the convolution of the highpass and low-pass filters with a signal to produce approximation coefficients and detail coefficients [15]:

$$a = h * x, \tag{10.14}$$

$$d = g * x. \tag{10.15}$$

where x is input data (signal), h is low-pass filter coefficient, a is approximation coefficient, g is high-pass filter coefficient, and d is detail coefficient. Since the approximation coefficients are obtained using a low-pass filter, which eliminates high oscillatory trends, they form a relatively close approximation of the signal. The combination of the approximation and the details of a signal are what make the signal unique. So, the signal must be convolved with a high-pass filter to obtain those important detail coefficients [13]. An example of the decomposition of a given signal into approximation and details is shown in Fig. 10.1.

10.2.6.2 Matrix Approach

Discrete wavelet transforms can also be represented as matrices. The wavelet transform matrix, shown below, changes from family to family since the filter coefficients of each family are different.

- Wavelet transform matrix

$$W_N = \left[\frac{Lo\_D}{Hi\_D}\right]. \tag{10.16}$$

The low-pass decomposition filter (Lo_D) and high- pass decomposition filter (Hi_D) represent the matrices formed by the filter coefficients ($h(n)$,$g(n)$). These matrices are combined in a block matrix (W_N) and are convolved with a signal to form a matrix (Y). Y represents the block matrix representation of the approximation and detail coefficients matrices.

- Deconstruction

$$Y = W_N * X. \tag{10.17}$$

If W_N is an orthogonal matrix, then the inverse is equivalent to the transpose of itself [15]. Thus, the reconstruction filters (Lo_R,Hi_R) are the transposed representations of the deconstruction filters:

$$W_N^{-1} = W_N^T = \left[\frac{Lo\_R}{Hi\_R}\right]. \tag{10.18}$$

By using the transpose of the filter coefficient matrix, the signal can be reconstructed by the combination of the approximation and detail coefficients.

- Reconstruction

$$X = W_N^T * Y. \tag{10.19}$$

10.2.7 Analysis and Synthesis Filter Banks

The figures in this subsection show the process that is used to transform signals using wavelet systems in conjunction with discrete wavelet transform. The symbol in Fig. 10.2, comprised of a downward arrow and the number two, indicates downsampling to the second degree. Downsampling to the second degree is the removal of roughly half the detail and approximation coefficients. Downsampling is performed after a signal is processed through the filters because half the coefficients have become redundant. Downsampling eliminates the redundant coefficients. The symbol in Fig. 10.3, comprised of an upward arrow and the number two, indicates upsampling to the second degree. Upsampling is the inserting of zeros in the place

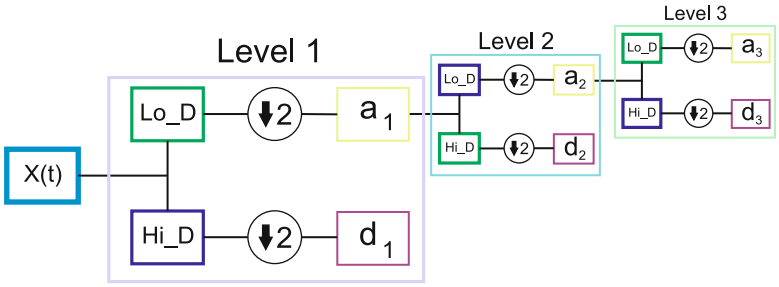

Fig. 10.2 Discrete wavelet transform deconstruction method

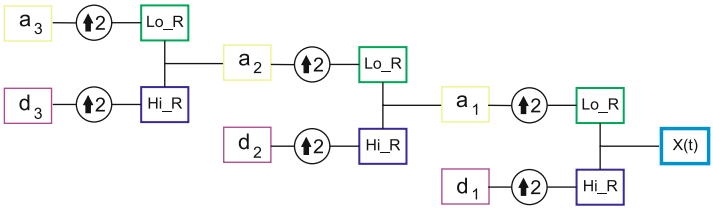

Fig. 10.3 Discrete wavelet transform reconstruction method

of the coefficients that were removed during downsampling. Subsampling is one of the factors that help wavelets being good for data compression. We usually use the approximation coefficients from the last level of deconstruction, along with all detail coefficients, to reconstruct a signal.

10.3 Electrocardiograms and How They Relate to Wavelets

The heart has cardiac cells that pass electrical impulses through the heart. These impulses regulate the pumping of the chambers (see Fig. 10.4). An electrocardiogram strip shows these electrical impulses as a signal. There are 12 different leads that show various perspectives of the heart. These leads come from different placements of electrodes over the body. An electrocardiogram is used to detect abnormalities in the heart since each part of the signal corresponds to a part of the movement of the impulse through the heart [4]. Different diseases can be diagnosed by looking at the differences in the length, frequency, and amplitude of each part of the wave. These factors depend on the voltage, speed, and path of the impulse through the heart's electrical system. Since each person is different, all of the previously mentioned things can vary from person to person (see Fig. 10.5) [4]. Electrocardiograms are biomedical signals, and like most of them, ECGs are non-stationary. Among different transform schemes, discrete wavelet transforms have shown promise because of their good localization properties in the time and frequency domain. Discrete wavelet transforms provide better performances than

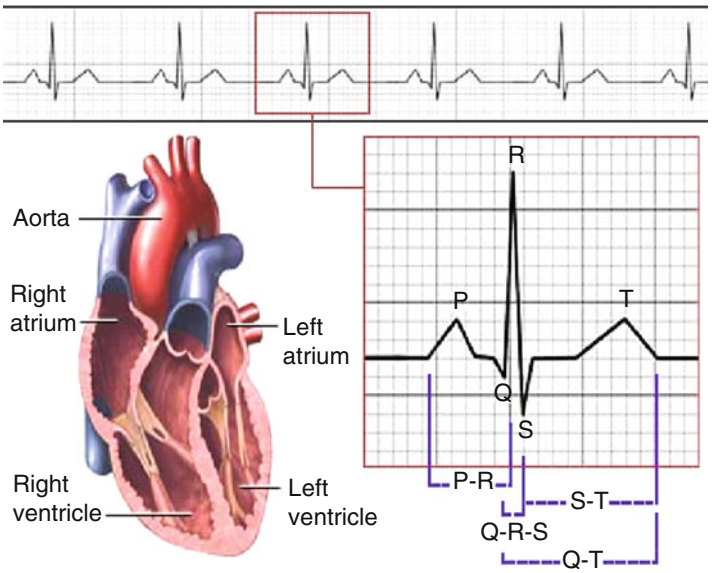

Fig. 10.4 Heart and ECG signal

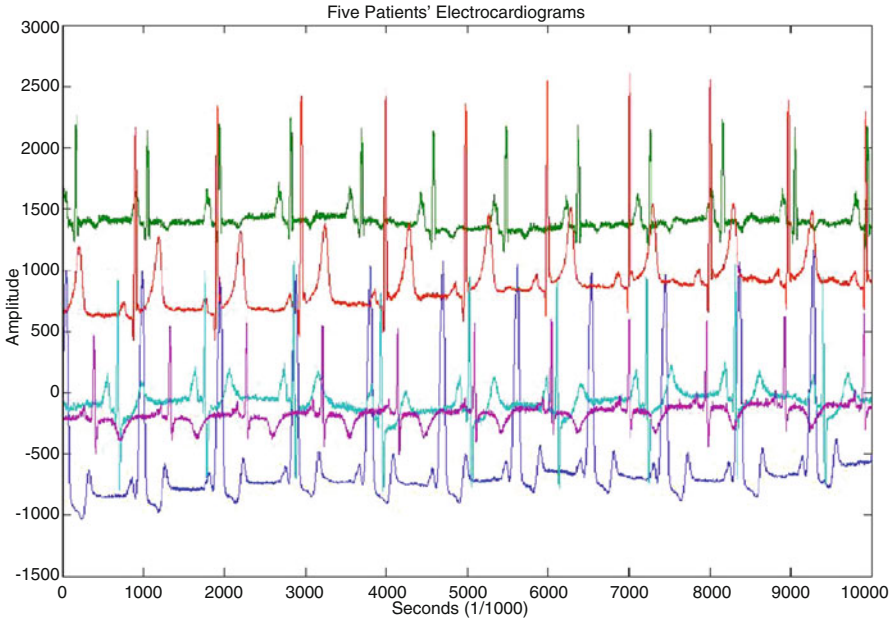

Fig. 10.5 ECG signals of five different patients

other transforms. Due to the compactness of supports and other properties described in Sect. 10.2.5, the local behavior of the highly nonstationary signals is expected to be well captured by wavelets than any other tools.

10.4 Main Results

In the context of data (or signals) approximation by using wavelets, one connects the smoothing or denoising of signals with the measures of smoothness depending on the magnitudes of their wavelet coefficients. Wavelet approximation using thresholding allows an adaptive representation of signal discontinuities. Applying thresholding in wavelet domain to problems of signal processing were used by Donoho [5] and many others. In this chapter, we explored two separate processes. The first process strictly denoised ECG signals using term-by-term thresholding. The second process denoised and compressed ECGs using global thresholding (GBL). The goal of both of these processes was to determine which wavelet and thresholding combination removed the most noise while smoothing out and maintaining the integrity of the signal. The best denoising combination was determined by using the SNR and PRD. An additional goal of the second process was to determine the best wavelet for compressing ECGs. The compression scores and L^2 norm recovery values were used to determine this additional goal. Our project used real ECGs from the PhysioNet PTB Diagnostic ECG Database (http://www.physionet.org/cgi-bin/ATM) [8].

10.4.1 Process 1

A group of male patients ranging from ages 43–63 were selected. Eight patients were healthy controls, and eight had suffered myocardial infarctions (heart attacks). The following wavelet families were used for both part one and two: Daubechies, Biorthogonal Spline, Coiflet, and Symlet. Part one used soft thresholding in combination with rigorous SURE, heuristic SURE, universal, and minimax thresholding rules. We denoised each patients' ECG signal varying the wavelet and thresholding. We evaluated each signal up to level 10 and varied the orders of the wavelet families.

10.4.1.1 Denoising

Signal interference (noise) can mask the true image of a signal. Denoising is performed to remove the unwanted and unnecessary noise. When deconstructing a signal using wavelets, the majority of the noise is isolated in the detail coefficients. The formula,

$$s_0(n) = s_r(n) + e(n) \tag{10.20}$$

describes how an observed signal is a pure signal and the noise that interferes with it. To get rid of this noise, three steps are utilized:

- Denoising via wavelet deconstruction
- Thresholding
- Inverse discrete wavelet transform (reconstruction)

The first and third steps are determined solely based on the wavelet family; the middle step is a procedure completely independent from the wavelet family and is the essential part of the denoising procedure [1].

Once a signal has been decomposed into its detail and approximation coefficients, it is now in a state where it can have thresholding imposed upon it. Two kinds of thresholding are used extensively: hard and soft. Hard thresholding is more rigid, whereas soft thresholding is smooth. For wavelet thresholding, there is a thresholding value, λ, which acts as a standard for the values of the detail coefficients. If the values do not meet the standards, that is, they are outside the λ, they are automatically set to zero [15]. Hard thresholding can be expressed as follows:

$$s'(x) = \begin{cases} d_{i,j} & if |d_{i,j}| > \lambda \\ 0 & if |d_{i,j}| \leq \lambda. \end{cases} \tag{10.21}$$

It should be noted that hard thresholding creates discontinuities at any $d_{i,j}$ equal to or less than the defined λ. These discontinuities create a more jagged signal, which is undesirable when denoising ECGs. A more desirable method to use when denoising ECGs is soft thresholding. Soft thresholding can be described as follows [7]:

$$d'_{i,j} = \begin{cases} d_{i,j} - \lambda & if \ \ d_{i,j} > \lambda \\ 0 & if \ |d_{i,j}| \leq \lambda \\ d_{i,j} + \lambda & if \ d_{i,j} < -\lambda. \end{cases} \tag{10.22}$$

As with hard thresholding, detail coefficients less than the threshold are scaled to zero. If the other conditions are met, then the corresponding coefficients are shrunk according to λ. Overall, this method creates a smooth signal which is the main reason soft thresholding is utilized in process one.

10.4.1.2 Term-by-Term Thresholding

To calculate λ, one must use thresholding rules. The rules described in the following methods are level dependent. They calculate a different λ from level to level. The first method is rigorous SURE (RIG), and it uses Stein's Unbiased Risk Estimator to minimize the error [15]. Universal thresholding (UNI) is another known method

which usually uses a higher thresholding value and often results in over smoothing. The universal threshold can be expressed as:

$$\lambda_{\text{univ}} = \sigma\sqrt{2\log(N)}, \tag{10.23}$$

where σ is the standard deviation of the noise and N is the signal length [1]. Minimax thresholding realizes the minimum of set of functions for the mean squared error. The final method used is heuristic SURE, which is just a combination of universal and SURE [11]. Once the signal has been denoised, it is ready to be reconstructed. Reconstruction is accomplished via inverse discrete wavelet transform, using all detail coefficients and only the last level's approximation coefficients [2]. The intention of denoising is to have a reconstructed signal which still depicts the important trends of the signal (i.e., accurately keeps the peaks and troughs of an ECG) [1].

10.4.1.3 Comparison Method 1

The SNR for each reconstructed signal was compared to the original SNR to see how much noise had been removed. The SNR was defined as:

$$\text{SNR} = \frac{\mu}{\sigma}. \tag{10.24}$$

The mean is represented by μ and the standard deviation by σ. Using the above definition, the UNI rule was found to have removed the most noise. The Daubechies wavelets of orders 9 and 10, level 10 removed the most noise when used in conjunction with UNI; Biorthogonal Spline order 3.1, level 10 removed the least amount of noise. The above definition of SNR was not an appropriate comparison method for ECG denoising. When too much noise is removed, the shape of the ECG is changed. When this occurs, an abnormality may not be able to be detected anymore which defeats the purpose of an ECG. For example, Fig. 10.6 shows a denoised signal that had too much noise removed compared to the original signal. Figure 10.6 shows another denoised signal that maintains the shape of the signal better. In fact, the differences between normal and abnormal ECGs are determined by things such as the length, frequency, and amplitude of each part of the waves.

10.4.1.4 Comparison Method 2

In order to overcome the problem of removing too much noise and changing the shape of the ECG signal, a different comparison method was used. This second method used a different definition of the SNR in conjunction with the PRD. For every level at every order, an individual PRD and SNR value was computed for the four different thresholding rules.

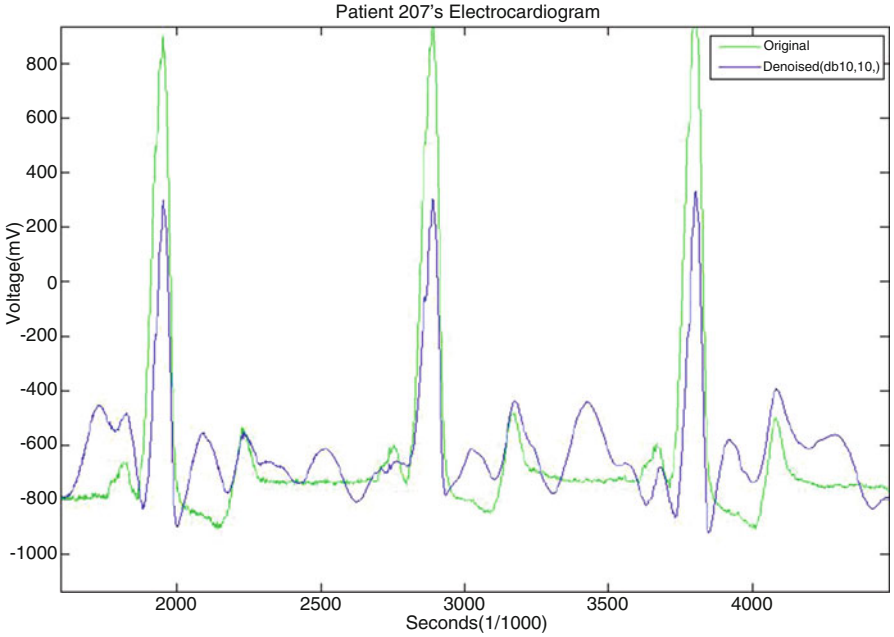

Fig. 10.6 This level of denoising changes the signal shape too much

The signal-to-noise ratio, or SNR, indicates how much noise has been removed. The equation for the SNR is [2]:

$$\text{SNR} := \log_{10}\left(\frac{\sum_{n=0}^{k} s_r^2(n)}{\sum_{n=0}^{k} e^2(n)}\right), \tag{10.25}$$

$$e(n) = s_0(n) - s_r(n), \tag{10.26}$$

- s_0: original signal
- s_r: reconstructed signal
- $e(n)$: noise

The percentage root mean square difference, or PRD, indicates how close the denoised signal is to the original. The equation for the PRD is [2]:

$$\text{PRD} = \sqrt{\frac{\sum_{n=0}^{N} (s_0(n) - s_r(n))^2}{\sum_{n=0}^{N} (s_0(n))^2}} \times 100\,\%. \tag{10.27}$$

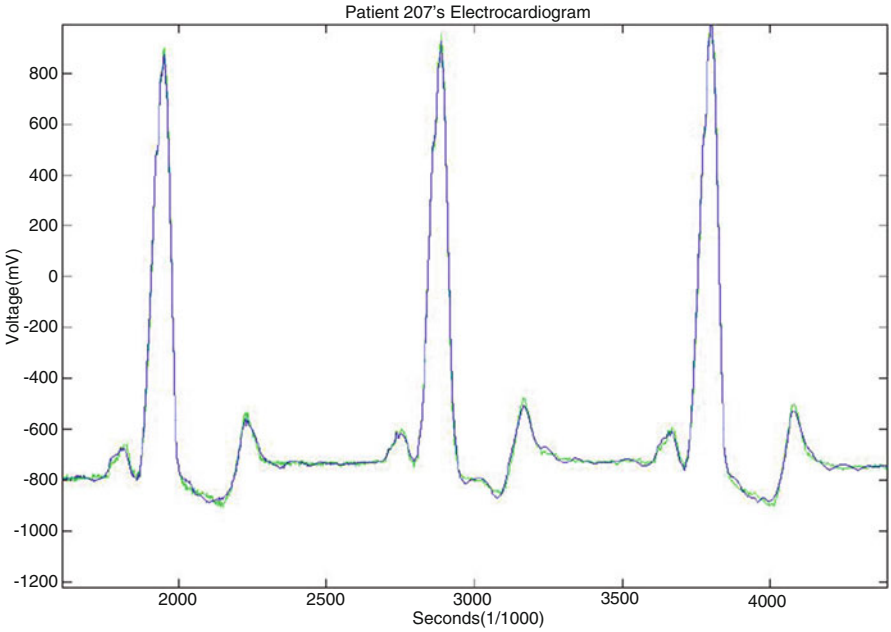

Fig. 10.7 Optimal denoising level maintains the shape of the signal

Table 10.2 The SNR and PRD values found for patient 252

Patient 252							
Method 1 (Coiflet,Rigorous)							
Order	Level	SNR	PRD	Order	Level	SNR	PRD
4	1	5.455	0.107	5	1	5.615	0.096
	2	5.045	0.183		2	5.115	0.171
	3	4.745	0.248		3	4.746	0.246
	4	4.628	0.280		4	4.579	0.291
	5	4.519	0.316		5	4.525	0.313
	6	4.337	0.398		6	4.393	0.368
	7	3.941	0.679		7	4.036	0.603
	8	3.894	0.715		8	3.979	0.645
	9	3.786	0.781		9	3.859	0.711
	10	3.751	0.824		10	3.825	0.745

The optimal SNR and PRD need to be reasonably small so that the most noise is removed without distorting the shape of the signal. The denoised signal in Fig. 10.6 has a smaller SNR and bigger PRD than the denoised signal in Fig. 10.7. Notice in Table 10.2, as the level increases, more noise is removed.

10.4.1.5 Results for Process 1

In general, the ranking of the families' ability to achieve the optimal result went: Coiflet, Symlet, Daubechies, and then Biorthogonal Spline. The best balance between the SNR and the PRD was at levels 4–6. At higher levels, a denoised signal starts to deviate too much from its original. An attempt to find an appropriate range for the optimal SNR and PRD was made and is explained in the next section.

10.4.1.6 Maximum PRD and Minimum SNR Selection

Recall, the denoised ECGs were analyzed using the SNR and PRD [2]. Plotting the PRD and SNR values on a scatter plot as x and y, respectively. Indeed, plotting a curve of best fit to PRD vs. SNR reveals an inverse relationship: while PRD increases, SNR decreases [2]. Since the relationship between PRD and SNR is an inverse one, it can be reasoned that there is a limit of maximum PRD and minimum SNR. This limit reflects the highest order one can go in denoising a signal before losing the characteristics of the true signal. We reasoned that there exists a function to reveal this limit, which intersects the PRD-SNR curve of best fit. The PRD-SNR curve is f, and the limiting curve of intersection is g. Intuitively, for g to intersect f, g needs to increase where f decreases. The point of intersection between f and g is the for mentioned limit. To create g, PRD was plotted as the independent variable and a function g of the ratio between PRD and SNR as the dependent variable on a scatter plot, which revealed a trend. Let $I_1 = [1, \infty)$. The function g is:

$$g = \mathrm{e}^{1+\frac{\mathrm{PRD}}{\mathrm{SNR}}}, \quad \forall \mathrm{SNR}_{(\mathrm{rigrsure,heursure})} \in I_1, \tag{10.28}$$

$$g = \mathrm{e}^{\frac{\mathrm{PRD}}{\mathrm{SNR}}}, \quad \forall \mathrm{SNR}_{(\mathrm{universal,minimax})} \in I_1. \tag{10.29}$$

One should note several conditions of the above statements. For each thresholding rule, there is a different intersecting function, with the only difference being an addition of one to the quotient of the first equation. This is because of the larger values of PRD and SNR produced by the universal and thresholding rules. Both functions are only defined for SNR values equal to or greater than 1. This can be reasoned by simply looking at the quotient in the natural exponent: the number created by SNR$<$1 increases far too fast to create an accurate line of best fit. Therefore, a second set of functions for g was created to include SNR$<$1, and dilute the larger numbers and yield similarly desired results. Let $I_2 = (-\infty, 1)$

$$g = \mathrm{e}^{\frac{1}{2}+\frac{\mathrm{PRD}}{10(\mathrm{SNR})}}, \quad \forall \mathrm{SNR}_{(\mathrm{rigrsure})} \in I_2, \tag{10.30}$$

$$g = \mathrm{e}^{\frac{\mathrm{PRD}}{10(\mathrm{SNR})}}, \forall \mathrm{SNR}_{(\mathrm{heursure,\ universal})} \in I_2, \tag{10.31}$$

$$g = \mathrm{e}^{1+\frac{\mathrm{PRD}}{10(\mathrm{SNR})}}, \forall \mathrm{SNR}_{(\mathrm{minimax})} \in I_2, \tag{10.32}$$

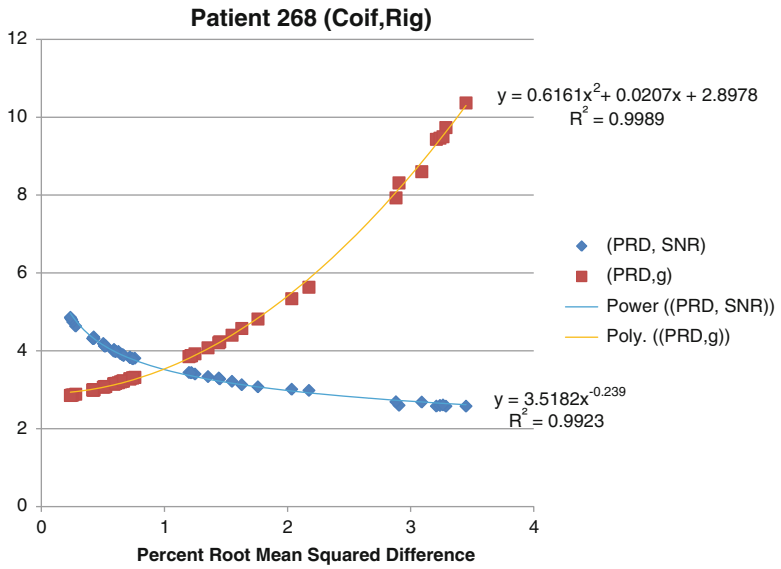

Fig. 10.8 The two scatter plots

Table 10.3 Data for the two scatter plots

Patient 268, Coiflet, RigorousSURE			
Level	**PRD**	**SNR**	**g**
1	0.239148	4.836122	2.856081
2	0.420638	4.322654	2.996096
3	0.519061	4.128506	3.082454
4	0.603438	3.986385	3.162538
5	0.665243	3.895566	3.224473
6	0.742455	3.801270	3.304608
7	1.354124	3.337925	4.078292
8	2.033773	3.012680	5.339175
9	3.28652	2.576317	9.734401
10	3.449886	2.577234	10.36671

To find the point of intersection, the two scatter plots described above must be plotted on the same graph using their shared independent variable PRD. The point of intersection between f and g reveals the maximum PRD and minimum SNR (Fig. 10.8). As long as the level of deconstruction is within these limits, it can be used to accurately denoise a signal. Within each threshold and order, the level cap was determined to be six with the PRD exceeding the limit at level 7. Therefore, the limit is level 6, which can be observed by denoising a signal and noticing how level 7 actually starts to create peaks and deviates from the actual signal. We comment that the above functions g may be replaced by more accurate approximate functions by using advanced interpolation techniques.

10.4.2 Process 2

While exploring the available built-in wavelet functions in MATLAB, we found a second available method for denoising electrocardiograms. We designed an m file that utilized two built-in wavelet functions. These functions performed two major tasks in denoising and compressing ECGs using GBL. GBL, the MATLAB functions, and our process will all be discussed in the following sections.

10.4.2.1 Global Thresholding

GBL is different from the term-by-term thresholding rules used in process 1. GBL uses a single thresholding value to denoise or compress a signal. This method of thresholding is commonly referred to as block thresholding. The disadvantages to term-by-term thresholding are that, when using wavelets, the thresholding values are determined without taking into consideration the terms' neighbors. This thresholding method is not optimal for wavelets because it requires a trade-off between variance and the mean squared error. These disadvantages led to the GBL method which is able to determine the best thresholding value for an entire neighborhood of coefficients. This type of thresholding method increases the adaptability of wavelet thresholding since the thresholding value is not dependent upon individual terms which can vary wildly in noisy signals. Since block thresholding takes into consideration a term's neighboring coefficients, the threshold tends to be more receptive to the jumps and skips of the original, noisy signal. Not only does it maintain a signal's properties better, this thresholding method preserves peaks better as well, even at the highest levels [9]. To select the threshold that minimizes error, the formula

$$\lambda_{jb} = [n^{-1} f(x_{jb})]^{1/2} \tag{10.33}$$

depicts the global threshold, where x_{jb} is such that $2^j x_{jb}$ lies in the middle of the block B_b, j is the resolution, and b is the block index [9].

10.4.2.2 Steps of Process 2

Based on the descriptions from the above subsection, we have the following steps. To begin process 2, we created a first MATLAB function and used the patient's ECG data to determine if soft or hard thresholding should be used, the number of approximation coefficients that should be kept, and determined the best thresholding value. We inputted the patient's ECG data and whether or not we would be denoising or compressing the signal. The outputted values were then stored and transferred to a second function. Note, the only major difference between compression and denoising is the type of threshold; compression uses hard thresholding and denoising uses soft thresholding. In step 2, we used a second MATLAB function and

the values obtained in step 1 to compress the ECG. The output was a compressed version of the ECG. The second MATLAB function used GBL and the thresholding value obtained in step 1. In step 3, the compressed ECG was entered into the first MATLAB function. This time, we indicated that we were denoising the compressed ECG. The output was then entered into the second MATLAB function which denoised the compressed ECG, once again using GBL. The denoised signal was then assessed using four values: SNR, PRD, compression score, and L^2 norm recovery.

10.4.2.3 Comparison Method 3

The SNR and PRD, as defined in comparison method 2, were once again used to determine which wavelet, order, and level achieved the optimal results for denoising the ECGs. In order to compare how well the ECGs were compressed, two new values had to be introduced: compression score and L^2 norm recovery. The compression score (PERF) represents the percentage of zero coefficients used during reconstruction. The significance of zero coefficients was discussed in Sect. 10.4.2. The closer the compression score is to 100 % the better. It is defined as

$$\text{PERF} = 100\,\% \times \frac{Z_n}{C_n}. \tag{10.34}$$

- Z_n = number of zero coefficients at current level
- C_n = number of coefficients

The L^2 norm recovery value (PERFL2) shows how close a compressed signal is to the original signal. The closer the L^2 norm recovery value is to 100 %, the closer the signal is the original:

$$\text{PERFL2} = 100\,\% \times \frac{(\text{vector-norm}(\text{coeffs of the current decomposition}, 2)^2}{(\text{vector-norm}(\text{orginal signal}, 2))^2}. \tag{10.35}$$

10.4.2.4 Results

The results obtained in process 2 indicated only minor differences in the PRD and SNR values for each family. Overall, the results for each family were very good when comparing the denoising indicators: SNR and PRD. The rankings were:

- Coiflet Family
- Symlet Family
- Daubechies Family
- Biorthogonal Spline Family

Table 10.4 Comparison results of using different wavelets

Patient 268 Data Results						
	Method 1		Method 2			
	Coiflet (P268,5,UNI)		**Coiflet (P268,5,GBL)**			
Level	PRD	SNR	PRD	SNR	PERF	PERFL2
1	0.314	4.370	0.373	4.165	49.830	99.999
2	0.538	3.964	0.569	3.935	73.851	99.996
3	0.785	3.562	0.687	3.777	84.850	99.992
4	1.065	3.253	0.766	3.713	89.478	99.986
5	1.493	2.934	0.812	3.673	91.309	99.977
6	2.053	2.691	0.846	3.639	91.872	99.966
7	4.241	2.166	0.869	3.619	91.886	99.960
8	4.755	2.052	0.880	3.605	91.870	99.959
9	5.606	1.890	0.889	3.593	91.843	99.959
10	5.769	1.882	0.892	3.586	91.817	99.964
	Bior(P268,2.4,RIG)		**Bior(P268,2.4,GBL)**			
Level	PRD	SNR	PRD	SNR	PERF	PERFL2
1	0.250	4.752	0.350	4.205	49.970	99.999
2	0.421	4.342	0.534	4.004	73.692	99.997
3	0.521	4.147	0.646	3.855	83.919	99.992
4	0.601	4.000	0.709	3.788	88.109	99.987
5	0.655	3.903	0.741	3.741	89.614	99.978
6	0.709	3.837	0.763	3.714	90.150	99.966
7	1.221	3.396	0.778	3.704	90.149	99.959
8	1.542	3.198	0.786	3.698	90.098	99.959
9	3.005	2.603	0.792	3.694	90.037	99.962
10	3.066	2.620	0.794	3.693	89.965	99.969

A few exceptions obtained bad results, they were Biorthogonal Spline of orders 1.1, 1.3, and 1.5; Daubechies order 1; and Symlet of order 5. These wavelets did not obtain results as well as all other wavelets. All levels obtain good results in both denoising and compression, but levels 4 and 5 had the optimal SNR to PRD relationship. Biorthogonal Spline order 1.5 and level 5 consistently obtained the best PERF and PERFL2 for the different ECGs. It obtained PERF scores between 91 % and 93 %. Its PERFL2 were around 99 % to almost 100 %.

10.5 Process 1 vs. Process 2: The Result

The result of our study clearly showed that process 2 using GBL was far superior to process 1 which used four different term-by-term thresholdings. The table in this section shows the comparison data results for patient 268. The sample data set shown uses the Biorthogonal Spline order 2.4 (Bior2.4) with rigrsure thresholding, which obtained the worst results during process 1. The data set for the same wavelet using GBL is shown on the right. Coiflet order five (Coif5) with UNI, which obtained the best results in process 1 and its compliment using GBL, is also shown.

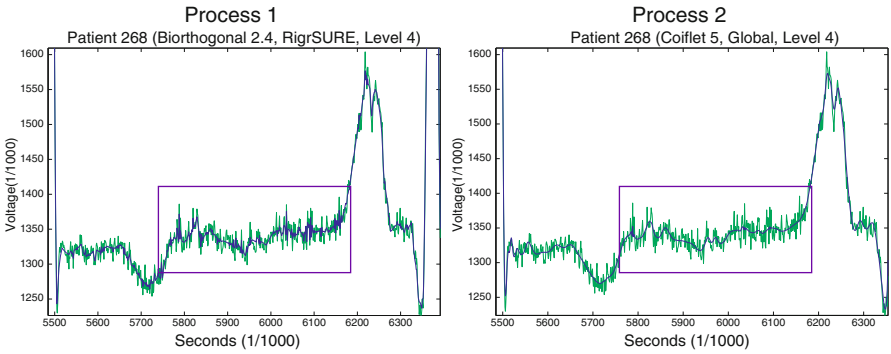

Fig. 10.9 The middle highlight of level 4 denoising

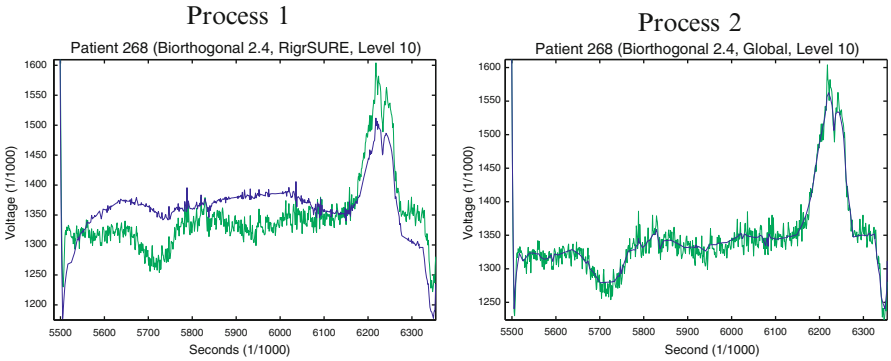

Fig. 10.10 Denoising at level 10 by using Biorthogonal 2.4

10.5.1 The Worst Case

The highlighted levels 4 and 10 are shown in the following figures (Fig. 10.9). For process 1, level 4 denoisings were found to be the best signal therefore, it is not surprising that Bior2.4 level 4 relatively meets our requirements. However, if you look at the boxed region, you can see how GBL is better able to smooth away the noisy regions while maintain, the integrity of the signal. The next figure shows the Bior2.4 level 10 denoising for process 1 and process 2. Level 10 denoisings obtained the worst results in process 1. The superiority of process 2, GBL, is clearly evident (Fig. 10.10).

10.5.2 The Best Case

Coiflet order 5 level 4 with rigrsure thresholding obtained the best results for process 1. This best case is compared to process 2 which obtained better results

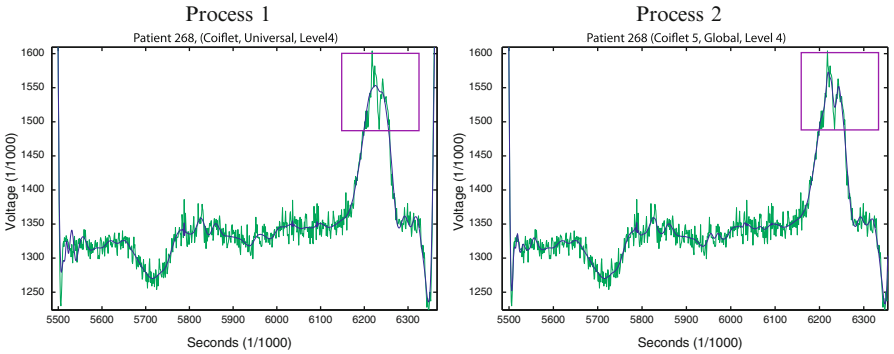

Fig. 10.11 The highlight of a singular part of the denoising at level 4

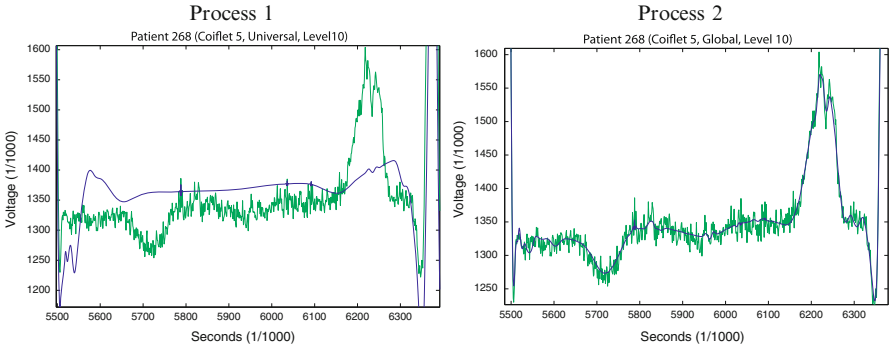

Fig. 10.12 Denoising at level 10 by using Coiflet

with GBL. Note the boxed region which contains one of the major peaks in an ECG. The amplitude and general shape of a peak can tell a cardiologist a lot about the condition of your heart, so it is important to maintain it. GBL clearly maintains this important region. The following figure of level 10 shows that GBL is able to maintain a clearer image of the ECGs at higher levels of decomposition which the thresholdings in process 1 were unable to do (Figs. 10.11 and 10.12).

10.6 Summary

For denoising ECG using term-by-term thresholds, Coiflet order five at level 4 decomposition is the best. When using any wavelet family or order, levels 4 and 5 are the best and levels 7 through 10 should not be used. When using GBL, any wavelet family can be used along with any level to denoise and compress an ECG. Levels 4 and 5 obtained the best, PRD and SNR relationship. In conclusion, we have determined GBL is the best thresholding to use when denoising electrocardiograms.

There are other developed systematic and computationally efficient procedures for analyzing multivariate nonstationary signals such as the method developed in [12]. Wavelet techniques are capable of revealing aspects of data that other time-frequency analysis techniques miss, that is aspects like trends, breakdown points, discontinuities in higher derivatives, and self-similarities [10]. With the choices of wavelets studied in this chapter, one can further evaluate the wavelet coefficients of ECG signals to obtain some comparison results between healthy and unhealthy patients.

Acknowledgements This research was conducted as part of the Central Michigan University LURE program during 2009–2011 and was supported by NSF-REU grant # 0606-36528. The authors are grateful for the support and would like to thank the anonymous referees' helpful comments as well.

References

1. Ahmaidi S, Bereksi-Reguig F, Chouakri SA, Fokapu O (2005) Wavelet denoising of the electrocardiogram signal based on the corrupted noise estimation. Comput Cardiol 32: 1021–1024
2. Alfaouri M, Daqrouq K (2008) ECG signal denoising by wavelet transform thresholding. Am J Appl Sci 5(3):276–281
3. Daubechies I (1992) Ten lectures on wavelets. Society for Industrial and Applied Mathematics, Philadelphia, PA
4. Davis D (2005) Quick and accurate 12-lead ECG interpretation, 4th edn. Lippincott Williams and Wilkins Philadelphia, PA
5. Donoho DI (1997) Cart and best-ortho-basis: a connection. Ann Statist 25(5):1870–1911
6. http://www.wikidoc.org/index.php/Willem_Einthoven.
7. Filho RG, Kozakevivius A, Nunes RC, Rodrigues CR Adaptive ECG filtering and QRS detection using orthogonal wavelet transformation. Retrieved July 1, 2010 from http://www.uninova.pt/~jmf/Alice
8. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 101(23):e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/full/101/23/e215].
9. Hall P, Penev S, Kerkyacharian G, Picard D (1997) Numerical performance of block thresholded wavelet estimators. Stat Comput 7:115–124
10. Mallat Stephane (2008) A wavelet tour of signal processing, 3rd edn. Academic Press
11. MATLAB (Version 7.8.0.347(R2009a)) [Computer software], (February 12, 2009), The MathWorks.
12. Ombao H, von Sachs R, Guo W (2005) SLEX analysis of multivariate nonstationary time series. J Am Stat Assoc 100(470):519–531
13. Strang G, Nguyen T (1997) Wavelets and filter banks. Wellesley-Cambridge Press, Cambridge
14. Singh BN, Tiwari AK (2006) Optimal selection of wavelet basis function applied to ECG signal denoising. Digital signal Process (16):275–287
15. Van Fleet PJ (2008) Discrete wavelet transformations an elementary approach with applications. John Wiley and Sons, Hoboken, NJ

Chapter 11
Multiscale Signal Processing with Discrete Hermite Functions

Dale H. Mugler and Anandi Mahadevan

Abstract Discrete Hermite functions (DHf) provide a new way of analyzing digital signals. As opposed to cumbersome computational methods that can only construct orthogonal discrete Hermite functions effectively for a small number of indices, there is a new method of computing DHf that is fast and efficient. Signal processing techniques that have been applied using the continuous Hermite functions (CHf) can now be adapted to the digital case, using this orthonormal set of DHf that share many of the properties of the continuous CHf. For some time, a multiscale version of the CHf has been available for analysis and has been applied to different kinds of signals and shown to be related to receptive fields of neurons. In this chapter, we explore the application of the digital DHf in multiscale analysis, showing analogies to the multiscale analysis provided by the CHf.

11.1 Introduction

The *continuous* Hermite functions (CHf) have been used in a number of important applications, including those for multiscale analysis, for example, in [16]. The CHf have many interesting properties, including their relation to Gaussians and their property of being eigenfunctions of the Fourier transform. The properties of the CHf, [13], and their use in multiscale analysis are important for this chapter, since this chapter widens the scope of their applications to involve the discrete Hermite functions (DHf), as defined by Mugler and Clary in [2, 3]. Simply sampling the

D.H. Mugler (✉)
University of Akron, Akron, OH 44325, USA
e-mail: dmugler@uakron.edu

A. Mahadevan
University of Akron, Akron, OH 44325, USA

Philips Home Healthcare Solutions, Murraysville, PA 15668 USA
e-mail: anandi.mahadevan@gmail.com

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_11,

CHf over a finite discrete domain does not result in a set of vectors that retain orthogonality or any of the other of the many properties of the CHf. Defined as the eigenvectors of a specific tridiagonal matrix [2, 11], the new set of DHf not only retain the shapes of the CHf but are also mutually orthonormal and are eigenvectors of a shifted (centered) Fourier matrix.

Multiscale analysis in this chapter refers to the ability, familiar from the theory of wavelets, to zoom in on features in a signal, moving from a coarse approximation to include details at several different levels. In particular, it provides a decomposition of the input signal into an approximation signal and detail signals at several different levels. That decomposition can be reversed to produce the original signal, with more details on this described in Sect. 11.3. Multiscale analysis allows one to isolate events that may happen in a very brief interval of time.

The shapes of the CHf have inspired a number of applications, for example, [14]. It is also important for applications of CHf that every function has a representation in terms of the CHf, since they form an orthonormal set of functions in the space of square-integrable functions over the whole real line $L^2(R)$. This property is also true for the DHf.

The usual introduction to the CHf begins with Hermite polynomials, $H_n(x)$. These are a classic set of orthogonal polynomials, and there are two slightly different ways of defining them. One form involves a monic set of polynomials, while the other form results in the leading coefficient being a power of 2. They can be defined either as $H_n(x) = (-1)^n \mathrm{e}^{x^2/2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2/2}$ for $n \geq 0$ for the monic polynomial form or as

$$H_n(x) = (-1)^n \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2} \tag{11.1}$$

for the case with leading coefficient being a power of 2. From the definition in (11.1), it follows that $H_0(x) = 1, H_1(x) = 2x, H_2(x) = 4x^2 - 2, \ldots$. Each of these forms is a rescaling of the other, so that the choice of form is not overly essential. We choose the second approach.

As is the case for classical sets of orthogonal polynomials, the Hermite polynomials satisfy a three-term recurrence relation, which is $H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$, with $H_0 = 1$ and $H_{-1} = 0$. More importantly, the CHf $h_n(x)$ are each defined as a normalized Gaussian multiple of the corresponding $H_n(x)$; in particular, the CHf are defined for $n \geq 0$ by

$$h_n(x) = \frac{1}{\pi^{1/4}\sqrt{2^n n!}} \mathrm{e}^{-x^2/2} H_n(x). \tag{11.2}$$

Multiplying the Hermite polynomials makes the CHf so that they are essentially of finite support, although the length of that support increases as n increases. These functions are orthonormal, in the sense that

$$\int_{-\infty}^{\infty} h_n(x) h_m(x) \mathrm{d}x = \delta(n-m) \tag{11.3}$$

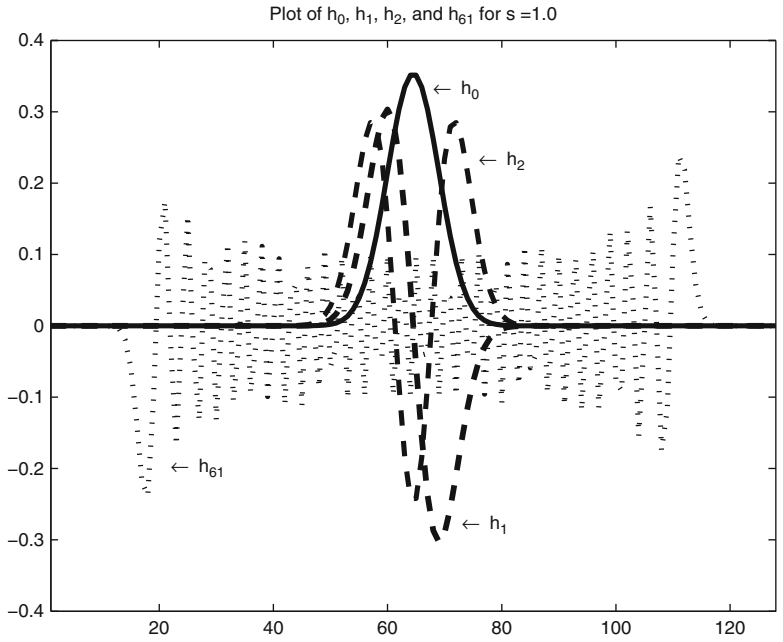

Fig. 11.1 Discrete Hermite functions, undilated

and every L^2 function has an expansion in terms of the CHf. Another important property of the CHf is that they are eigenfunctions of the Fourier transform, in the sense that $\mathscr{F}(h_n(x)) = (-i)^n h_n(x)$, where $\mathscr{F}$ is the integral Fourier transform.

The DHf, as defined in [2, 3], have the analogous property that they are eigenvectors of the centered Fourier matrix. In particular, $\mathscr{F}_{\mathscr{C}}(h_k) = (-i)^k h_k$, where $\mathscr{F}_{\mathscr{C}}$ is the centered Fourier matrix and h_k is the kth eigenvector. As eigenvectors of a related symmetric tridiagonal matrix, as explained in [2], they form an orthonormal set of eigenvectors, and every vector of length N can be expressed as a linear combination of the DHf. This means that for a vector x of length N, there is a representation

$$x[n] = \sum_{k=0}^{N-1} c_k h_k[n] \tag{11.4}$$

$0 \le n \le N-1$, where we use the bracket notation to indicate a discrete function of index n. In this representation, (11.4), the c_k are transform values given by the simple inner products of vector x with DHf vector h_k, i.e., with $c_k = \langle x, h_k \rangle$.

Figure 11.1 shows plots of h_0, h_1, h_2, and h_{61}, DHf vectors for the case when $N = 128$. Plots of these vectors are indistinguishable from the plots of the corresponding CHf for small indices of the subscript and very similar to the plots of the corresponding CHf for higher indices. For example, h_0 has the shape of a

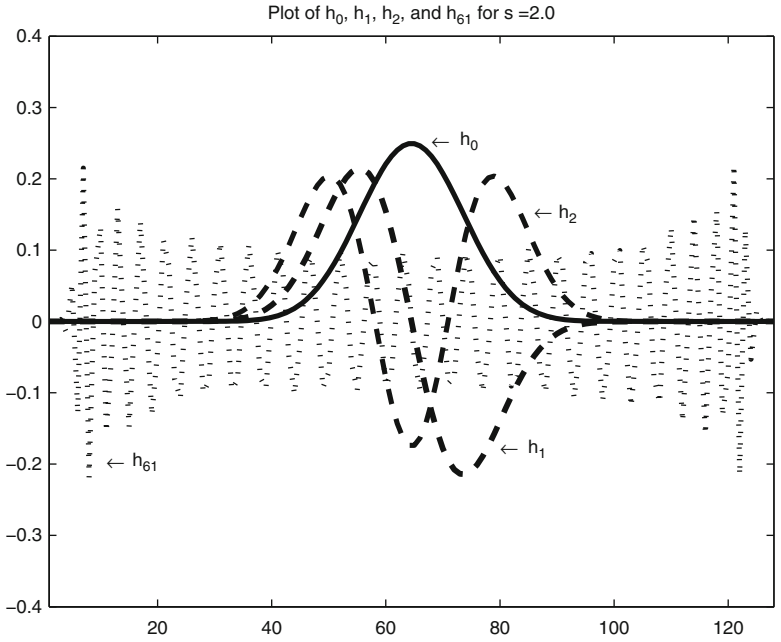

Fig. 11.2 Discrete Hermite functions, dilated with parameter s = 2

discrete Gaussian function. One can show that sampling the CHf at a special set of discrete time values produces vectors extremely close to the DHf for small indices.

An advantage of the DHf for computational purposes is that it is easy to compute these vectors since they are eigenvectors of the very sparse symmetric tridiagonal matrix defined in [2].

It is important for the multiscale applications to follow that the DHf also have a dilated form. This is a result of one of the authors with references as above and in [6, 10]. The dilated form of the DHf is also important for applications in electrocardiogram analysis [5, 12], and in ballistogram artifact removal from electroencephalograms that was done by the authors in [7]. Since the initial DHf vector is basically a discrete Gaussian, one can consider the width of this function as a parameter s that we deem to equal 1 for the undilated case. The dilation comes in interpreting the width of the Gaussian similar to the standard deviation. Producing the dilated set of DHf with parameter s involves adjusting the symmetric tridiagonal matrix defined in [2]. Since the dilated DHF are produced as eigenvectors of a very sparse matrix (as in the undilated case), it is also a fast computational process to produce the entire set of DHf, even for very large length N. As parameter s is increased, the plots of the associated DHf are widened, similar to an increase in the standard deviation of a Gaussian.

Figure 11.2 shows the plots of the same DHf vectors as in Fig. 11.1, except that the vectors are dilated with dilation parameter $s = 2$.

The capability to determine all of the transform values c_k in the representation of digital function x in (11.4) is essential to the subband decomposition that leads to a discrete multiscale Hermite analysis in Sect. 11.3. It is also important that the DHf can be computed quickly and efficiently, in order to provide the multiscale analysis.

11.2 Multiscale for Continuous Hermite Functions

11.2.1 The DoG Function

In the paper [16], "The Multiscale Hermite Transform for Local Orientation Analysis," the authors use the continuous Hermite transform as related to Gaussian derivatives and relate their work to the human visual system. In particular, they involve a scale-space representation as a "multiscale representation that comprises a continuous scale parameter and preserves the same spatial sampling at all scales." See also [4] and [17]. Another idea presented in [18] is to construct a family of signals constructed by convolution with Gaussian kernels with different dilations of the standard deviation.

As noted in [18], the Gaussian convolution of a signal f(x) depends both on x, the signal's independent variable, and on σ, the Gaussian's standard deviation. In particular that convolution is defined by

$$F(x,\sigma) = f(x) * G(x;\sigma) = \int_{-\infty}^{\infty} f(u)G(x-u;\sigma)\mathrm{d}u \tag{11.5}$$

where

$$G(x;\sigma) = \frac{1}{\sqrt{4\pi\sigma}}\mathrm{e}^{-x^2/4\sigma} \tag{11.6}$$

is the Gaussian function with standard deviation σ. Equation (11.5) is the operation of a smoothing filter, with the dilated Gaussian as the low-pass filter.

The starting point for the multiscale decomposition in [16] is the Difference of two Gaussians (DoG) defined by

$$DoG = G(x;\sigma_1) - G(x:\sigma_2) \tag{11.7}$$

where the first term is a narrower Gaussian, corresponding to a smaller standard deviation σ_1, and the second term is related to a wider Gaussian, corresponding to σ_2. This is generalized in [15] with the inclusion of two multipliers, g_1 and g_2,

$$DoG = g_1 \cdot G(x;\sigma_1) - g_2 \cdot G(x:\sigma_2). \tag{11.8}$$

The DoG in (11.8) is related to the receptive fields of retinal ganglion cells [15], where the ratios of the dilation parameters and the multipliers are used as $\sigma_2/\sigma_1 = 3$

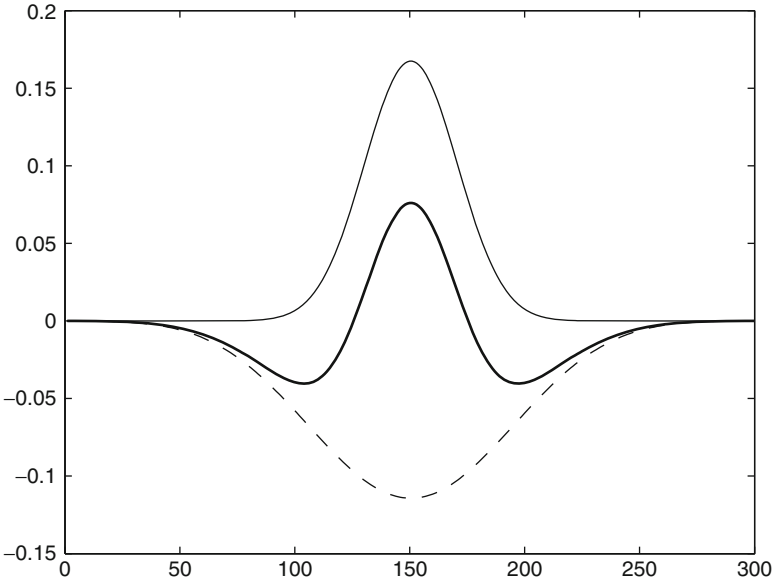

Fig. 11.3 The darker curve is the *DOG* difference of Gaussians

and $g_2/g_1 = 4/5$. A plot of a typical DoG curve using these values of the parameters and the two Gaussians from (11.8) that define it is given in Fig. 11.3.

As noted in Sect. 11.1, the initial Hermite function is a Gaussian, and it follows that the initial DHf h_0 is a discretized Gaussian function. This has been shown to be very close to a sampled version of a Gaussian function, although obtained in a very different way. This means that a discrete DoG can be created using DHf obtained with different dilation parameters. That is, a discrete DoG function can be given by $g_1 \cdot h_{0,s_1} - g_2 \cdot h_{0,s_2}$, with constants g_1 and g_2 as in (11.8) and s as the dilation parameter in the discrete case. This means that the development in [16] that depended on the continuous version of the DoG might be able to be extended to the digital case with this discrete DoG. Though the current article does not involve use of discrete DoG, the authors plan to incorporate it into their research work in the near future.

11.2.2 A Discrete Version of Gaussian Derivatives

Gaussian derivatives are related to the Hermite polynomials. From [8], with scaling parameter σ,

$$D_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \frac{1}{\sigma\sqrt{\pi}} H_n\left(\frac{x}{\sigma}\right) \mathrm{e}^{-x^2/\sigma^2}$$

$$= \frac{1}{\sqrt{2^n n!}} \cdot \frac{\mathrm{d}^n}{\mathrm{d}(\frac{x}{\sigma})^n} \left[\frac{1}{\sigma\sqrt{\pi}} \mathrm{e}^{-x^2/\sigma^2} \right]. \tag{11.9}$$

The Hermite *functions* $h_n(x)$ are defined using the Hermite polynomials, (11.2). This definition can also be related to the Gaussian derivatives in (11.9) to obtain a relation between the Gaussian derivatives and the Hermite functions. Solving for $H_n(x)$ in (11.2) and substituting in (11.9) yield that the $D_n(x)$ are a weighted multiple of $\mathrm{e}^{-x^2/2} \cdot h_n(x)$. For example, with $\sigma = 1$, the relation from (11.9) is

$$D_n(x) = \frac{1}{\sqrt{\pi 2^n n!}} \mathrm{e}^{-x^2/2} h_n(x). \tag{11.10}$$

This leads to a discrete formulation of the Gaussian derivatives. For the DHf from Sect. 11.1, the vector h_0 is a digital version of the Gaussian function $\mathrm{e}^{-x^2/2}$. In conjunction with (11.10), we use the discretization of the Gaussian derivatives $D_n(x)$ given by $\mathrm{d}g_0 = h_0$, and

$$\mathrm{d}g_n = h_0 .* h_n \tag{11.11}$$

for $1 \leq n \leq N-1$, where the vector multiplication (with the $.*$ notation) in (11.11) is term-by-term multiplication. See Fig. 11.4 for a graphical demonstration of the qualities of this discretized set of Gaussian derivatives.

The discrete Hermite transform (DHmT) representation of an arbitrary digital signal x leads to a representation of a windowed version of x using the discrete Gaussian derivatives. Start with the DHmT representation

$$x[m] = \sum_{k=0}^{N-1} c_{k,s} h_{k,s}[m], \tag{11.12}$$

where $c_{k,s} = \langle x, h_{k,s} \rangle$ is given by the usual inner product. With h_{0,s_2} as the discrete Gaussian with appropriate dilation parameter s_2, form the term-by-term product of h_{0,s_2} with x in (11.12) to obtain

$$x .* h_{0,s_2} = \sum_{k=0}^{N-1} c_{k,s} h_{0,s_2} .* h_{k,s}, \tag{11.13}$$

$$x .* h_{0,s_2} = \sum_{k=0}^{N-1} c_{k,s} \mathrm{d}g_k,$$

giving a representation of a Gaussian-windowed version of input vector x in terms of the discrete Gaussian derivatives (11.11). The continuous Gaussian derivatives are used as the filter functions in the general polynomial multiscale analysis in [8]. It may be possible to do something similar for the discrete case with the representation (13).

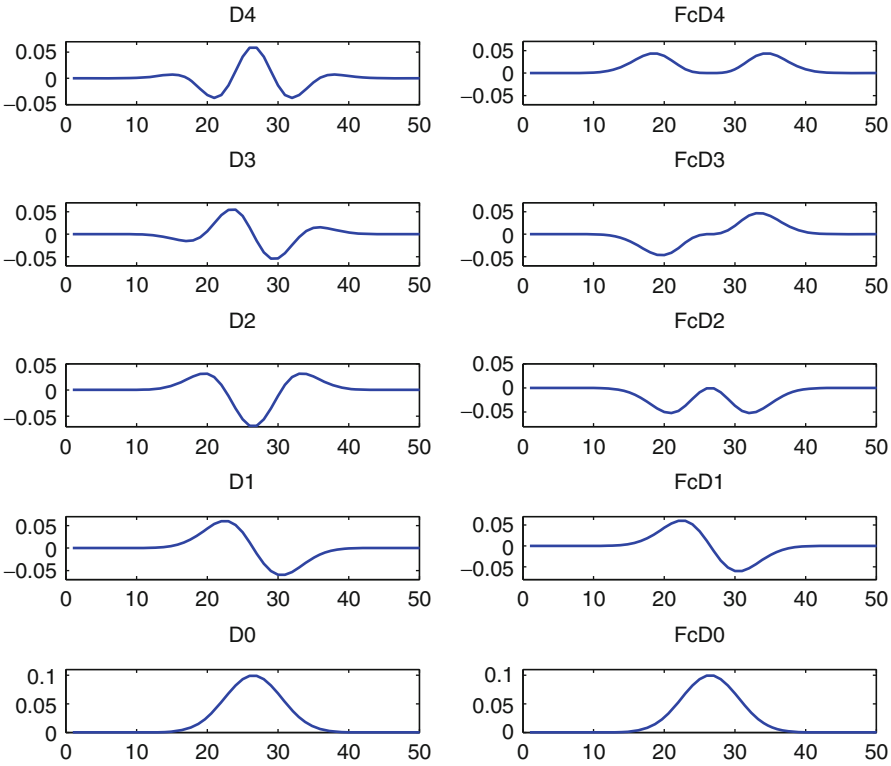

Fig. 11.4 Discrete Gaussian derivatives (*left*) and their Fourier transforms (*right*)

11.2.3 Localized Analysis

In [8], with local window function as the Gaussian

$$V(t) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} \mathrm{e}^{-t^2/2\sigma^2},$$

the author constructs a weighting function

$$W(t) = \sum_k V(t - kT)$$

where T is the period for the periodic repetition of the Gaussian. Provided that W is nonzero for all appropriate values of t, there is a representation of signal $x(t)$ given by

$$x(t) = \frac{1}{W(t)} \sum_k x(t) \cdot V(t - kT). \tag{11.14}$$

In particular, note that for any value of t in the domain of interest, say at $t = t_M$, the value of $\sum_k x(t) \cdot V(t - kT)$ at t_M reduces to the one term in this sum that corresponds to the value of k_M for which the support of $V(t - kT)$ contains t_M. That term would be $x(t_M) \cdot V(t_M - k_M T)$. The corresponding term in the denominator for $W(t_M)$ is that term without the $x(t_M)$ multiple, so the fraction on the right side of (11.14) reduces to $x(t_M)$.

To apply this kind of analysis directly to a discrete, digital signal $x[n]$, let $h_{0,s}$ be the discrete Gaussian obtained as the initial function of the set of discrete, dilated Hermite functions with dilation parameter s. Define a digital weighting function by

$$W[n] = \sum_k h_{0,s}[n - kT],$$

a periodic function of T that covers the domain of values of interest. Note that the dilation parameter s in $h_{0,s}$ must be large enough with respect to shift parameter T so that there is overlap in the shifted discrete Gaussians.

Within each localized region, expand the discrete signal $x \cdot^* h_{0,\sigma_1}$ in a DHmT expansion with scale parameter σ_2 so that

$$x \cdot^* h_{0,s_1} = \sum_{m=0}^{N-1} c_m h_{m,s_2} \tag{11.15}$$

with $c_{m,1} = \langle x \cdot h_{0,s_1}, h_{m,s_2} \rangle$, the inner product of the windowed portion of the input signal x with a s_2-dilated DHf. This makes it so the coefficient in (11.15) is equal to

$$c_m = \sum_{k=0}^{N-1} x[k] \cdot h_{0,s_1}[k] \cdot h_{m,s_2}[k]. \tag{11.16}$$

For multiscale analysis, the windowed input signal, $x \cdot h_{0,s_1}$, and the expansion in (11.15) for a localized region will be developed further so as to analyze the windowed signal at different scales.

11.3 Multiscale Analysis with Discrete Hermite Functions

11.3.1 Background

As noted in Sect. 11.2, the authors in [16] use CHf as related to Gaussian derivatives for the multiscale Hermite transform for continuous signals. A discrete form of Gaussian derivatives was developed in Sect. 11.2.2 using the DHf that extend the concept of Gaussian derivatives to provide an expansion (11.14) of a Gaussian-windowed discrete signal in terms of those discrete Gaussian derivatives. These results involve the relations between Gaussian derivatives and Hermite functions as

given in both the continuous case (11.10) and discrete case (11.11). Section 11.1 summarized properties of the CHf as well as the DHf as developed in [2, 12].

Earlier, Martens [8, 9] established multiscale analysis in the broad context of polynomial transforms, with downsampling and upsampling as important operations in creating the multiscale representations. In [16], the authors provide a pyramidal implementation of analysis and synthesis for the multiscale Hermite transform. The idea of a pyramidal implementation was discussed in [1].

The idea of a multiscale analysis is closely connected to subbands and filter banks, and we build upon that connection in this section. In a standard implementation of wavelets for multiscale analysis, the level one analysis involves a low-pass filter and a high-pass filter that divide the frequency domain into two nearly equal parts. The "detail" subsignal D_1 corresponds to the output of the high-pass filter and is the part of the signal that contains rapidly changing quantities. The "approximation" subsignal A_1 corresponds to the output of the low-pass filter and is the coarse approximation to the signal. Together, if x is the input signal, wavelet analysis results in the reconstruction of x as $x = A_1 + D_1$. Further levels of the multiscale analysis subdivide the coarse subsignal A_1 in exactly the same way as the input signal, so that $A_1 = A_2 + D_2$, where A_2 is an even coarser approximation to the input signal and D_2 is higher level details. Interpreted in the frequency domain, the frequency intervals for A_2 and D_2 fit into the frequency interval of A_1. Combining the two equations above gives that $x = A_2 + D_2 + D_1$.

Our approach for a multiscale analysis using DHf is similar to the standard implementation of wavelets. It involves creating "approximation" and "detail" subsignals at several levels, giving a multiscale representation for a discrete, windowed input signal. At the end of this section, we will show that the sum of specific terms leads to reconstruction of the windowed input signal. As discussed in Sect. 11.2.3 on localized analysis, we assume that the input signal is windowed with a discrete Gaussian window function in order to isolate a portion of the signal, but, for ease of understanding, we will refer to the discrete input signal simply as $x[n]$ under that assumption.

11.3.2 Application of Multiscale Analysis for Discrete Signals

The method we use begins with the representation of the discrete input signal in terms of the DHf as in (11.4). Our method essentially uses subband decompositions for the multiscale analysis, but in the Hermite transform domain as opposed to the usual Fourier transform domain. The DHmT transform values are the coefficients in that representation, listed there as c_k. Those transform values are the values multiplying the corresponding DHf $h_{k,\sigma}$. We assume that the number N of terms in the expansion is at least an even number, for the first downsampling by 2, and in general is a multiple of a power of 2, with that power being equal to the number of levels in the multiscale analysis.

Downsampling in this method will refer to separating the transform values into two distinct halves, similar to subband downsampling for general wavelet

transformations. These begin as the transform values representing the input signal as in (11.4). The approximation at level 1 will be the function related to the first half of the transform values, and the details at level 1 will be the function related to the second half of those values. In particular, if c_1 represents the first half of the DHmT values, i.e.,

$$c_1 = \{c_{0,s}, \ldots, c_{(N-1)/2,s}\},$$

with s as the dilation parameter, then the approximation at level 1 A_1 is the inverse DHmT of the vector formed with c_1 as the values in the first half and with zeros in the second half. Similarly, the details at level 1, D_1, are the inverse DHmT of the vector c_2 with first half as zeros and with second half as the DHmT values

$$c_2 = \{c_{N/2,s}, \ldots, c_{N-1,s}\}.$$

The transform coefficients are effectively downsampled by 2 in the sense that they are cut into first and second halves. Since $x = \sum_{k=0}^{N-1} c_{k,s} h_{k,s}$, the representation $x = A_1 + D_1$ results from the linearity of the DHmT transform.

The next level in the multiscale analysis involves the splitting of the "low-pass" DHmT values c_1 into halves. Since the DHf are basically their own Fourier transforms, the lower-indexed h_k functions that have their primary support near the origin also have a spectrum that is nearer the origin. Using the lower-indexed values for the approximation then connects this multiscale analysis using the DHf to wavelet multiscale analysis.

Let c_{11} be the first half of c_1 and c_{12} be the second half of c_1. The approximation at level 2 will be the inverse DHmT of the length N vector whose first fourth is c_{11} but is otherwise zero. The details at level two are the inverse DHmT of the length N vector whose second fourth is c_{12} but is otherwise zero. For the third-level approximation and detail vectors, continue this process, halving the vector c_{11} once again. This process can be continued to as many levels as the length of the input vector will allow.

11.3.3 Results

This section provides an example of multiscale analysis up to three levels for the multiscale analysis method as outlined in Sect. 11.3.2 and as applied to a physiological signal. The example signal sample contains noise and is not perfectly periodic, but is an actual physiological signal with basic periodicity. A plot of the input signal is given in Fig. 11.5.

The DHmT of a signal includes a dilation parameter σ. It is important that the dilation parameter be chosen large enough, at the beginning of the computation, so that the DHf h_k cover the entire interval of the windowed function. The DHf beginning with the discrete Gaussian h_0 have effective support that expands as k increases, and the multiscale analysis will use relatively small index values as the number of levels in the multiscale analysis increases.

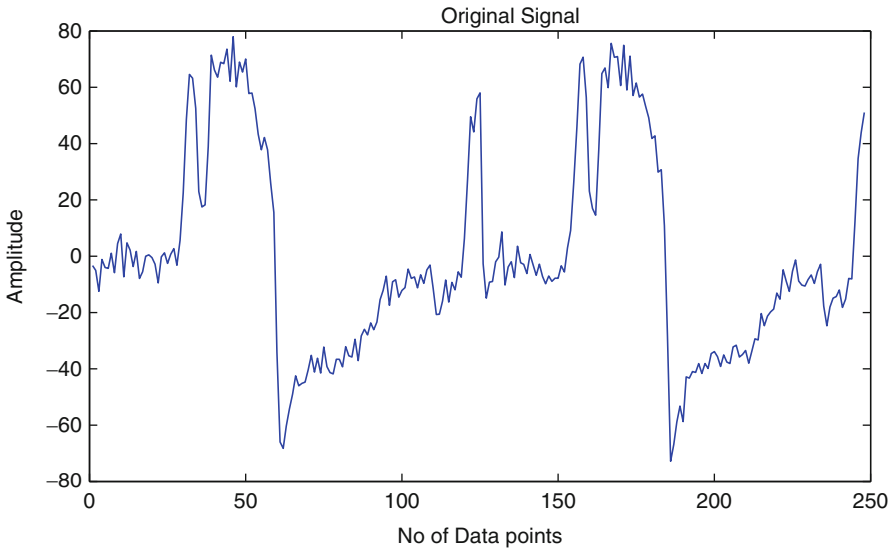

Fig. 11.5 Signal used for analysis

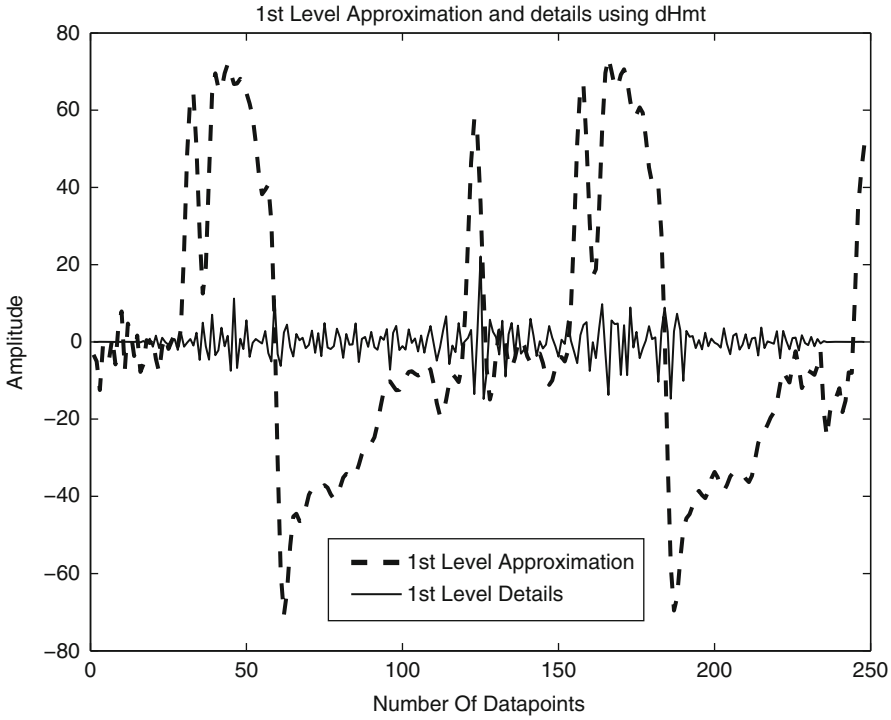

Fig. 11.6 First-level DHmT approximation and detail

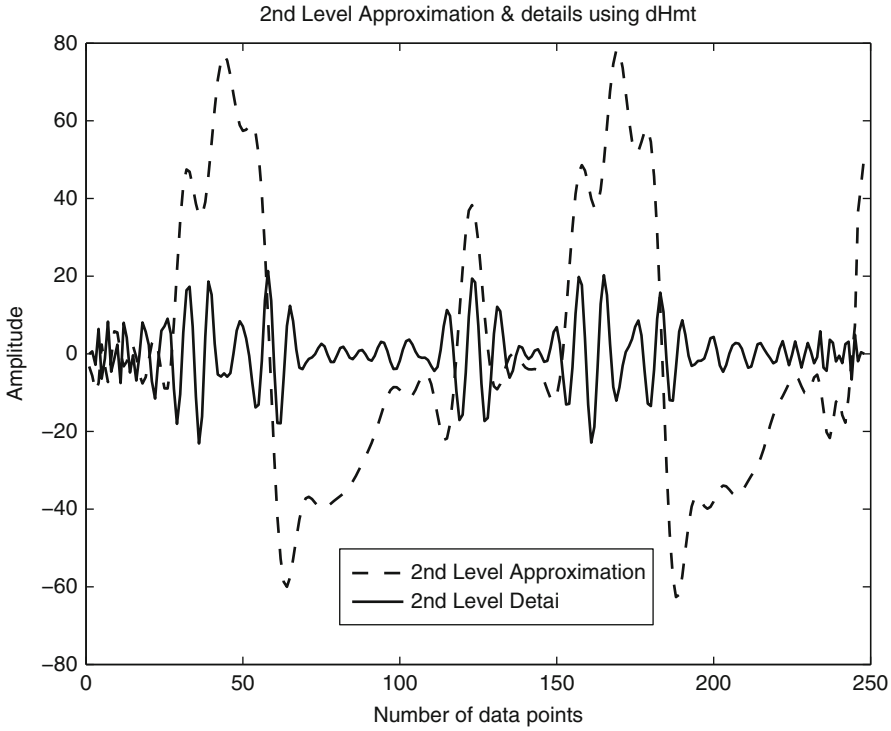

Fig. 11.7 Second-level DHmT approximation and detail

Using the method outlined in Sect. 11.3.2, the first-level approximation and details are plotted in a common graph in Fig. 11.6. At this level, the details are small in magnitude and contain mostly high-frequency noise, and the underlying pattern of periodicity is seen in the first-level approximation.

Completing the multiscale analysis with the DHmT to the second level, the approximation and details at level two are shown in Fig. 11.7. The larger scale is clear from the plot as the approximation becomes smoother.

At the third level of the multiscale analysis of the signal pictured in Fig. 11.5, a very clear underlying periodic pattern is seen in the approximation at level three in Fig. 11.8. The periodic pattern is even more clear at level three than at the previous levels of the preceding figures. The details at level three have magnitudes that are the largest of these particular vectors. Note the smoothness of the approximation of the input signal at this level.

The subband method, here employed for the DHmT instead of the Fourier transform, provides the usual reconstruction of the input signal from the multiscale signals illustrated in Figs. 11.5–11.8. If the scaling parameter s is kept constant over the different levels, then the reason that

$$x = A_3 + D_3 + D_2 + D_1,$$

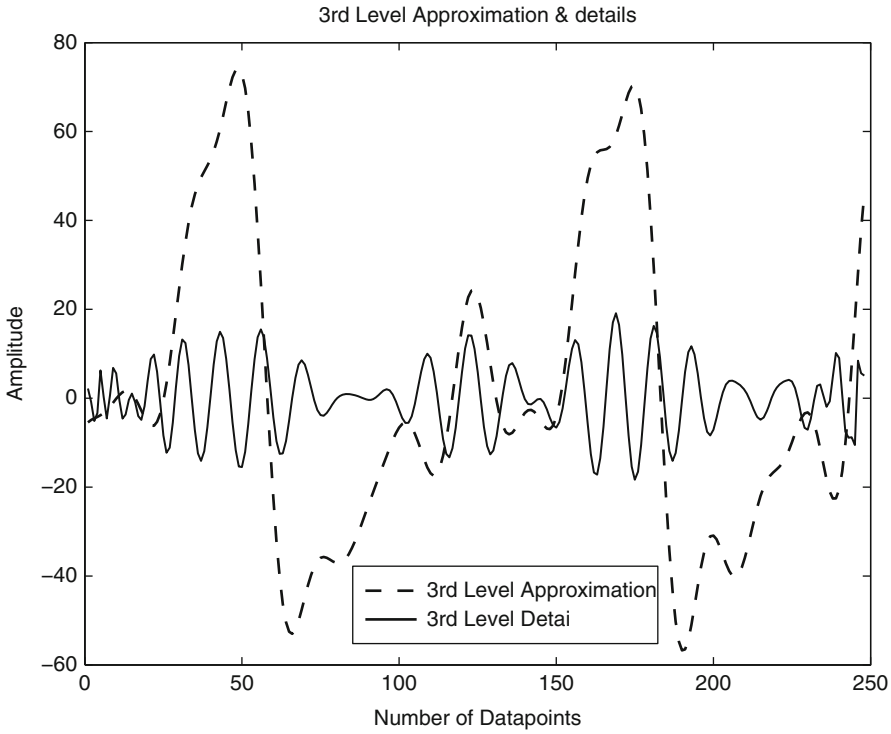

Fig. 11.8 Third-level DHmT approximation and detail

follows from the linearity of the DHmT transform.

The computational cost of this multiscale analysis is only in the determination of the forward transform and several inverse DHmT.

11.3.4 Comparison to Wavelet Analysis

The DHmT analysis of an input signal as illustrated in Figs. 11.6–11.8 is compared in this section to wavelet analysis of the same signal for the three different levels. The wavelet decompositions presented here use the biorthogonal 3.3 wavelet.

The first figure shows the wavelet decomposition, both the approximation and details, at level one (Fig. 11.9). This can be compared to Fig. 11.6.

The next figure, Fig. 11.10, shows the wavelet decomposition at level two. This plot should be compared to Fig. 11.7.

Figure 11.11 shows the wavelet decomposition at level three. This plot should be compared to Fig. 11.8.

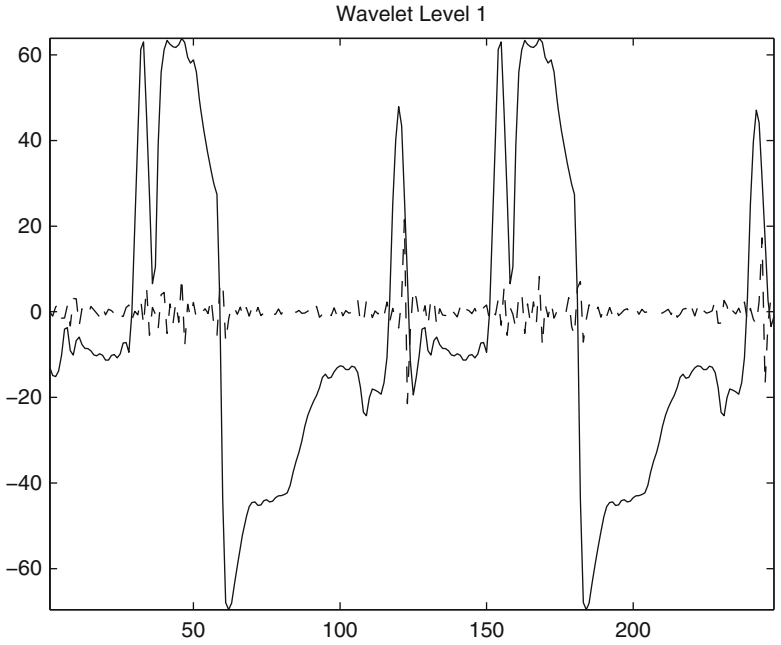

Fig. 11.9 First-level wavelet approximation and detail

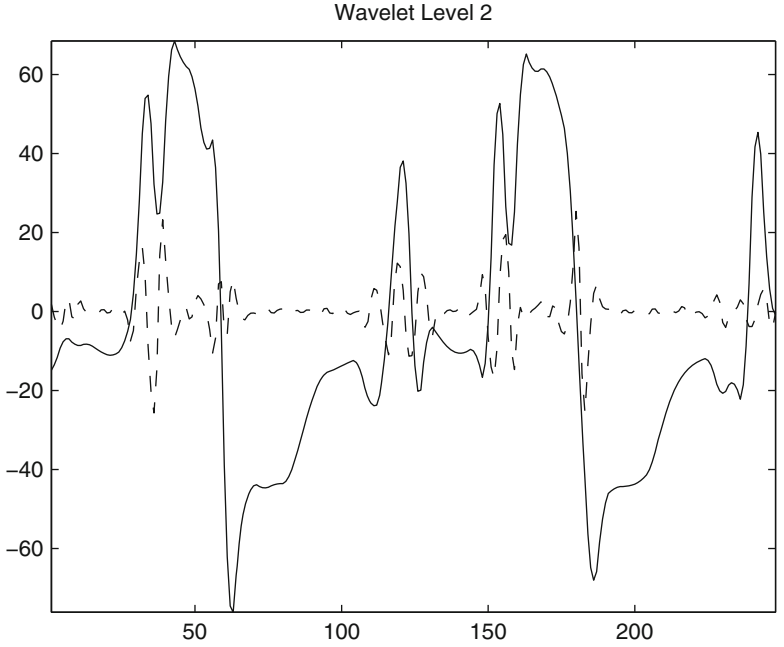

Fig. 11.10 Second-level wavelet approximation and detail

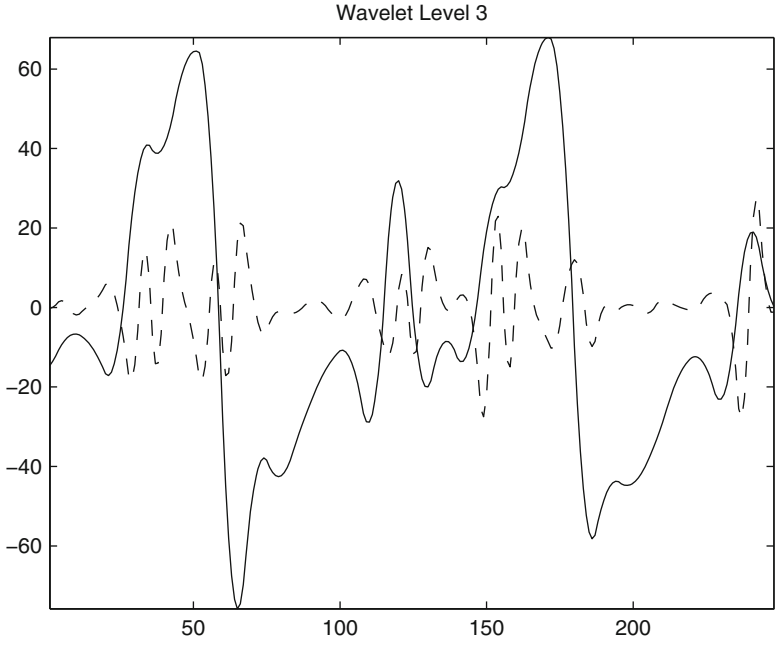

Fig. 11.11 Third-level wavelet approximation and detail

In general, the DHmT decomposition approximations are somewhat smoother for this particular input signal than those in the corresponding wavelet decompositions. There are subtle but distinct differences between the two decompositions for this input signal. That the decompositions are somewhat similar indicates that the DHmT may provide an alternative multiscale analysis to that provided by wavelets.

11.4 Conclusions

The previous section described a method for using the DHmT to produce a multiscale analysis of a digital signal. Figures 11.5–11.8 illustrate a multiscale analysis for an example of a nearly periodic physiological signal. This multiscale analysis is made possible by the new method of producing a complete set of DHf [3] in a way that preserves the property of the CHf that they are essentially their own Fourier transform. The new method also provides a computationally efficient method to produce the entire set of DHf. Previous methods to produce versions of the CHf for the discrete case are cumbersome and are generally able to only produce a small number of the DHf for actual use. The method of the previous section is based on a subband decomposition of the DHmT values.

The basic properties of the CHf and analogous properties of the new set of DHf were described at the beginning of this chapter. It is important to this multiscale analysis that the DHf form an orthonormal set of vectors. The DoG function was discussed in Sect. 11.2, beginning with the relations used in the article [16] that used this development for multiscale analysis involving the CHf. That article also involved the Gaussian derivatives for the continuous case, and these were shown to extend to the discrete case in Sect. 11.2.2 with a set of discrete Gaussian derivatives based on the DHf. A multiscale analysis involving these discrete Gaussian derivative functions is an idea still awaiting further development, as a different approach was used for the multiscale analysis in Sect. 11.3. The idea that multiscale analysis is related to windowing was developed further in Sect. 11.2.3 with that idea being used throughout this chapter.

This chapter touches on several different ways to develop multiscale analysis using the DHf and develops and demonstrates the working of one such method to conclusion. The importance of multiscale analysis suggests that the other ways should also be developed further as well. That work is still to come.

References

1. Burt PJ, Adelson EH (1983) The Laplacian pyramid as a compact image code. IEEE Trans Commun 31:532–540
2. Clary S, Mugler DH (2003) Shifted Fourier matrices and their tridiagonal commutors. SIAM J Matrix Anal 24(3):809–821
3. Clary S, Mugler DH (2004) Eigenvectors for a class of discrete cosine and sine transforms. Sampling Theor Signal Image Process 3:83–94
4. Escalante-Ramrez B, Martens JB (1992) Noise reduction in computerized tomography images by means of polynomial transforms. J Visual Commun Image Rep 3:272–285
5. Gopalikrishnan R, Acharya S, Mugler DH (2004) Real time monitoring of ischemic changes in electrocardiograms using discrete Hermite functions. In: Proceedings of 26th International IEEE EMBS Conference, San Francisco, IEEE, pp 438–441
6. Gopalikrishnan R, Mugler DH (2010) The evolution of Hermite transform in biomedical applications. In: Shukla A, Tiwari R (eds.) Intelligent medical technologies and biomedical engineering: tools and applications. IGI Global, San Francisco, IEEE, pp 260–278
7. Mahadevan A, Acharya S, Sheffer D, Mugler DH (2008) Ballistocardiogram artifact removal in EEG-fMRI signals using discrete Hermite transforms. IEEE J Selected Topics Signal Process (Special Issue on: fMRI analysis for human brain mapping) 2(6):839–853
8. Martens JB (1990) The Hermite transform-theory. IEEE Trans Acoust Speech Signal Process 38(9):1595–1606
9. Martens JB (1990) The Hermite transform-applications. IEEE Trans Acoustics, Speech Signal Process 38(9):1607–1618
10. Mugler DH, Clary S (2000) Discrete Hermite functions. In: Proceedings of the International conference on scientific computing and mathematical modeling, IMACS 2000, pp 318–321
11. Mugler DH, Clary S (2001) Discrete Hermite functions and the fractional Fourier transform. In: Proceedings of the International Workshop on Sampling Theory, pp 303–308
12. Mugler DH, Clary S, Wu Y (2002) Discrete Hermite expansion of digital signals: applications to ecg signals. In: Proceedings of IEEE Signal Processing Society. 10th DSP Workshop. Georgia, pp 271–276

13. Poularikas AD (1999) The handbook of formulas and tables for signal processing. CRC Press
14. Refregier A (2003) Shapelets I. A method for image analysis. Mon Not R Astron Soc 338: 35–47
15. Rodieck RW (1965) Quantitative analysis of cat retinal ganglion cell response to visual stimuli. Vision Res 5:583–601
16. Silván-Cárdenas JL, Escalante-Ramrez B (2006) The multiscale hermite transform for local orientation analysis. IEEE Trans Im Proc 15:1236–1253
17. Van Rullen R, Thorpe SJ (2001) Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. Neural Comput 13:1255–1283
18. Witkin A (1984) Scale-space filtering: A new approach to multiscale description. In: Proceedings of IEEE International Conference Acoustics, Speech, Signal Processing (ICASSP), vol 9. San Diego, CA, pp 150–153

Chapter 12
Earth Mover's Distance-Based Local Discriminant Basis

Bradley Marchand and Naoki Saito

Abstract Local discriminant basis (LDB) is a tool to extract useful features for signal and image classification problems. Original LDB methods rely on the time–frequency energy distribution of classes or empirical probability densities, with some information theoretic measure (such as Kullback–Leibler divergence) for feature selection. Depending on the problem, energy distributions may not provide the best information for classification. Further, training set sizes and accuracy in the computed empirical probability density functions (epdfs) may hinder the learning process. To improve these deficiencies and provide a more data adaptive algorithm, we propose the use of signatures and earth mover's distance (EMD). Signatures and EMD provide a data adaptive statistic that is more descriptive than the distribution of energies and more robust than an epdf-based approach. In this chapter, we first review LDB and EMD and then outline how they can be incorporated into a fast EMD based LDB algorithm.We then demonstrate the capabilities of our new algorithm in comparison to both energy distribution and epdf-based LDB algorithms using four different classification problems using synthetic datasets.

12.1 Introduction

Local discriminant basis (LDB) is a best basis algorithm developed by Saito and Coifman for the purpose of classification [9, 10]. It works by decomposing training signals into a time–frequency dictionary, such as block discrete cosine transform, local cosine transform, or wavelet packet transform (WPT). The dictionaries

B. Marchand (✉)
Panama City Division, Naval Surface Warfare Center, FL 32407, USA
e-mail: bradley.marchand@navy.mil

N. Saito
Department of Mathematics, University of California, Davis, CA 95616, USA
e-mail: saito@math.ucdavis.edu

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_12,

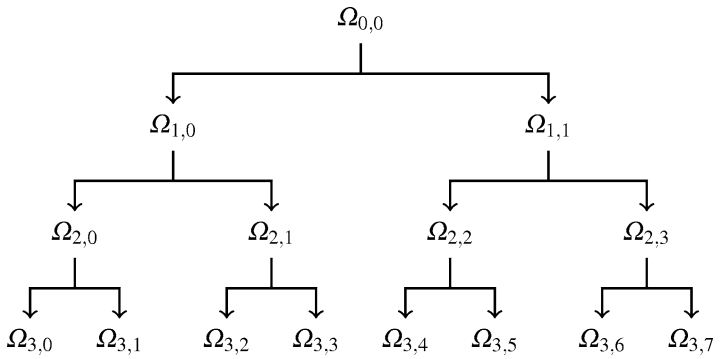

Fig. 12.1 Depiction of the wavelet packet transform. The process is the same as the wavelet decomposition with the added decomposition of the high-pass coefficients. This yields a redundant decomposition

decompose signals into a redundant set of orthogonal subspaces, as shown in Fig. 12.1. Each subspace contains basis vectors localized in time and frequency. Its goal, given a dictionary, is to find the signal representation within the dictionary that is most useful for classification and discrimination. The idea is that these dictionaries provide us with localized elementary building blocks for isolating critical differences among signal classes. These differences are learned from LDB's time–frequency map. In the original LDB algorithm, this map was a simple accumulation of class signal energy at each coordinate in each subspace. Formally, let N_c be the number of signals belonging to class c, $\left\{\mathbf{x}_i^{(c)}\right\}_{i=1}^{N_c}$ be the set of signals belonging to class c, and $\mathbf{w}_{j,k,l}$ be basis vectors from our selected dictionary parameterized by indices j, k, and l indicating the scale (or level of decomposition), frequency band, and position of basis vector, respectively. Then, our energy map for class c is formed as,

$$\Gamma_{j,k,l}^{(c)} := \frac{\sum_{i=1}^{N_c}\left(\mathbf{w}_{j,k,l}\cdot\mathbf{x}_i^{(c)}\right)^2}{\sum_{i=1}^{N_c}\left\|\mathbf{x}_i^{(c)}\right\|^2}.$$

Later Saito et al. [11] proposed a refinement of the algorithm by changing the time–frequency map from an accumulation of class signal energy to statistical distributions of the expansion coefficients. Although any distribution metric can be used to evaluate the discriminating power of a coordinate, average shifted histograms (ASH) [12] were used to compute an empirical probability density function (epdf) in [11] for their computational efficiency. This improvement allows LDB to detect finer differences because the statistical behavior of the class signals in each coordinate in the dictionary can be analyzed.

In this chapter, to further refine LDB's time–frequency map, we propose the use of *signatures* instead of epdfs and the use of *earth mover's distance* (EMD) [7] to compute the discriminating power of a coordinate. Signatures provide us with a fully data-driven representation, which can be efficiently used with EMD. This representation is more efficient than a histogram and is able to represent complex data structure with fewer samples. EMD is a metric between signatures that naturally extends the notion of distance between points to that of sets or distributions of elements. We begin by reviewing the concept of signatures and EMD in the next section. Then we will outline an EMD-based LDB algorithm in Sect. 12.3. Next, in Sect. 12.4, we will compare performance of all three LDB algorithms on synthetic datasets using various base classifiers. Finally, we conclude in Sect. 12.5 with a summary of performance.

12.2 Earth Mover's Distance

A *signature* represents a set of clusters of feature vectors, say, in $\mathbb{R}^d$. Each such cluster is represented by its mean, $\mathbf{m}_j$, of vectors belonging to that cluster and the weight (or importance) of that cluster, $w_{\mathbf{m}_j}$. The number of clusters in a signature varies with the complexity of the object being represented. Signatures are generalized histograms. A histogram is a fixed partitioning of the underlying space with cluster centers defined as the central value in each bin. The weight of each cluster is the percentage of points in the bin. The flexibility provided by signatures is the ability to place the "bins" where the data is located. For example, representation of data that exists on a curved manifold might require a relatively fine partitioning of the space to achieve a histogram that captures the distribution of the data. However, a signature representation is likely to be much more efficient since we are not required to partition the entire space, and feature clusters can be placed at ideal locations along the manifold. A comparison of histograms and signatures is detailed in [8]. Unfortunately, most dissimilarity measures cannot be applied to signatures. This is because they rely on direct correspondence between bins. That is, they can be used for histograms that contain the same number of bins. This is, however, not guaranteed with signatures. EMD, on the other hand, is designed for use with signatures.

EMD was first introduced by Rubner, et al. [8] for retrieval of color and textured images. It has several properties that have many advantages over other distance measures:

- Applies to signatures
- Naturally reflects nearness
- Allows for partial matching
- Is a metric (if total weights of two signatures are equal and cost is a metric)

EMD has the intuitive interpretation of the minimum amount of work required to move piles of soil (or earth) into holes. The location and size of the piles of soil are

represented by cluster centers and weights of a signature, respectively. Similarly, the other signature represents the location and size of the holes to be filled. Formally, if we let

$$P = \{(\mathbf{p}_1, w_{\mathbf{p}_1}), \ldots, (\mathbf{p}_m, w_{\mathbf{p}_m})\},$$
$$Q = \{(\mathbf{q}_1, w_{\mathbf{q}_1}), \ldots, (\mathbf{q}_n, w_{\mathbf{q}_n})\}$$

be our two signatures and $C = [c_{ij}]$ the cost matrix where c_{ij} represents the cost of moving one unit of mass from the ith cluster in P to the jth cluster in Q, then the EMD algorithm seeks the flow $F = [f_{ij}]$ that minimizes the work

$$W(P,Q,F) := \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} f_{ij},$$

subject to the constraints:

$$f_{ij} \geq 0 \qquad 1 \leq i \leq m, 1 \leq j \leq n;$$
$$\sum_{j=1}^{n} f_{ij} \leq w_{\mathbf{p}_i} \qquad 1 \leq i \leq m;$$
$$\sum_{i=1}^{m} f_{ij} \leq w_{\mathbf{q}_j} \qquad 1 \leq j \leq n;$$
$$\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} = \min \left(\sum_{i=1}^{m} w_{\mathbf{p}_i}, \sum_{j=1}^{n} w_{\mathbf{q}_j} \right).$$

Once the optimal flow F is found, EMD is the resulting work normalized by the total flow:

$$\mathrm{EMD}(P,Q) := \frac{W(P,Q,F)}{\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij}}.$$

This normalization is necessary to avoid favoring smaller signatures if the two signatures have different total weights.

The optimal flow is found by solving the well-known transportation problem, or *the Monge-Kantorovich mass transportation problem* [5]. Typically, this requires the use of a linear programming such as the simplex method to solve for the optimal flow. A detailed explanation of the simplex method can be found in [5]. However, there are a few situations where fast algorithms that do not require linear programming can be used.

In particular, for the one dimensional case where the cost is the Euclidean distance and the signatures have equal total weights, w_Σ, Rubner and Tomasi [7, Sect. 2.3.1] showed that the EMD can be directly calculated by

$$\mathrm{EMD}(P,Q) = \frac{\sum_{k=1}^{m+n-1} |\hat{p}_k - \hat{q}_k|(r_{k+1} - r_k)}{w_\Sigma},$$

where $r_1, r_2, \ldots, r_{m+n}$ is the sorted list of

$$p_1, p_2, \ldots, p_m, q_1, q_2, \ldots, q_n,$$

and

$$\hat{p}_k := \sum_{p_i \le r_k} w_{p_i}, \qquad \hat{q}_k := \sum_{q_i \le r_k} w_{q_i}.$$

The algorithm relies on the fact that the minimum work between two one dimensionals distributions is known to be the L_1 distance between the cumulative distribution functions (cdfs), as discussed by [7, Sect. 2.3.1]. We note that the L_1 distance between two cdfs is a special instance of the so-called $\bar{\rho}$ or Ornstein distance [1]; see also [6] on the deep relationship between EMD and the Marrows distance often used in statistics. The following theorem is presented for convenience:

Theorem 1. *Define the empirical cdfs of 1D signatures P and Q as*

$$P(t) := \begin{cases} 0 & t \in (-\infty, p_{(1)}), \\ \sum_{i=1}^{k} w_{p_{(i)}} & t \in [p_{(k)}, p_{(k+1)}),\ 1 \le k \le m-1, \\ \sum_{i=1}^{m} w_{p_{(i)}} & t \in [p_{(m)}, \infty), \end{cases}$$

$$Q(t) := \begin{cases} 0 & t \in (-\infty, q_{(1)}), \\ \sum_{j=1}^{k} w_{q_{(j)}} & t \in [q_{(k)}, q_{(k+1)}),\ 1 \le k \le n-1, \\ \sum_{j=1}^{n} w_{q_{(j)}} & t \in [q_{(n)}, \infty), \end{cases}$$

where $\{p_{(i)}\}$ and $\{q_{(j)}\}$ are sorted versions (in nondecreasing order) of $\{p_i\}$ and $\{q_j\}$, respectively. If P and Q have equal total weights $\sum_{i=1}^{m} w_{p_i} = \sum_{j=1}^{n} w_{q_j} =: w_\Sigma$, then

$$\mathrm{EMD}(P,Q) = \frac{\int_{-\infty}^{\infty} |P(t) - Q(t)|\,\mathrm{d}t}{w_\Sigma}.$$

Throughout this chapter we will be using this fast 1D version of EMD, which restricts our cost function to L_1 distance. We could use a simplex solver, which would allow other cost functions, but the use of the simplex solver greatly impacts speed. The worst-case computational cost for a simplex solver is exponential, but the use of a transportation-simplex solver and a good initial solution (close to an optimal solution) greatly improves performance. Rubner and Tomasi report [7, Sect. 2.3] an empirical performance of $O(n^3 \log n)$ in the case when both P and Q have $n = m$ clusters. This is opposed to the fast 1D EMD solver which has a

computational cost of $O(n \log n)$. Clustering helps lighten the computational burden of the transportation-simplex solver, but introduces complexity to the algorithm (in the form of clustering parameters) and presents variability of results. With 1D EMD we can completely avoid clustering the data.

12.3 EMD-Based Local Discriminant Basis

Generally, the LDB algorithm is broken into the following steps:

Algorithm 2. *Given a training dataset $\mathscr{T}$ that consists of C classes of signals $\{\{\mathbf{x}_i^{(c)}\}_{i=1}^{N_c}\}_{c=1}^{C}$:*

- *Choose a dictionary and specify the maximum level of decomposition.*
- *Expand all the signals into tree-structured subspaces.*
- *For each class, compute a time–frequency map.*
- *Use the time–frequency maps to compute the discriminating power of each coordinate.*
- *Prune the tree by examining the discriminating power of each subspace.*
- *Order the basis vectors by their discriminating power.*

This process is made efficient by the speed with which the signals can be expanded into the selected dictionary, and by exploiting the tree structure in the pruning process and using additive discriminant measures for comparison.

Definition 1. Let $\mathbf{p}$ and $\mathbf{q}$ be any two vectors in $\mathbb{R}^n$. A *discriminant measure*, $D(\mathbf{p},\mathbf{q})$, is a map $D : (\mathbb{R}^n \times \mathbb{R}^n) \to \mathbb{R}$ such that

- $D(\mathbf{p},\mathbf{p}) = 0$.
- (Nonnegative) $D(\mathbf{p},\mathbf{q}) \geq 0$ for all $\mathbf{p}$, $\mathbf{q}$ in $\mathbb{R}^n$.

A discriminant measure is said to be *additive* if

$$D(\mathbf{p},\mathbf{q}) = \sum_{i=1}^{n} D(p_i, q_i).$$

For our EMD-based LDB, we construct our time–frequency map by collecting signatures for each coordinate of each class in a subspace. Specifically, for coordinate l of class c in subspace $\Omega_{j,k}$, our signatures are

$$s_{(j,k,l)}^{(c)} = \left\{ \left(\mathbf{w}_{j,k,l} \cdot \mathbf{x}_i^{(c)}, 1/N_c \right) \right\}_{i=1}^{N_c}. \tag{12.1}$$

The collection of these signatures form our time–frequency map. In this formulation we have chosen our samples for each class to be equally weighted, which is reasonable since we are not assuming or attempting to compute the relative importance of a particular signal to its class or its overall ability to discriminate. To efficiently evaluate the discriminant power of a coordinate or subspace, we define our additive measure using EMD as follows.

Definition 2. Let $s^{(c)}_{(j,k,l)}$ be the signature for training signals belonging to class c at level j, block k, and position l. Then the *EMD distance between classes* is defined as the sum of the pairwise EMD distances:

$$\mathscr{D}(j,k,l) := \sum_{m=1}^{C-1} \sum_{n=m+1}^{C} \text{EMD}\left(s^{(m)}_{(j,k,l)}, s^{(n)}_{(j,k,l)}\right). \tag{12.2}$$

We refer to $\mathscr{D}(j,k,l)$ as the discriminant power of the coordinate. The *discriminant power of a subspace* is then

$$\Gamma(j,k) := \sum_{l=0}^{2^{n_0-j}-1} \mathscr{D}(j,k,l), \quad 0 \le j \le J \le n_0; 0 \le k \le 2^j - 1, \tag{12.3}$$

where the length of each signal in the dataset is assumed to be $n = 2^{n_0}$, and J is the maximum depth of decomposition set by the user.

We also use the notation $\mathscr{D}(\Omega_{j,k}) := \Gamma(j,k)$ to emphasize that $\Gamma(j,k)$ is a discriminant measure for the subspace $\Omega_{j,k}$. Often the sum in (12.3) is truncated; more precisely, we only sum the k_0 largest values from each subspace rather than summing all the 2^{n_0-j} values as (12.3). (Note that for a certain j for k_0 set by the user, we could have $k_0 > 2^{n_0-j}$. In that case, we sum all the 2^{n_0-j} values.) This helps with situations where there are many weak coordinates summing to a large value.

Note that $\mathscr{D}(\Omega_{j,k})$ is an additive discriminant for subspace $\Omega_{j,k}$ since it has been defined as the sum of the discriminant powers of the subspace coordinates. The benefit of an additive discriminant measure comes in the pruning process. Using the tree-structure notation shown in Fig. 12.1, pruning starts at the base of the tree and is conducted by the following rule:

Algorithm 3. *Let* $0 \le j \le J \le n_0$ *and* $0 \le k \le 2^j - 1$. *If* $\mathscr{D}(\Omega_{j,k}) \ge \mathscr{D}(\Omega_{j+1,2k} \cup \Omega_{j+1,2k+1})$, *then select* $\Omega_{j,k}$ *over* $\Omega_{j+1,2k} \cup \Omega_{j+1,2k+1}$; *otherwise, select* $\Omega_{j+1,2k} \cup \Omega_{j+1,2k+1}$.

If the measure $\mathscr{D}(\Omega_{j+1,2k} \cup \Omega_{j+1,2k+1})$ is additive, it can be efficiently computed as

$$\mathscr{D}(\Omega_{j+1,2k}) + \mathscr{D}(\Omega_{j+1,2k+1}),$$

which only requires the addition of the discriminant measure for the individual subspaces to obtain the measure of their union.

Substituting our signature time–frequency map, (12.1), and EMD discriminant measure, (12.3), into Algorithm 2, we obtain our EMD-based LDB algorithm. This algorithm benefits from the adaptive structure of signatures and the robustness of EMD while remaining computationally fast and capable of detecting fine differences with few parameters. Further, our signatures can be quickly updated to incorporate new information. This becomes important when we consider situations where training data is limited and/or the data is noisy. Here incremental learning is important so that new information can be incorporated without complete retraining

and better training samples can be used as they are available. We can avoid the need to retrain on the entire set by initializing our algorithm with a small training set and storing the signature map constructed above with the modification that we store the sum of the sample values and the number of samples in each cluster as our features and weights. This allows us to update the mean features and weights without storing any extra information. The signature's mean value and weight are calculated during the computation of the EMD. Any new data that we wish to incorporate can easily be added by decomposing it into the selected dictionary and incorporating it into our stored signature map. To accommodate limited memory resources and reduce time to extend signature storage, a specified limit can be set for the signature length during initialization, so that memory can be preallocated. If a signature has reached the storage capacity specified or is within a threshold distance to a cluster in the signature, then we can use our grouping technique which is similar to clustering.

Algorithm 4. *Let m^* be our capacity and $m < m^*$. Given signature*

$$S = \{(s_1, w_{s_1}), \ldots, (s_m, w_{s_m})\},$$

where s_i is the sum of each cluster value and w_{s_i} is the number of samples in each cluster for $i \in \{1, \ldots, m\}$. Let $0 \leq \tau$ be our grouping threshold and s_{new} a new sample:

- *If there exists $s_k, k \in \{1, \ldots, m\}$ such that $|s_{\text{new}} - s_k| < \tau$, then set $s_k = s_k + s_{\text{new}}$ and $w_{s_k} = w_{s_k} + 1$.*
- *Else add new sample to the end of the signature:*

$$S = \{(s_i, w_{s_1}), \ldots, (s_m, w_{s_m}), (s_{\text{new}}, 1)\}.$$

Although we are only working in 1D, Algorithm 4 works in any dimension as long as our signatures are not at capacity, $m = m^*$. When a signature reaches capacity we reduce the signature size by combining the closest clusters of the signature. Having our signature evolve in this manner means that our capacity acts as a resolution parameter and our grouping threshold as a sensitivity parameter. In a situation with a complex structure over a relatively large span, our capacity limit may force a less than ideal signature resolution reducing the signature's ability to describe the structure. Therefore, considerations for the training data must be taken when choosing signature capacity.

12.4 Local Discriminant Basis Algorithm Performance

The purpose of constructing a feature space is to pull out the important properties of the datasets for the purposes of discrimination or compression. A good feature space should provide dimension reduction and/or improved classifier performance. LDB algorithms naturally provide dimension reduction by expanding data into an

orthonormal basis followed by selecting the most discriminant basis coordinates. In practice, relatively few LDB vectors are usually needed for discrimination. To evaluate our algorithm on increasing classifier performance, we use two different base classifiers, linear discriminant analysis (LDA) and classification tree (CT); see, e.g., [4, Sects. 4.3 and 9.2] for the details of these classifiers. These two classifiers construct distinct decision boundaries, which gives us an indication of the complexity of the separation. LDA seeks an optimally separating hyperplane as its decision boundary. A classification tree seeks an optimal partitioning of each coordinate.

For comparison, we have three LDB algorithms available: time–frequency LDB algorithm (LDBK), epdf-based LDB algorithm (LDBKASH), and EMD-based LDB algorithm (LDBKEMD). Each algorithm is analyzing a different quantity to determine class separability. For LDBK we are concerned with normalized coefficient energy. For LDBKASH we measure separation of coefficients' epdfs. And for LDBKEMD we are computing the separation of the empirical cumulative distribution functions (ecdfs). All three are analyzing the same information, just in different ways. As we will see in the examples below, how you analyze the information greatly impacts the quality of the selected feature space. Our LDBKASH and LDBKEMD algorithms are using distribution information and not a single statistic for each coordinate. So, they are able to incorporate information about the statistical distribution of the coefficients into the selection process. This allows for the detection of subtler differences that can be lost in the case of a single statistic. However, estimating reliable epdfs using a noisy dataset with a limited number of samples is quite difficult.

To compare the performance of the different LDB algorithms, we conduct classification experiments on four different synthetic signal datasets. The first two datasets (triangular waveforms and shape waveforms) were used by Saito [9] to demonstrate the benefit of using the original LDB algorithms for classification. Our third dataset is a variation of the shape waveforms dataset constructed for subtle differences between classes. The last dataset looks at classes that differ in frequency content only. Each dataset contains three different classes of signals. For each class, we generated 100 signals to use for training each LDB algorithm. Another 1,000 signals are generated to test the constructed LDB feature spaces. For each LDB algorithm, we set the parameter $k_0 = 10$, i.e., we evaluate the goodness of each subspace using the top ten most discriminant coordinates. For each dataset, classifiers are trained using the top ten most discriminant features from each LDB algorithm. We repeat this process of generating training and test sets, computing features spaces, and applying classifiers ten times. The mean and standard deviation of the misclassification over the ten trials are presented in Table 12.1. As for the LDBK and LDBKASH algorithms, we use the symmetric relative entropy as the discriminant measure $D(\mathbf{p},\mathbf{q})$ throughout this article, i.e.,

$$D(\mathbf{p},\mathbf{q}) = J(\mathbf{p},\mathbf{q}) := \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} + q_i \log \frac{q_i}{p_i},$$

Table 12.1 Table of training and test waveforms misclassification for all examples discussed

	LDA		CT	
	Train	Test	Train	Test
Triangle waveform classification				
STD	11.40 % ± 1.10	22.18 % ± 1.18	3.93 % ± 0.78	30.58 % ± 1.82
LDBK	14.13 % ± 2.26	17.08 % ± 1.05	5.87 % ± 0.88	23.48 % ± 1.98
LDBKASH	13.17 % ± 2.18	16.60 % ± 0.85	4.90 % ± 1.31	22.36 % ± 1.34
LDBKEMD	13.90 % ± 1.75	16.38 % ± 0.84	6.27 % ± 1.91	23.82 % ± 2.10
Shape waveform classification				
STD	0.27 % ± 0.41	6.94 % ± 0.87	0.63 % ± 0.60	6.58 % ± 0.77
LDBK	4.10 % ± 2.62	4.62 % ± 1.66	0.87 % ± 0.59	3.96 % ± 1.27
LDBKASH	1.67 % ± 0.92	2.54 % ± 0.42	0.67 % ± 0.61	3.50 % ± 1.23
LDBKEMD	2.07 % ± 0.78	3.07 % ± 0.68	0.57 % ± 0.42	2.98 % ± 0.74
Bell waveform classification				
STD	5.50 % ± 1.57	28.93 % ± 4.81	2.67 % ± 1.31	20.78 % ± 1.80
LDBK	34.10 % ± 17.91	36.30 % ± 17.00	7.33 % ± 3.68	30.38 % ± 10.19
LDBKASH	26.07 % ± 6.35	28.40 % ± 5.87	5.00 % ± 1.61	24.38 % ± 4.57
LDBKEMD	16.33 % ± 6.25	19.40 % ± 5.81	4.27 % ± 0.81	17.91 % ± 1.66
Chirp waveform classification				
STD	0.00 % ± 0.00	0.00 % ± 0.00	1.63 % ± 0.81	18.91 % ± 1.15
LDBK	4.33 % ± 4.26	6.92 % ± 6.07	4.07 % ± 2.57	20.36 % ± 9.93
LDBKASH	18.37 % ± 9.39	22.86 % ± 9.26	5.93 % ± 2.02	32.42 % ± 7.74
LDBKEMD	0.00 % ± 0.00	0.00 % ± 0.00	0.00 % ± 0.00	1.21 % ± 0.51

Each example has results for all LDB algorithm using linear discriminate analysis and classification tree classifiers. STD above indicates the use of the standard coordinate system for representing signals, i.e., the raw signals are directly fed to the classifiers

Example 1. Triangular Waveforms

This example was originally examined in [2]. Later, Saito [9] extended the length of the signals from 21 to 32, so that the signals are of dyadic length and could be used to evaluate the performance of LDB. The example consists of three classes of signals which are formed from a convex linear combination of triangular waveforms. Specifically, the classes of signals are generated by the following formulas:

$$x^{(1)}(i) = uh_1(i) + (1-u)h_2(i) + \varepsilon(i) \qquad \text{for class 1,} \tag{12.4}$$

$$x^{(2)}(i) = uh_1(i) + (1-u)h_3(i) + \varepsilon(i) \qquad \text{for class 2,} \tag{12.5}$$

$$x^{(3)}(i) = uh_2(i) + (1-u)h_3(i) + \varepsilon(i) \qquad \text{for class 3,} \tag{12.6}$$

where $i = 1, \ldots, 32$, $h_1(i) = \max(6 - |i-7|, 0)$, $h_2(i) = h_1(i-8)$, $h_3(i) = h_1(i-4)$, u is a uniform random variable on $(0,1)$, and $\varepsilon(i)$'s are the i.i.d. standard normal variates. Five sample waveforms from each class are shown in Fig. 12.2. This example is convenient for performance evaluation because Breiman et al. [2]

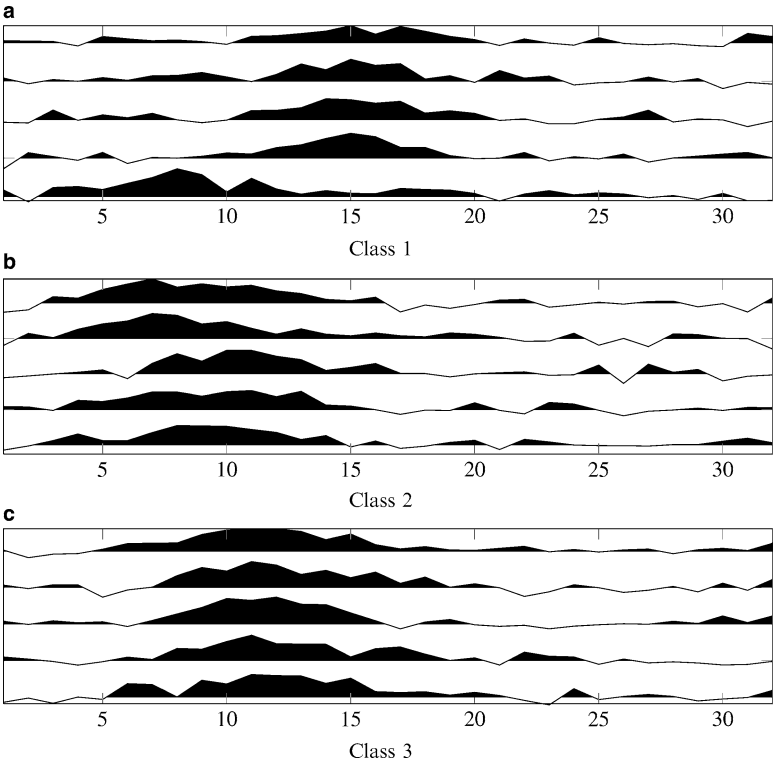

Fig. 12.2 Five sample waveforms from each class of triangular waveform dataset. Waveforms are generated using (12.4)–(12.6)

computed the Bayes error rate to be 14 %, which gives us an ideal classifier performance to expect. We use the wavelet packet dictionary with a 6-tap coiflet filter [3, Sect. 8.2] to compute the LDB.

All three algorithms exhibit similar classification performance and had similar basis selections. The LDA classifier performs the best with both training and test classification results near the Bayes error rate, see Table 12.1. If we look at the most discriminating basis vectors selected, Fig. 12.3a, we can identify the distinction that is being used to discriminate between the classes. The vectors with the greatest discriminant power are concentrated near the peaks of h_1, h_2, and h_3. The basis vectors are detecting the presence of the distinct characteristics of each class, i.e., the triangular peaks of h_1, h_2, and h_3. The coefficient plot for all three algorithms are similar, so we only show one in Fig. 12.3b. We see that each class lies on a linear manifold segment. The difficult signals to classify lie at the intersections of these linear manifold segments. These correspond to signals where one triangular peak is much more prominent. For this example, there is no extra benefit to examining the statistical distribution of the coefficients for each class. The time–frequency energy maps provide, as computed by LDBK, sufficient information for discrimination.

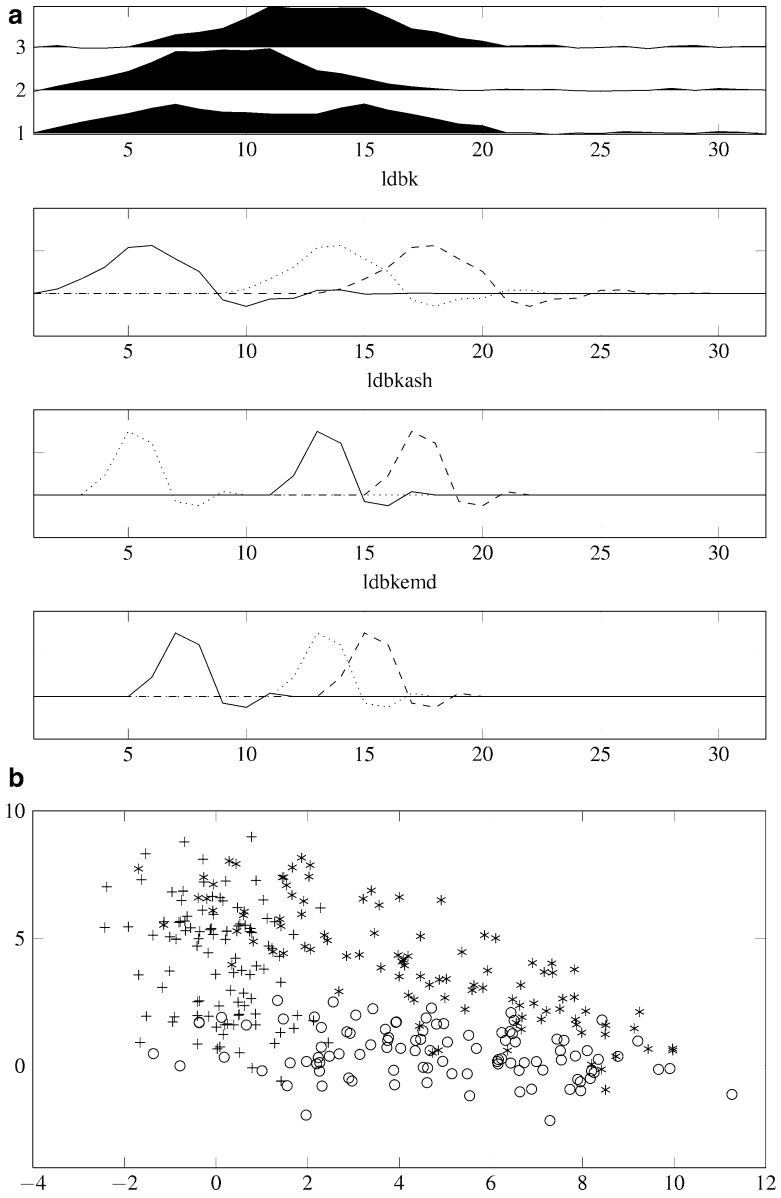

Fig. 12.3 (**a**) In the upper plot the mean training waveform for each of the triangular waveform dataset is shown. In the plot below it, the three most discriminant LDB vectors are shown for each LDB algorithm. (**b**) A scatter plot of the coefficients of the training signals in the top two most discriminating LDBK coordinates. The symbols, "*," "o," "+," represent class 1, class 2, class 3 signals, respectively

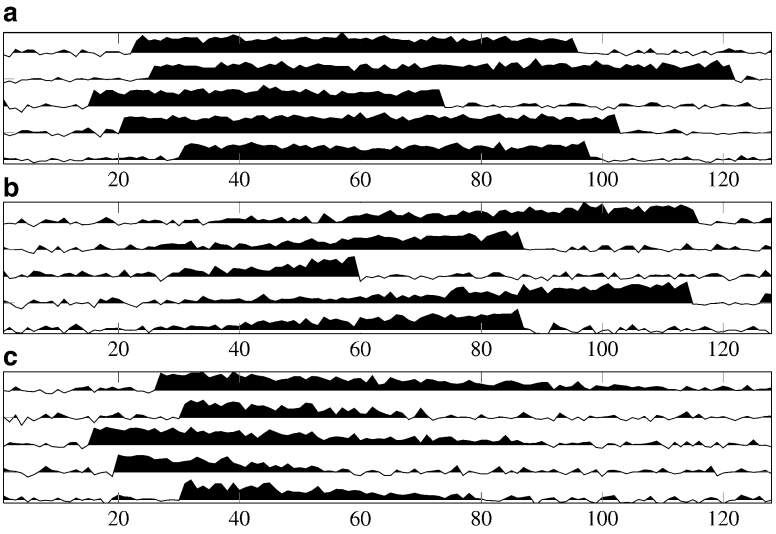

Fig. 12.4 Five sample waveforms from each class of the shape waveform dataset

Example 2. Shape Waveforms

Our second dataset is comprised of signal classes of different shapes (cylinder, bell, and funnel). All classes of signals are of finite duration, varied over when they appear. The cylinder signal class is a flat-amplitude (i.e., boxcar) signal. The bell signal class has a linearly increasing amplitude. And the funnel signal class has a linearly decreasing amplitude. More precisely, our signal classes are generated by the following formulae:

$$
\begin{aligned}
c(i) &= (6+\eta)\cdot\chi_{[a,b]}(i)+\varepsilon(i) && \text{cylinder,}\\
b(i) &= (6+\eta)\cdot\chi_{[a,b]}(i)\cdot(i-a)/(b-a)+\varepsilon(i) && \text{bell,}\\
f(i) &= (6+\eta)\cdot\chi_{[a,b]}(i)\cdot(b-i)/(b-a)+\varepsilon(i) && \text{funnel,}
\end{aligned}
$$

where $i = 1,\ldots,128$, a is an integer-valued uniform random variable on the interval [16,32], $b-a$ also obeys an integer-valued uniform distribution on [32,96], η and $\varepsilon(i)$'s are the i.i.d. standard normal variates, and $\chi_{[a,b]}(i)$ is the characteristic, or indicator, function on the interval $[a,b]$. Five example waveforms from each class can be seen in Fig. 12.4. We use the wavelet packet coefficients from an 18-tap coiflet filter [3, Sect. 8.2] to compute the LDB.

The ten most discriminating LDB vectors for each algorithm are shown in Fig. 12.5. Classification performance, shown in Table 12.1, for each algorithm is similar with a slight benefit to using LDBKASH or LDBKEMD. However, the top LDB vectors selected do vary for each algorithm. For all the algorithms, within the ten most discriminating vectors selected, there are features concentrated at

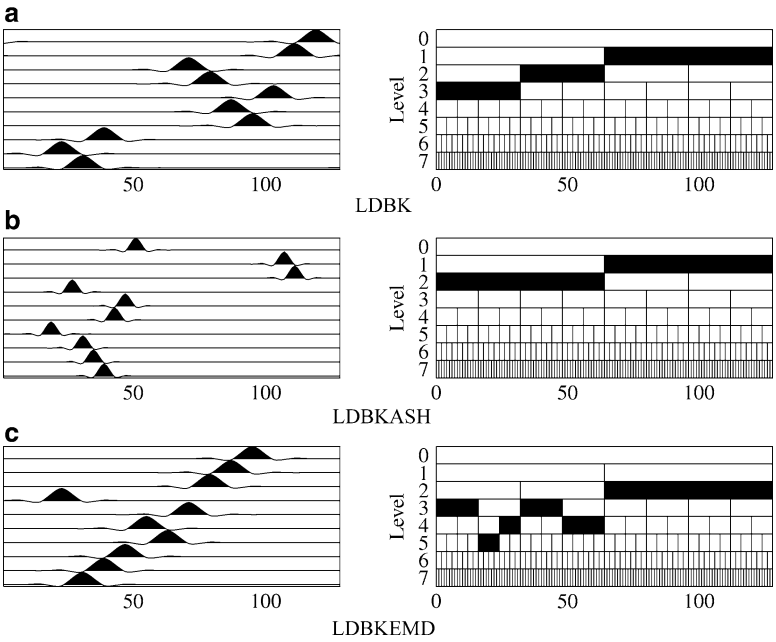

Fig. 12.5 Top ten LDB vectors selected for the shape waveform using different LDB algorithms. The subfigures in the right column show the selected subspaces indicated in *black*

the beginning and end of where we expect the cylinder, bell, or funnel to be. For LDBKEMD, we note that unlike LDBK and LDBKASH, there are vectors concentrated in the middle. While the ends emphasize the regions where we are likely to see the most dramatic change in signal characteristics, our signals do vary within the interval $[a,b]$. For our next example we will try to emphasize this by making our classes more similar.

Example 3. Bell Waveforms

In the previous example we noted that the different LDB algorithms selected somewhat different basis vectors. However, the differences had little impact on the classification performance since all the LDB algorithms selected features that focused on regions where there was a great deal of change between classes. For this example, we attempt to make the shape distinction more subtle by considering three bells with varying slopes. Specifically, the signals are generated by the following formulae:

$$b_1(i) = \chi_{[a,b]}(i) \cdot (i-a) \cdot S_1 + \varepsilon(i) \qquad \text{bell 1,}$$

$$b_2(i) = \chi_{[a,b]}(i) \cdot (i-a) \cdot S_2 + \varepsilon(i) \qquad \text{bell 2,}$$

$$b_3(i) = \chi_{[a,b]}(i) \cdot (i-a) \cdot S_3 + \varepsilon(i) \qquad \text{bell 3,}$$

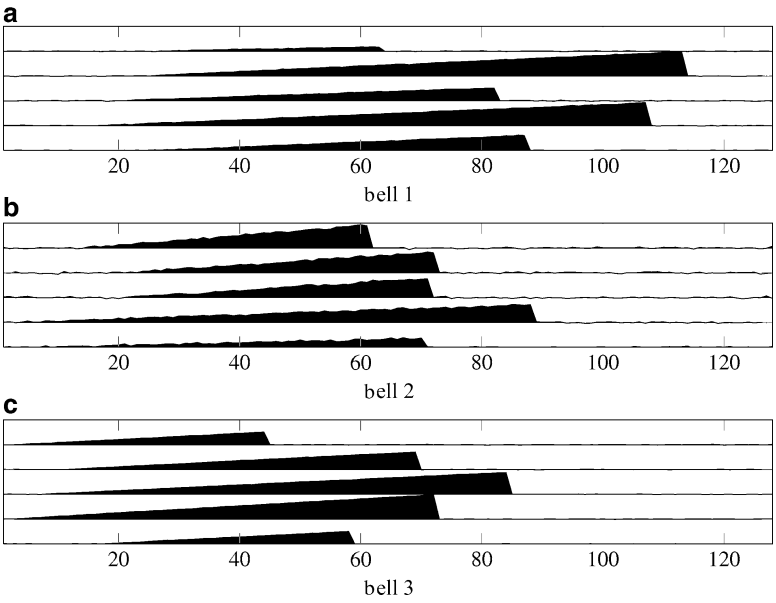

Fig. 12.6 Five sample waveforms from each class of the Bell waveform dataset

where $i = 1,\ldots,128$, a is an integer-valued uniform random variable on the interval [16,32], $b-a$ also obeys an integer-valued uniform distribution on [32,96], $\varepsilon(i)$'s are the i.i.d. standard normal variates, and $\chi_{[a,b]}(i)$ is the characteristic, or indicator, function on the interval $[a,b]$. The slopes (S_1, S_2, and S_3) are normal random variables with standard deviation 0.2 and respective means of 1, 0.5, and 2. Five sample waveforms from each class can be seen in Fig. 12.6. Visually it is quite difficult to distinguish between the classes of signals. For training, the LDB is computed from the wavelet packet coefficients with a 18-tap coiflet filter [3, Sect. 8.2].

Looking at the classification performance, shown in Table 12.1, we see a clear performance increase between using energy versus the distribution of energy. The LDBKASH and LDBKEMD algorithms show significantly lower average misclassification rates and standard deviation than LDBK. Looking at the ten selected basis vectors shown in Fig. 12.7, we see that the LDBKEMD vectors are varied in position across the length of the signal with the center location being the most important. The LDBKASH vectors have a narrower support and are concentrated toward the center of the signal. The LDBK vectors are concentrated toward the end of the signal where we would expect the greatest impact from the varying slope. However, there is also a chance of the signal not being there due to the randomness in the parameter b, the signal ending position. From the coefficient plot, Fig. 12.8, we can see that the top 3 features for LDBKASH and LDBKEMD correspond to the change in slope for each class. Bell 2 had the smallest slope, and we see in the coefficient plot, indicated with the "o" marker, that it is at the bottom.

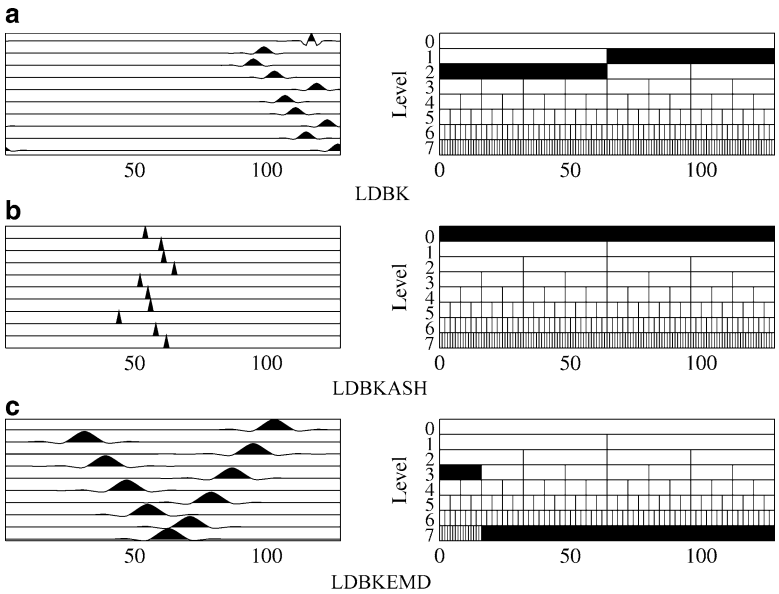

Fig. 12.7 Top ten LDB vectors selected for the bell waveforms using different LDB algorithms. The subfigures in the right column show the selected subspaces indicated in *black*

Then we see that bell 1, indicated with the "*" marker, follows after bell 2 with the intermediate slope. And last we see bell 3, indicated with the "+" marker, which had the largest slope.

Example 4. Chirp Waveforms
For our final example we will look at the ability of our LDB algorithms to distinguish between frequency differences in signals. Our dataset will consist of three different quadratic chirp waveforms with additive Gaussian noise. All chirps start with 200-Hz oscillations, sweep down to varied minimum frequency ranges, and go up to the 200-Hz level again in the end as the spectrogram in Fig. 12.9 shows. More specifically,

$$c_1(i) = \cos\left(2\pi\left[\frac{80}{3}t(i)^3 + 120t(i)\right]\right) + \varepsilon(i) \qquad \text{chirp 1,}$$

$$c_2(i) = \cos\left(2\pi\left[\frac{50}{3}t(i)^3 + 150t(i)\right]\right) + \varepsilon(i) \qquad \text{chirp 2,}$$

$$c_3(i) = \cos\left(2\pi\left[\frac{20}{3}t(i)^3 + 180t(i)\right]\right) + \varepsilon(i) \qquad \text{chirp 3,}$$

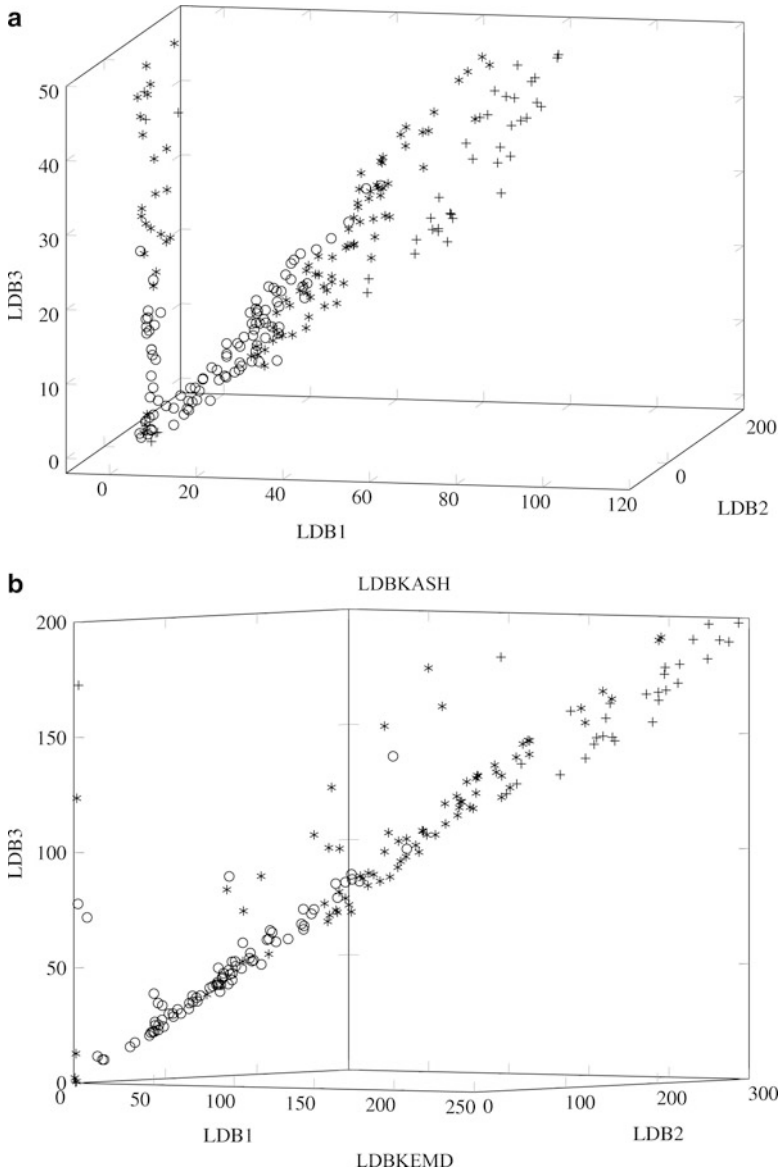

Fig. 12.8 The distribution of the coefficients in the top three LDB coordinates for bell waveform training and test datasets using (**a**) LDBKASH and (**b**) LDBKEMD algorithms. The symbols, "*," "o," "+," represent bell 1, bell 2, bell 3 signals, respectively

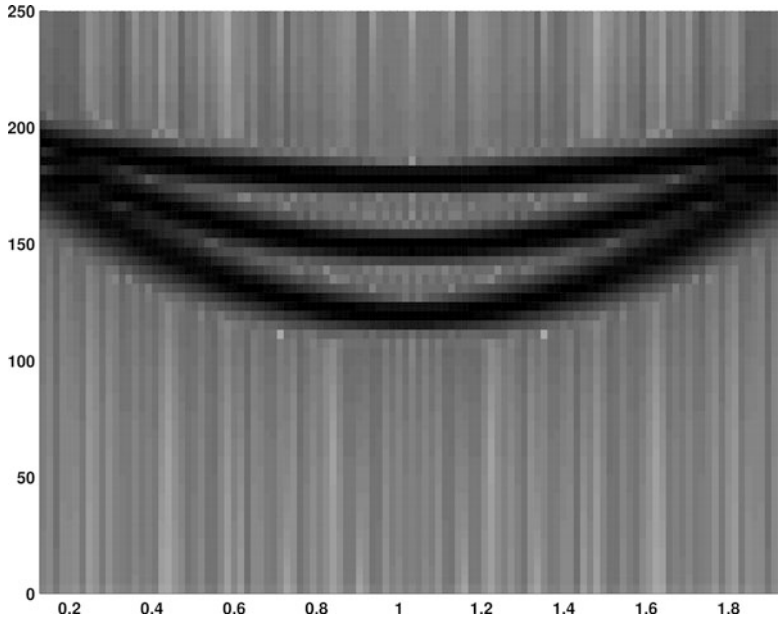

Fig. 12.9 Spectrogram of chirp waveforms for chirp waveform dataset. Horizontal axis is time (seconds). Vertical axis is frequency (Hz)

where $i = 1, \ldots, 1024$, $t(i) = i/512 - 1$, and $\varepsilon(i)$'s are i.i.d. standard normal variates. For training, the LDB is computed from the wavelet packet coefficients with a 18-tap coiflet filter.

From the result shown in Table 12.1, we see that LDBKEMD performs very well with zero misclassification for an LDA classifier. The LDBKASH algorithm performs the worst with a test misclassification rate around 23 % with an LDA classifier. For LDBK the test misclassification rate is around 7 % with an LDA classifier. If we look at the resulting LDB selected for each algorithm and top ten LDB vectors, Fig. 12.10, we see that the LDBKASH algorithm selected the root level with the most discriminating features being concentrated toward the center of the signal. The LDBK algorithm also selected features concentrated toward the center, but with wider support. On the other hand, the features selected by the LDBKEMD algorithm are quite different. Some of the LDBKEMD vectors are concentrated around the beginning and ending locations while the others have much wider supports in time.

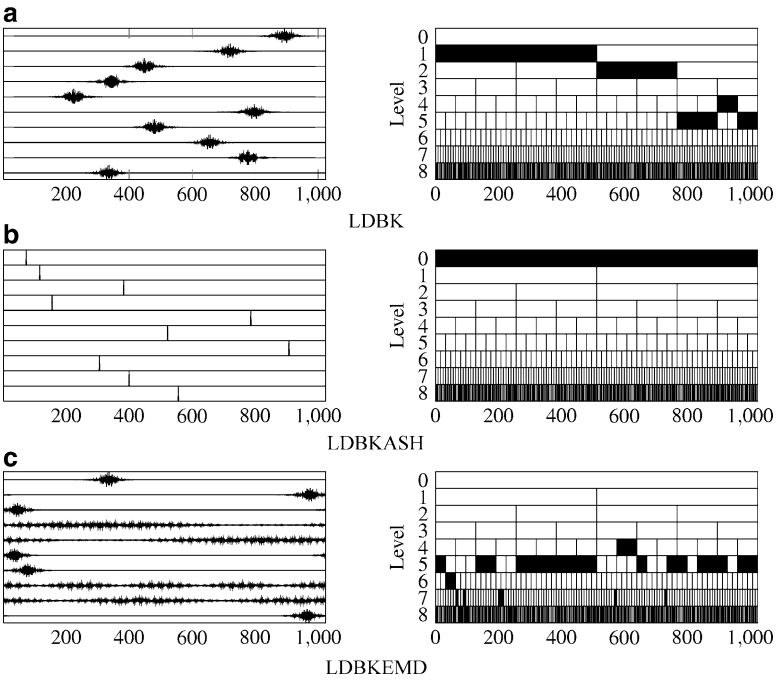

Fig. 12.10 Top ten LDB vectors selected for the chirp waveforms using different LDB algorithms. The subfigures in the right column show the selected subspaces indicated in *black*

12.5 Conclusion

LDB is an effective and computationally fast method for extracting discriminant features from signals. Furthermore, as we demonstrated in Sect. 12.4, the resulting feature space is interpretable; we know what is being used in each feature through the corresponding LDB basis vector and the expansion coefficients of the signals relative to that vector and therefore have a better understanding of when it will not be as effective and why. Also, we presented yet another LDB algorithm using a new discriminant measure based on signatures and EMD and demonstrated its capability of detecting features that can be missed using other versions of LDBs. Our EMD-based LDB can also be adapted to use new training data as it is provided. In comparison to LDBKASH, our LDBKEMD algorithm has fewer parameters to tweak and avoids the difficult task of estimating reliable epdfs, which make this new algorithm more robust. As demonstrated by our last two examples, it provides for better separation of classes with less training and a measure of discrimination that is less susceptible to outliers. However, as shown with our first two examples, in some situations expending extra effort, i.e., incorporating statistical behavior of the expansion coefficients, does not necessarily improve the performance.

Acknowledgments This research was partially supported by grants NSF DMS-0410406, ONR N00014-06-1-0615, N00014-07-1-0166, N00014-09-1-0041, and N00014-09-1-0318 as well as NSF VIGRE grant DMS-0135345. Further, support was provided from the Inhouse Laboratory Independent Research (ILIR) program at the Naval Surface Warfare Center, Panama City.

References

1. Basseville M (1989) Distance measures for signal processing and pattern recognition. Signal Process 18:349–369
2. Breiman L, Friedman JH, Olshen RA, Stone C (1993) Classification and regression trees. Chapman & Hall, Inc., New York
3. Daubechies I (1992) Ten lectures on wavelets. SIAM, Philadelphia, PA,
4. Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: data mining, inference, and prediction, 2nd edn. Springer, Berlin
5. Hillier FS, Lieberman GJ (1995) Introduction to mathematical programming. McGraw-Hill, San Francisco
6. Levina E, Bickel P (2001) The earth mover's distance is the mallow distance: Some insights from statistics. In: Proceedings of 8th IEEE International Conference on computer vision, pp 251–256
7. Rubner Y, Tomasi C (2000) Perceptual metrics for image database navigation. Kluwer Academic Publishers, Boston
8. Rubner Y, Tomasi C, Guibas LJ (2000) The earth mover's distance as a metric for image retrival. Int J Comput Vis 40:99–121
9. Saito N (1994) Local feature extraction and its applications using a library of bases. Ph.D thesis. Department of Mathematics, Yale University
10. Saito N, Coifman RR (1995) Local discriminant bases and their applications. J Math Imag Vis 5:337–358
11. Saito N, Coifman RR, Geshwind FB, Warner F (2002) Discriminant feature extraction using empirical probability density estimation and a local basis library. Pattern Recogn 35:2841–2852
12. Scott DW (1985) Averaged shifted histograms: effective nonparametric density estimators in several dimensions. Ann Stat 13:1024–1040

Part III
Statistical Analysis

Chapter 13
Characterizations of Certain Continuous Distributions

G.G. Hamedani

Abstract In designing a stochastic model for a particular modeling problem, an investigator will be vitally interested to know if their model fits the requirements of a specific underlying probability distribution. To this end, the investigator will vitally depend on the characterizations of the selected distribution. The Amoroso, SSK (Shakil–Singh–Kibria), SKS (Shakil–Kibria–Singh), SK (Shakil–Kibria), and SKS-type distributions have been suggested to have potential applications in modeling and are characterized here based on either a simple relationship between two truncated moments or a truncated moment of a function of the first order statistic or of a function of the *n*th order statistic, the two more interesting order statistics. We also present a characterization of SKS-type distribution based on the conditional expectation of adjacent generalized order statistics.

13.1 Introduction

In designing a stochastic model for a particular modeling problem, an investigator will be vitally interested to know if their model fits the requirements of a specific underlying probability distribution. To this end, the investigator will depend on the characterizations of the selected distribution. Generally speaking, the problem of characterizing a distribution is an important problem in various fields and has recently attracted the attention of many researchers. Consequently, various characterization results have been reported in the literature. These characterizations have been established in many different directions, one of which is in terms of the

G.G. Hamedani (✉)
Department of Mathematics, Statistics and Computer Science,
Marquette University, Milwaukee, WI 53201-1881, USA
e-mail: g.hamedani@marquette.edu

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_13,

truncated moments. We like to mention here the works of Galambos and Kotz [8], Kotz and Shanbhag [20], Glänzel [9,10], Glänzel et al. [12], Glänzel and Hamedani [11], and Hamedani [13–15].

Recently, Ahsanullah and Hamedani [3] characterized the power function and the beta of the first-kind distributions based on a truncated moment of the nth order statistic and first order statistic, respectively, extending some known characterizations of the power function and the uniform distributions (see [1,2]). Following [3], Hamedani et al. [17] characterized the following distributions based on a truncated moment of the first order statistic: Burr type XII (a special case), generalized beta 1, generalized beta 2 (the last two family of distributions unify many distributions employed for size distribution of income [21]), generalized Pareto, Pareto of first kind, and Weibull. The following families of distributions were also mentioned in [17] as special cases of Weibull: Burr type X, chi-square, extreme value type 2, gamma and Rayleigh. Hamedani [15] established characterizations of 31 more continuous univariate distributions based on a truncated moment of the first order statistic or of the nth order statistic or of a function of the first order statistic or of a function of the nth order statistic.

Various systems of distributions have been constructed to provide approximations to a wide variety of distributions (see, e.g., [18]). These systems are designed with the requirements of ease of computation and feasibility of algebraic manipulation. To meet the requirements, there must be as few parameters as possible in defining a member of the system.

One of these systems is Pearson system. A continuous distribution belongs to this system if its probability density function (pdf) $f(x)$ satisfies a differential equation of the form

$$\frac{1}{f(x)}\frac{\mathrm{d}f(x)}{\mathrm{d}x} = -\frac{x+a}{bx^2+cx+d} \tag{13.1}$$

where a, b, c, and d are real parameters such that $f(x)$ is a pdf. The shape of the pdf depends on the values of these parameters. Pearson [22] classified the different shapes into a number of types I–VII (see Appendix A). Many well-known distributions are special cases of Pearson-type distributions which are characterized in [15], Sects. 3–6.

Another system is Burr system, [6], which like Pearson system, has various types I–XII. This system, however, is not as involved and as basic as Pearson system. There are also families of distributions like extreme value and Pareto which have different kind or type members. These distributions are also characterized in [15] Sects. 3–6.

The families discussed in Sects. 5.3 and 5.5 of [15] were first introduced in [5] in the context of minimum dynamic discrimination information approach to probability modeling. The families in Sects. 5.8 and 5.9 of [15] appeared in [4], which were shown to be maximum dynamic entropy models.

The presentation of the content of this work is as follows. Sect. 13.2 deals with introduction of Amoroso distribution, the natural unification of the gamma and extreme value distributions. In Sect. 13.3, we present characterizations of

the Amoroso distribution based on the truncated moment of a function of first order statistic and of a function of nth order statistic. Section 13.4 is devoted to definitions of SSK, SKS, SK, and SKS-type distributions. In Sect. 13.5, we present characterizations of SSK distribution based on a simple relationship between two truncated moments. Section 13.6 deals with the characterizations of SKS-type distribution based on the truncated moment of a function of first order statistic and of a function of nth order statistic. We also give a characterization of this distribution based on conditional expectation of adjacent generalized order statistics. In Sect. 13.7 we present a characterization of SK distribution based on a simple relation between two truncated moments. Finally, in Sect. 13.8 we have a very short concluding remark. For further characterization results in this direction, we refer the reader to Ahsanullah and Hamedani [3], Hamedani et al. [17], and Hamedani [15].

13.2 The Amoroso Distribution

This section deals with introducing the Amoroso distribution. It is pointed out by Crooks [7] that the Amoroso distribution, a four parameter, continuous, univariate, unimodel pdf with semi-infinite range, was originally developed to model lifetimes (see [7] for more details). Moreover, many well-known and important distributions are special cases or limiting forms of the Amoroso distribution. Table 13.1 is taken (with permission from G.E. Crooks for which we are grateful to him) from [7], which shows 35 special and four limiting cases of the Amoroso distribution. These distributions and their importance in different fields of studies have been discussed in detail in [7].

The pdf of the Amoroso distribution is given by

$$f(x;a,\alpha,\tau,k) = \frac{1}{\Gamma(k)}\left|\frac{\tau}{\alpha}\right|\left(\frac{x-a}{\alpha}\right)^{\tau k-1}\exp\left\{-\left(\frac{x-a}{\alpha}\right)^{\tau}\right\} \tag{13.2}$$

for x, a, α, τ in $\mathbb{R}, k > 0$, support $x \geq a$ if $\alpha > 0$, $x \leq a$ if $\alpha < 0$. As usual, $\Gamma(k) = \int_0^\infty u^{k-1}\,\mathrm{e}^{-u}\,\mathrm{d}u$, for $k > 0$.

The four real parameters of the Amoroso distribution consist of a location parameter a, a scale parameter α, and two shape parameters, τ and k. The shape parameter k is positive, and most of the time, an integer, $k = n$, or half-integer $k = \frac{m}{2}$. If the random variable X has the Amoroso distribution with parameters a, α, τ and $k > 0$, we write $X \sim \text{Amoroso}(a,\alpha,\tau,k)$.

For further details about the distributions listed in Table 13.1 and their applications, we refer the reader to Crooks [7].

We give Table 13.2 displaying four cases based on the signs of α and τ for the random variable $X \sim \text{Amoroso}(a,\alpha,\tau,k)$. Without loss of generality we assume $a = 0$ throughout this work.

Table 13.1 The Amoroso family of distributions

Amoroso	a	α	k	τ	
Stacy	0	·	·	·	
Gen. Fisher-Tippett	·	·	n	·	
Fisher-Tippett	·	·	1	·	
Fréchet	·	·	1	<0	
Generalized Fréchet	·	·	n	<0	
Scaled inverse chi	0	·	$\frac{m}{2}$	-2	
Inverse chi	0	$\frac{1}{\sqrt{2}}$	$\frac{m}{2}$	-2	
Inverse Rayleigh	0	·	1	-2	
Pearson type V	·	·	·	-1	
Inverse gamma	0	·	·	-1	
Scaled inverse chi-square	0	·	$\frac{m}{2}$	-1	
Inverse chi-square	0	$\frac{1}{2}$	$\frac{m}{2}$	-1	
Lévy	·	·	$\frac{1}{2}$	-1	
Inverse exponential	0	·	1	-1	
Pearson type III	·	·	·	1	
Gamma	0	·	·	1	
Erlang	0	>0	n	1	
Standard gamma	0	1	·	1	
Scaled chi-square	0	·	$\frac{m}{2}$	1	
Chi-square	0	2	$\frac{m}{2}$	1	
Shifted exponential	·	·	1	1	
Exponential	0	·	1	1	
Standard exponential	0	1	1	1	
Wien	0	·	4	1	
Nakagami	·	·	·	2	
Scaled chi	0	·	$\frac{m}{2}$	2	
Chi	0	$\sqrt{2}$	$\frac{m}{2}$	2	
Half-normal	0	·	$\frac{1}{2}$	2	
Rayleigh	0	·	1	2	
Maxwell	0	·	$\frac{3}{2}$	2	
Wilson-Hilferty	0	·	·	3	
Generalized Weibull	·	·	n	>0	
Weibull	·	·	1	>0	
Pseudo-Weibull	·	·	$1+\frac{1}{\tau}$	>0	
Stretched exponential	0	·	1	>0	
Log–gamma	·	·	·	·	$\lim_{\tau\to\infty}$
Power law	·	·	$\frac{1-p}{\tau}$	·	$\lim_{\tau\to 0}$
Log-normal	·	·	$\frac{1}{(\tau\sigma)^2}$	·	$\lim_{\tau\to 0}$
Normal	·	·	·	1	$\lim_{k\to\infty}$

m, n positive integers

For $\alpha > 0$ and $\tau > 0$, Amoroso$(0,\alpha,\tau,k) = GG(\alpha,\tau,k)$, generalized gamma distribution. The characterizations given here are valid for the distributions of $-X$ (when $\alpha < 0, \tau > 0$), $\frac{1}{X}$ (when $\alpha > 0, \tau < 0$), and $-\frac{1}{X}$ (when $\alpha < 0, \tau < 0$).

Table 13.2 Special rvs with generalized gamma distributions

	$\tau > 0$	$\tau < 0$
$\alpha > 0$	$X \sim GG(\alpha, \tau, k)$	$\frac{1}{X} \sim GG\left(\frac{1}{\alpha}, -\tau, k\right)$
$\alpha < 0$	$-X \sim GG(-\alpha, \tau, k)$	$-\frac{1}{X} \sim GG\left(-\frac{1}{\alpha}, -\tau, k\right)$

Table 13.2 shows that for $\alpha < 0$ a simple change of parameters $\alpha' = -\alpha$ will produce the cases on the second row of the table. So, we investigate here the characterizations of the distribution of X when $\alpha > 0$ and $\tau > 0$ (**Case I**) and when $\alpha > 0$ and $\tau < 0$ (**Case II**).

Case I The pdf of the Amoroso random variable is given by

$$f(x; \alpha, \tau, k) = \frac{\tau}{\alpha \Gamma(k)} \left(\frac{x}{\alpha}\right)^{\tau k - 1} \exp\left\{-\left(\frac{x}{\alpha}\right)^{\tau}\right\}, x \geq 0 \tag{13.3}$$

where all three parameters α, τ, and k are positive.

Case II Letting $\gamma = -\tau > 0$, the *pdf* of the Amoroso random variable X is now

$$f(x; \alpha, \gamma, k) = \frac{\gamma}{\alpha \Gamma(k)} \left(\frac{x}{\alpha}\right)^{-(\gamma k + 1)} \exp\left\{-\left(\frac{x}{\alpha}\right)^{-\gamma}\right\}, x \geq 0 \tag{13.4}$$

where all three parameters α, γ, and k are positive.

The cumulative distribution function (*cdf*), F, corresponding to (13.2) and (13.4) are, respectively,

$$F(x) = \frac{1}{\Gamma(k)} \int_0^{\left(\frac{x}{\alpha}\right)^{\tau}} u^{k-1} \mathrm{e}^{-u} \mathrm{d}u, \; x \geq 0 \tag{13.5}$$

and

$$F(x) = 1 - \frac{1}{\Gamma(k)} \int_0^{\left(\frac{x}{\alpha}\right)^{-\gamma}} u^{k-1} \mathrm{e}^{-u} \mathrm{d}u, \; x \geq 0 \tag{13.6}$$

13.3 Characterizations of the Amoroso Distribution

This section is devoted to the characterizations of the Amoroso distribution based on truncated moment of a function of first order statistic as well as on truncated moment of a function of nth order statistic. As we pointed out in Sect. 13.2, we will present our characterizations of the Amoroso distribution in two separate cases as follows. First, however, we give the *pdf* of the jth order statistic.

Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be the order statistics of a random sample of size n from a continuous *cdf* F with the corresponding *pdf* f. The random variable $X_{j:n}$ denotes the jth order statistic from a random sample of n independent random

variables $X_1, X, \ldots, X_n$ with common $cdf\ F$. Then, the $pdf\ f_{j:n}$ of $X_{j:n}$, $j = 1, 2, \ldots, n$ is given by

$$f_{j:n}(x) = \frac{n!}{(j-1)!(n-j)!} f(x) (F(x))^{j-1} (1 - F(x))^{n-j}.$$

The $pdfs$ of the first and the nth order statistics are, respectively

$$f_{1:n}(x) = n f(x) (1 - F(x))^{n-1} \quad and \quad f_{n:n}(x) = n f(x) (F(x))^{n-1}.$$

13.3.1 Characterizations of the Amoroso PDF (Case I)

In this subsection we present a characterization of the Amoroso distribution with pdf (13.3) in terms of a truncated moment of a function of the nth order statistic. We define the function

$$\gamma_1 \left[k; \left(\frac{x}{\alpha}\right)^{\tau}\right] = \int_0^{\left(\frac{x}{\alpha}\right)^{\tau}} u^{k-1} e^{-u} du \quad for\ \alpha > 0, \tau > 0, k > 0,\ and\ x \geq 0.$$

Proposition 13.3.1.1. *Let $X : \Omega \to [0, \infty)$ be a continuous random variable with $cdf\ F$. The pdf of X is (13.3) if and only if*

$$E\left\{\gamma_1 \left[k; \left(\frac{X_{n:n}}{\alpha}\right)^{\tau}\right] | X_{n:n} < t\right\} = \frac{n}{n+1} \gamma_1 \left[k; \left(\frac{t}{\alpha}\right)^{\tau}\right], \ t > 0. \tag{13.7}$$

Proof. Let X have pdf (13.3), then $F(x)$ is given by (13.5). Now using (13.5) on the left-hand side of (13.7), we arrive at

$$\begin{aligned} E\left\{\gamma_1 \left[k; \left(\frac{X_{n:n}}{\alpha}\right)^{\tau}\right] | X_{n:n} < t\right\} &= \frac{\int_0^t \gamma_1 \left[k; \left(\frac{x}{\alpha}\right)^{\tau}\right] d\left((F(x))^n\right)}{(F(t))^n} \\ &= \gamma_1 \left[k; \left(\frac{t}{\alpha}\right)^{\tau}\right] - \frac{\Gamma(k)}{n+1} F(t) \\ &= \frac{n}{n+1} \gamma_1 \left[k; \left(\frac{t}{\alpha}\right)^{\tau}\right] \quad t > 0. \end{aligned}$$

Now, assume (3.1.1) holds, then

$$\int_0^t \gamma_1 \left[k; \left(\frac{x}{\alpha}\right)^{\tau}\right] d\left((F(x))^n\right) = \frac{n}{n+1} \gamma_1 \left[k; \left(\frac{t}{\alpha}\right)^{\tau}\right] (F(t))^n, \quad t > 0.$$

Differentiating both sides of the above equation with respect to t and upon simplification, we obtain

$$\frac{f(t)}{F(t)} = \frac{\frac{d}{dt}\gamma_1\left[k;\left(\frac{t}{\alpha}\right)^{\tau}\right]}{\gamma_1\left[k;\left(\frac{t}{\alpha}\right)^{\tau}\right]}, \quad t>0.$$

Integrating both sides of the last equation with respect to t from x to ∞, and in view of the fact that $\lim_{t\to\infty}\gamma_1\left[k;\left(\frac{t}{\alpha}\right)^{\tau}\right] = \Gamma(k)$, we obtain (13.5) which completes the proof. □

Remark 13.3.1.2. For $k=1$, the following characterization in terms of the first order statistic is given for (13.3) (see [17], Subsection (vi)).

Proposition 13.3.1.3. *Let $X:\Omega\to\mathbb{R}^+$ be a continuous random variable with cdf F such that $\lim_{x\to\infty} x^{\tau}(1-F(x))^n = 0$. Then X has pdf (13.3) (with $k=1$) if and only if*

$$E\left[X_{1:n}^{\tau}|X_{1:n}>t\right] = t^{\tau} + \frac{\alpha^{\tau}}{n}, t>0.$$

13.3.2 Characterizations of the Amoroso PDF (Case II)

In this subsection we present a characterization of the Amoroso distribution with *pdf* (13.4) in terms of a truncated moment of a function of the first order statistic.

Proposition 13.3.2.1. *Let $X:\Omega\to[0,\infty)$ be a continuous random variable with cdf F. The pdf of X is (13.4) if and only if*

$$E\left\{\gamma_1\left[k;\left(\frac{X_{1:n}}{\alpha}\right)^{-\gamma}\right]|X_{1:n}>t\right\} = \frac{n}{n+1}\gamma_1\left[k;\left(\frac{t}{\alpha}\right)^{-\gamma}\right], \ t>0. \tag{13.8}$$

Proof. Let X have *pdf* (13.4), then $F(x)$ is given by (13.6), and

$$\begin{aligned} E\left\{\gamma_1\left[k;\left(\frac{X_{1:n}}{\alpha}\right)^{-\gamma}\right]|X_{1:n}>t\right\} &= \frac{\int_t^{\infty}\gamma_1\left[k;\left(\frac{x}{\alpha}\right)^{-\gamma}\right]nf(x)(1-F(x))^{n-1}\,dx}{(1-F(t))^n} \\ &= \gamma_1\left[k;\left(\frac{t}{\alpha}\right)^{-\gamma}\right] - \frac{\Gamma(k)}{n+1}(1-F(t)) \\ &= \frac{n}{n+1}\gamma_1\left[k;\left(\frac{t}{\alpha}\right)^{-\gamma}\right], \quad t>0. \end{aligned}$$

Now, assume (13.8) holds, then

$$\int_t^{\infty}\gamma_1\left[k;\left(\frac{x}{\alpha}\right)^{-\gamma}\right]nf(x)(1-F(x))^{n-1}\,dx = \frac{n}{n+1}\gamma_1\left[k;\left(\frac{t}{\alpha}\right)^{-\gamma}\right](1-F(t))^n, \quad t>0.$$

Differentiating both sides of the above equation with respect to t and upon simplification, we obtain

$$-\frac{f(t)}{1-F(t)}=\frac{\frac{\mathrm{d}}{\mathrm{d}t}\gamma_1\left[k;\left(\frac{t}{\alpha}\right)^{-\gamma}\right]}{\gamma_1\left[k;\left(\frac{t}{\alpha}\right)^{-\gamma}\right]},\quad t>0.$$

Integrating both sides of this equation with respect to t from 0 to x, and in view of the fact that $\lim_{t\to 0}\gamma_1\left[k;\left(\frac{t}{\alpha}\right)^{-\gamma}\right]=\Gamma(k)$, we obtain (13.6). □

Remark 13.3.2.2. For $k=1$, the following characterization in terms of the nth order statistic is given for (13.4) (see [15], Subsect. 4.2).

Proposition 13.3.2.3. *Let $X:\Omega\to\mathbb{R}^+$ be a continuous random variable with cdf F such that $\lim_{x\to 0}x^{-\gamma}(F(x))^n=0$. Then X has pdf (13.4) (with $k=1$) if and only if*

$$E\left[X_{n:n}^{-\gamma}|X_{n:n}<t\right]=t^{-\gamma}+\frac{1}{n\alpha^{\gamma}},\quad t>0.$$

13.4 The SSK (Shakil–Singh–Kibria), SKS (Shakil–Kibria–Singh), SKS-Type, and SK (Shakil–Kibria) Distributions

In this section we will give the definitions of SSK, SKS, SKS-type, and SK distributions in Subsects. 13.4.1–13.4.4, respectively. Recently, some researchers have considered a generalization of (13.1) given by

$$\frac{1}{f(x)}\frac{\mathrm{d}f(x)}{\mathrm{d}x}=\frac{\sum_{j=0}^{m}a_jx^j}{\sum_{j=0}^{m}b_jx^j},\tag{13.9}$$

where $m,n\in\mathbb{N}/\{0\}$ and the coefficients $a_j{}'s$, $b_j{}'s$ are real parameters. The system of continuous univariate *pdfs* generated by (13.9) is called generalized Pearson system which includes a vast majority of continuous *pdfs*.

13.4.1 SSK Distribution (Product Distribution Based on the Generalized Pearson Differential Equation)

Shakil et al. [25] consider 13.10 when $m=2$, $n=1$, $b_0=0$, $b_1\neq 0$, and $x>0$. The solution of this special case is an interesting three parameter distribution with *pdf* f given by

$$f(x;\alpha,\beta,\nu)=C_1x^{\nu}\exp\left(-\alpha x^2-\beta x\right),\ \ x>0,\alpha>0,\beta>0,\nu>0,\tag{13.10}$$

where $\alpha = -\frac{a_2}{2b_1}$, $\beta = -\frac{a_1}{b_1}$, $\nu = \frac{a_0}{b_1}$, and $b_1 \neq 0$ are parameters and C_1 is the normalizing constant.

Remark 13.4.1.1. A special case of equation (13.4), with $\gamma = 2$, will also have a solution of the form (13.10) as well.

The family of the distributions represented by *pdf* 13.10 can be expressed in terms of confluent hypergeometric functions of Tricomi and Kummer. As pointed out in [25], it is a rich family which includes the product of exponential and Rayleigh *pdfs*, the product of gamma and Rayleigh *pdfs*, the product of gamma and Rice *pdfs*, the product of gamma and normal *pdfs*, and the product of gamma and half-normal *pdfs*, among others. For detailed treatment (theory and applications) of this family we refer the reader to [25]. The family of SSK distributions will be characterized in Sect. 13.5.

13.4.2 SKS Distribution

Shakil et al. [24] consider (13.9) when $m = 2p$, $n = p+1$, $a_j = 0$, $j = 1,2,\ldots,p-1$, $p+1,\ldots,2p-1 = 0$; $b_j = 0$, $j = 1,2,\ldots,p$, $b_{p+1} \neq 0$, and $x > 0$. The solution of this special case is an interesting four parameter distribution with *pdf f* (using their notation) given by

$$f(x;\alpha,\beta,\nu,p) = C_2 x^{\nu-1} \exp\left(-\alpha x^p - \beta x^{-p}\right), \quad x > 0, \alpha \geq 0, \beta \geq 0, \nu \in \mathbb{R}, \tag{13.11}$$

where $\alpha = -\frac{a_{2p}}{pb_{p+1}}$, $\beta = \frac{a_p}{pb_{p+1}}$, $\nu = \frac{(a_p+b_{p+1})}{b_{p+1}}$, $b_{p+1} \neq 0$,and $p \in \mathbb{N} \ /\{0\}$ are parameters and C_2 is the normalizing constant.

Shakil et al. [24] classified their newly proposed family into the following three classes:

Class I. $\alpha > 0, \beta = 0, \nu > 0,$ and $p \in \mathbb{N}/\{0\}$.

Class II. $\alpha = 0, \beta > 0, \nu < 0,$ and $p \in \mathbb{N}/\{0\}$.

Class III. $\alpha > 0, \beta > 0, \nu \in \mathbb{R},$ and $p \in \mathbb{N}/\{0\}$.

Shakil et al. [24] pointed out that they found their "newly proposed model fits better than gamma, log-normal and inverse Gaussian distributions in the fields of biomedicine, demography, environmental and ecological sciences, finance, lifetime data, reliability theory, traffic data, etc. They hope that the findings of their paper will be useful for the practitioners in various fields of theoretical and applied sciences." They also pointed out that "It appears from literature that not much attention has been paid to the study of the family of continuous *pdfs* that can be generated as a solution of the generalized Pearson differential equation (13.11), except three papers cited in [24]." For a detailed treatment of the above-mentioned three cases and their

significance as well as related statistical analysis, we refer the reader to [24]. These cases were characterized in Hamedani [16] based on a simple relationship between two truncated moments.

13.4.3 SKS-Type Distribution

The SKS distribution has support in $(0,\infty)$, and one may be interested in similar distribution with bounded support. We would like to present here a distribution with bounded support, which we call it SKS type given by the pdf

$$f(x;\alpha,\beta,p) = Cpx^{-(p+1)}\left(\beta-\alpha x^{2p}\right)\exp\left(-\alpha x^{p}-\beta x^{-p}\right), \quad 0<x<\left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}, \tag{13.12}$$

where $\alpha>0$, $\beta>0$, and $p\in\mathbb{R}^{+}$ are parameters and $C=\exp\left(2\sqrt{\alpha\beta}\right)$ is the normalizing constant.

Remark 13.4.3.1. We do not require p to be a positive integer in (13.12). If, however, $p\in\mathbb{N}\,/\,\{0\}$, then (13.12) will be a member of the generalized Pearson system defined via (13.9)

$$\frac{1}{f(x)}\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{\beta^2 p-\beta(p+1)x^{p}-2\alpha\beta p x^{2p}-\alpha(p-1)x^{3p}+\alpha^2 p x^{4p}}{\beta x^{p+1}-\alpha x^{3p+1}}.$$

The cdf F corresponding to the pdf (13.12) is

$$F(x) = C\exp\left(-\alpha x^{p}-\beta x^{-p}\right), 0<x<\left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}. \tag{13.13}$$

The family of SKS-type distributions will be characterized in Sect. 13.6.

13.4.4 SK Distribution

Shakil and Kibria [23] consider a solution of (13.9) for $m=p$, $n=p+1$, $a_j=0$, $j=1,2,\ldots,p-1$, $b_j=0$, $j=0,1,\ldots,p$, $a_p\neq 0$, $b_1\neq 0$, $b_{p+1}\neq 0$, and $x>0$. This special five-parameter solution is given by

$$f(x;\alpha,\beta,\nu,\tau,p) = C_3 x^{\nu-1}\left(\alpha x^{p}+\beta\right)^{-\tau}, \quad x>0,\alpha>0,\beta>0,\nu>0,\tau>0,p\in\mathbb{N}/\{0\}, \tag{13.14}$$

where α, β, ν, τ, p are parameters, $\tau > \frac{\nu}{p}$, and C_3 is the normalizing constant. We refer the reader to [23] for further details and statistical analyses related to this family.

Final Remark of Sect. 13.4. In view of (13.9), we would like to make the observation that the *pdf* f of a sub-family of the Amoroso family satisfies the generalized Pearson differential equation (13.9) with, of course, appropriate boundary condition. For $a = 0$, $\alpha > 0$ $(or\ \ \alpha < 0)$, $\tau = -\gamma$, $\gamma \in \mathbb{N} / \{0\}$, and $k > 0$, the *pdf* f given by (13.2) satisfies (13.9) with $a_0 = \gamma\alpha^{\gamma}$, $a_j = 0$, $j = 1, 2, \ldots, \gamma - 1$, $a_{\gamma} = -(\gamma k + 1)$; $b_j = 0$, $j = 0, 1, \ldots, \gamma$, and $b_{\gamma+1} = 1$, i.e.,

$$\frac{1}{f(x)} \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{\gamma\alpha^{\gamma} - (\gamma k + 1)x^{\gamma}}{x^{\gamma+1}}.$$

For $a = 0$, $\alpha > 0$ $(or\ \alpha < 0)$, $\tau = \gamma$, $\gamma \in \mathbb{N} / \{0\}$, and $k > 0$, the *pdf* f given by (2.1) satisfies (4.1) with $a_0 = \gamma k - 1$, $a_j = 0$, $j = 1, 2, \ldots, \gamma - 1$, $a_{\gamma} = -\gamma\alpha^{-\gamma}$; $b_0 = 0$, and $b_1 = 1$, i.e.,

$$\frac{1}{f(x)} \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{(\gamma k - 1) - \gamma\alpha^{-\gamma}x^{\gamma}}{x}.$$

13.5 Characterizations of the SSK Distribution

In this section we present characterizations of the *pdf* (13.10) in terms of a simple relationship between two truncated moments. Our characterization results presented here will employ an interesting result due to Glänzel [9], which is stated here (Theorem G) for the sake of completeness.

Theorem G. Let $(\Omega, \mathscr{F}, P)$ be a given probability space and let $H = [a, b]$ be an interval for some $a < b$ $(a = -\infty$ and $b = +\infty$ might as well be allowed). Let $X : \Omega \to H$ be a continuous random variable with the distribution function F and let g and h be two real functions defined on H such that

$$E[g(X) | X \geq x] = E[h(X) | X \geq x] \lambda(x), x \in H$$

is defined with some real function λ. Assume that $g, h \in C^1(H)$, $\lambda \in C^2(H)$, and F is twice continuously differentiable and strictly monotone function on the set H. Finally, assume that the equation $h\lambda = g$ has no real solution in the interior of H. Then F is uniquely determined by the functions g, h, and λ, particularly

$$F(x) = \int_a^x C \left| \frac{\lambda'(u)}{\lambda(u) h(u) - g(u)} \right| \exp(-s(u)) \mathrm{d}u,$$

where the function s is a solution of the differential equation

$$s' = \frac{\lambda' h}{\lambda h - g}$$

and C is a constant, chosen to make $\int_H \mathrm{d}F = 1$.

Remark 13.5.1. In Theorem G, the interval H need not be closed.

Proposition 13.5.2. *Let $X : \Omega \to (0,\infty)$ be a continuous random variable and let $h(x) = x^{1-\nu}\exp(\beta x)$ for $x \in (0,\infty)$. The pdf of X is (13.10) if and only if there exist functions g and λ defined in Theorem G satisfying the differential equation*

$$\frac{\lambda'(x)}{\lambda(x)h(x) - g(x)} = 2\alpha x^{\nu}\exp(-\beta x), \quad x > 0. \tag{13.15}$$

Proof. Let X have pdf (13.10) and let

$$g(x) = x^{1-\nu}\left(\alpha + \beta x^{-1}\right), \quad x > 0$$

and

$$\lambda(x) = 2\alpha\exp(-\beta x), \quad x > 0.$$

Then

$$(1 - F(x))\, E[h(X)\,|X \geq x] = \frac{C_1}{2\alpha}\exp\left(-\alpha x^2\right), \quad x > 0,$$

$$(1 - F(x))\, E[g(X)\,|X \geq x] = C_1\exp\left(-\alpha x^2 - \beta x\right), \quad x > 0,$$

where C_1 is a constant. We also have

$$\lambda(x)h(x) - g(x) = -\beta x^{-\nu} < 0 \quad for \ x > 0.$$

The differential equation (13.15) clearly holds.

Conversely, if g and λ satisfy the differential equation (13.15) , then

$$s'(x) = \frac{\lambda'(x)h(x)}{\lambda(x)h(x) - g(x)} = 2\alpha x, \quad x > 0,$$

and hence

$$s(x) = \alpha x^2, \quad x > 0.$$

□

Now from Theorem G, X has pdf (13.10).

Corollary 13.5.3. *Let $X : \Omega \to (0,\infty)$ be a continuous random variable and let $h(x) = x^{1-\nu}\left(2\alpha + \beta x^{-1}\right)$ and $g(x) = x^{1-\nu}\exp(\beta x)$ for $x \in (0,\infty)$. The pdf of X is (13.10) if and only if the function λ has the form*

$$\lambda(x) = \frac{1}{2\alpha}\exp(\beta x), x > 0.$$

Remark 13.5.4. The general solution of the differential equation (13.15) is

$$\lambda(x) = \exp(\alpha x^2)\left[-\int 2\alpha x^{\nu}\exp(-\alpha x^2 - \beta x)\, g(x)\,\mathrm{d}x + D\right], \quad x > 0,$$

where D is a constant. One set of appropriate functions is given in Proposition 13.5.2.

13.6 Characterizations of the SKS-Type Distribution

In this section we present two characterizations of *pdf* (13.12) in terms of a truncated moment of a function of first order statistic and of a function of nth order statistic, respectively. These characterizations are consequences of the following two theorems given in Hamedani [15], which are stated here for the sake of completeness. We also present a characterization of the *pdf* (13.12) based on the conditional expectation of adjacent generalized order statistics.

Theorem 1 (Theorem 2.2 of [15], p 464). *Let $X : \Omega \to (a,b)$, $a \geq 0$ be a continuous random variable with cdf F such that $\lim_{x\to b} x^{\delta}(1-F(x))^n = 0$, for some $\delta > 0$. Let $g(x,\delta,n)$ be a real-valued function which is differentiable with respect to x and $\int_a^b \frac{\delta x^{\delta-1}}{ng(x,\delta,n)}\mathrm{d}x = \infty$. Then*

$$E\left[X_{1:n}^{\delta} | X_{1:n} > t\right] = t^{\delta} + g(t,\delta,n), \quad a < t < b,$$

implies that

$$F(t) = 1 - \left(\frac{g(a,\delta,n)}{g(t,\delta,n)}\right)^{\frac{1}{n}} \exp\left(-\int_a^t \frac{\delta x^{\delta-1}}{ng(x,\delta,n)}\mathrm{d}x\right), \quad a \leq t < b.$$

Theorem 2 (Theorem 2.8 of [24], p 469). *Let $X : \Omega \to (a,b)$, $a \geq 0$ be a continuous random variable with cdf F such that $\lim_{x\to a}(x-a)^{-\delta}(F(x))^n = 0$, for some $\delta > 0$. Let $g(x,\delta,n)$ be a real-valued function which is differentiable with respect to x and $\int_a^b \frac{\delta(x-a)^{-\delta-1}}{ng(x,\delta,n)}\mathrm{d}x = \infty$. Then*

$$E\left[(X_{1:n}-a)^{-\delta} | X_{n:n} < t\right] = (t-a)^{-\delta} + g(t,\delta,n), \quad a < t < b,$$

implies that

$$F(t) = \left(\frac{g(b,\delta,n)}{g(t,\delta,n)}\right)^{\frac{1}{n}} \exp\left(-\int_t^b \frac{\delta(x-a)^{-\delta-1}}{ng(x,\delta,n)}\mathrm{d}x\right), \quad a \leq t < b.$$

Proposition 13.6.3. *Let* $X : \Omega \to \left(0, \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}\right)$ *be a continuous random variable with cdf* F *such that* $\lim_{x \to \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}} x^{\delta} (1 - F(x))^n = 0$, *for some* $\delta > 0$. *The pdf of* X *is (13.12) if and only if*

$$E\left[X_{1:n}^{\delta} | X_{1:n} > t\right] = t^{\delta} + \frac{\delta}{np}\left(\frac{x^{\delta+p}}{\beta - \alpha x^{2p}}\right), \quad 0 < t < \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}.$$

Proof. See Theorem 1. □

Proposition 13.6.4. *Let* $X : \Omega \to \left(0, \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}\right)$ *be a continuous random variable with cdf* F *such that* $\lim_{x \to 0} x^{-\delta} (F(x))^n = 0$, *for some* $\delta > 0$. *The pdf of* X *is (13.12) if and only if*

$$E\left[X_{n:n}^{-\delta} | X_{n:n} < t\right] = t^{-\delta} - \frac{\delta}{np}\left(\frac{x^{p-\delta}}{\beta - \alpha x^{2p}}\right), \quad 0 < t < \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}.$$

Proof. See Theorem 2. □

The concept of generalized order statistics (*gos*) was introduced by Kamps [19] in terms of their joint *pdf*. The order statistics, record values, k-record values, Pfeifer records, and progressive type II order statistics are special cases of the *gos*. The *rvs* (random variables) $X(1,n,m,k), X(2,n,m,k), \ldots, X(n,n,m,k)$, $k > 0$, and $m \in \mathbb{R}$ are n *gos* from an absolutely continuous *cdf* F with corresponding *pdf* f if their joint *pdf* $f_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n)$ can be written as

$$\begin{aligned} f_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n) &= k\left(\Pi_{j=1}^{n-1} \gamma_j\right)\left[\Pi_{j=1}^{n-1} (1 - F(x_j))^m f(x_j)\right] \\ &\quad \times (1 - F(x_n))^{k-1} f(x_n), F^{-1}(0+) \\ &< x_1 < x_2 < \cdots < x_n < F^{-1}(1-), \end{aligned} \tag{13.16}$$

where $\gamma_j = k + (n - j)(m + 1)$ for all j, $1 \le j \le n$, k is a positive integer, and $m \ge -1$.

If $k = 1$ and $m = 0$, then $X(r,n,m,k)$ reduces to the ordinary rth order statistic and (13.16) will be the joint *pdf* of order statistics $(X_{j:n})_{1 \le j \le n}$ from F. If $k = 1$ and $m = -1$, then (13.16) will be the joint *pdf* of the first n upper record values of the *i.i.d.* (independent and identically distributed) *rvs* with *cdf* F and *pdf* f.

Integrating out $x_1, x_2, \ldots, x_{r-1}, x_{r+1}, \ldots, x_n$ from (13.16), we obtain the *pdf* $f_{r,n,m,k}$ of $X(r,n,m,k)$:

$$f_{r,n,m,k}(x) = \frac{c_r}{\Gamma(r)} (1 - F(x))^{\gamma_r - 1} f(x) g_m^{r-1}(F(x)), \tag{13.17}$$

where $c_r = \Pi_{j=1}^{n-1} \gamma_j$ and

$$g_m(x) = \frac{1}{m+1}\left[1-(1-x)^{m+1}\right], \quad m \neq -1$$
$$= -\ln(1-x), m = -1, x \in (0,1).$$

Since $\lim_{m\to -1} \frac{1}{m+1}\left[1-(1-x)^{m+1}\right] = -\ln(1-x)$, we write $g_m(x) = \frac{1}{m+1}\left[1-(1-x)^{m+1}\right]$, for all $x \in (0,1)$ and all m with $g_{-1}(x) = \lim_{m\to -1} g_m(x)$.

The joint pdf of $X(r,n,m,k)$ and $X(r+1,n,m,k)$, $1 \leq r < n$, is given by (see Kamps [19], p 68)

$$f_{r,,r+1,n,m,k}(x,y) = \frac{c_{r+1}}{\Gamma(r)}(1-F(x))^m f(x) g_m^{r-1}(F(x))(1-F(x))^{\gamma_{r+1}-1} f(y), x < y,$$

and consequently the conditional pdf of $X(r+1,n,m,k)$ given $X(r,n,m,k) = x$, for $m \geq -1$, is

$$f_{r+1|r,n,m,k}(y|x) = \gamma_{r+1}\left(\frac{1-F(y)}{1-F(x)}\right)^{\gamma_{r+1}-1} \cdot \frac{f(y)}{(1-F(x))}, \quad y > x, \tag{13.18}$$

where $\gamma_{r+1} = \gamma_r - 1 - m$. The conditional pdf of $X(r,n,m,k)$ given $X(r+1, n,m,k) = y$, for $m \neq -1$, is

$$f_{r|r+1,n,m,k}(x|y) = r(1-F(x))^m \left(\frac{1-(1-F(x))^{m+1}}{m+1}\right)^{r-1}$$
$$\times \left(\frac{1-(1-F(y))^{m+1}}{m+1}\right)^{-r} f(x), \quad x < y. \tag{13.19}$$

Our last characterization of the pdf (13.12) will be based on the conditional expectation of $X(r,n,m,k)$ given $X(r+1,n,m,k)$ when $m = 0$.

Proposition 13.6.5. *Let $(X_j)_{j\geq 1}$ be a sequence of i.i.d. rvs on $\left(0, \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}\right)$ with an absolutely continuous cdf F, corresponding pdf f and with $\lim_{x\to 0} s(x)(F(x))^r = 0$, where $s(x) = r C_*(\alpha x^{-p} + \beta x^p)$, where C_* is an arbitrary positive constant. Let $(X(r,n,m,k))_{1\leq r\leq n}$ be the first n gos from F. Then*

$$E[s(X(r,n,m,k))|X(r+1,n,m,k) = t] = s(t) + C_*, 0 < t < \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}} \tag{13.20}$$

implies that

$$F(x) = C\exp\left(-\alpha x^{p} - \beta x^{-p}\right), \quad 0 < x < \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}},$$

where $C = \exp\left(2\sqrt{\alpha\beta}\right)$.

Proof. From (13.20), in view of (13.19), we have

$$\int_0^t s(x)\, r(F(x))^{r-1} (F(t))^{-r} f(x)\, dx = s(t) + C_*, \quad 0 < t < \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}.$$

Upon integrating by parts on the left-hand side of the last equality and in view of the assumption $\lim_{x\to 0} s(x)(F(x))^{r} = 0$, we have

$$C_*(F(t))^{r} = -\int_0^t s'(x)(F(x))^{r}\, dx. \tag{13.21}$$

Now, differentiating both sides of (13.21) with respect to t, we arrive at

$$\frac{f(t)}{F(t)} = -\frac{1}{rC_*} s'(t).$$

Integrating both sides of this equality from x to $\left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}$, we have

$$F(x) = \left\{\exp\left(2\sqrt{\alpha\beta}\right)\right\}\exp\left(-\alpha x^{p} - \beta x^{-p}\right), \quad 0 < x < \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2p}}. \qquad \square$$

13.7 Characterizations of the SK Distribution

In this section we present characterizations of the *pdf* (13.14) in terms of a simple relationship between two truncated moments. Our characterization results presented here will, as in Sect. 13.5, employ Theorem G.

Proposition 13.7.1. *Let* $X : \Omega \to (0,\infty)$ *be a continuous random variable and let* $h(x) = x^{p-\nu}$ *for* $x \in (0,\infty)$. *The pdf of X is (13.14), with* $\tau > 1$, *if and only if there exist functions g and* λ *defined in Theorem G, satisfying the differential equation*

$$\frac{\lambda'(x)}{\lambda(x)h(x) - g(x)} = \alpha p(\tau - 1)x^{\nu-1}(\alpha x^{p} + \beta)^{-1}, \quad x > 0. \tag{13.22}$$

Proof. Let X have pdf (13.14) and let

$$g(x) = x^{p-\nu}(\alpha x^p + \beta)^{-1}, \quad x > 0,$$

and

$$\lambda(x) = \frac{\tau}{\tau - 1}(\alpha x^p + \beta), \quad x > 0.$$

Then

$$(1 - F(x))E[h(X)|X \geq x] = \frac{C_3}{\alpha p(\tau - 1)}(\alpha x^p + \beta)^{1-\tau}, \quad x > 0,$$

$$(1 - F(x))E[g(X)|X \geq x] = \frac{C_3}{\alpha p \tau}(\alpha x^p + \beta)^{-\tau}, \quad x > 0,$$

and

$$\lambda(x)h(x) - g(x) = -\frac{1}{\tau}x^{p-\nu} < 0 \quad for \ x > 0.$$

The differential equation (13.22) clearly holds.

Conversely, if g and λ satisfy the differential equation (13.22), then

$$s'(x) = \frac{\lambda'(x)h(x)}{\lambda(x)h(x) - g(x)} = \alpha p(\tau - 1)x^{p-1}(\alpha x^p + \beta)^{-1}, \quad x > 0,$$

and hence

$$s(x) = \ln(\alpha x^p + \beta)^{\tau - 1}, \quad x > 0. \qquad \square$$

Now from Theorem G, X has pdf (13.14).

Corollary 13.7.2. *Let $X : \Omega \to (0, \infty)$ be a continuous random variable and let $h(x) = x^{p-\nu}(\alpha x^p + \beta)^{-1}$ and $g(x) = x^{p-\nu}$ for $x \in (0, \infty)$. The pdf of X is (13.14), with $\tau > 1$, if and only if the function λ has the form*

$$\lambda(x) = \frac{\tau}{\tau - 1}(\alpha x^p + \beta), \quad x > 0.$$

13.8 Conclusion

In designing a stochastic model for a particular modeling problem, an investigator will be vitally interested to know if their model fits the requirements of a specific underlying probability distribution. To this end, the investigator will vitally depend on the characterizations of the selected distribution. A good number of distributions which have important applications in many different fields have been mentioned

in this work. Various characterizations of these distributions have been established. We certainly hope that these results will be of interest to an investigator who may believe their model has a distribution mentioned here and is looking for justifying the validity of their model.

Appendix A

We like to mention that this kind of characterization based on the ratio of truncated moments is stable in the sense of weak convergence; in particular, let us assume that there is a sequence $\{X_n\}$ of random variables with distribution functions $\{F_n\}$ such that the functions g_n, h_n, and λ_n $(n \in \mathbb{N})$ satisfy the conditions of Theorem G and let $g_n \to g$, $h_n \to h$ for some continuously differentiable real functions g and h. Let, finally, X be a random variable with distribution F. Under the condition that $g_n(X)$ and $h_n(X)$ are uniformly integrable and the family is relatively compact, the sequence X_n converges to X in distribution if and only if λ_n converges weakly to λ, where

$$\lambda(x) = \frac{E[g(X)|X \geq x]}{E[h(X)|X \geq x]}.$$

This stability theorem makes sure that the convergence of distribution functions is reflected by corresponding convergence of the functions g , h, and λ, respectively. It guarantees, for instance, the "convergence" of characterization of the Wald distribution to that of the Lévy-Smirnov distribution if $\alpha \to \infty$, as was pointed out in [11].

A further consequence of the stability property of Theorem G is the application of this theorem to special tasks in statistical practice such as the estimation of the parameters of discrete distributions. For such purpose, the functions g, h, and, specially, λ should be as simple as possible. Since the function triplet is not uniquely determined, it is often possible to choose λ as a linear function. Therefore, it is worth analyzing some special cases which helps to find new characterizations reflecting the relationship between individual continuous univariate distributions and appropriate in other areas of statistics.

In view of Theorem G, a characterization of the Pearson system, due to Glänzel [9], is given below.

Proposition A-2. *Let $X : \Omega \to H \subseteq \mathbb{R}$ be a continuous random variable and let $g(x) = x^2 - tx - w$, $h(x) = rx + u$ for $x \in H$, where r, t, u, and w are real parameters such that the distribution is well defined on H. The distribution function of X belongs to Pearson's system if and only if the function λ has the form $\lambda = x$, $x \in H$.*

Remark A-3. Since it can always be assumed that the expectation of a non-strictly positive continuous random variable is zero, we let $u = 0$, where appropriate, in the brief discussion below. Note that $w > 0$ if $u = 0$.

The following cases can be distinguished:

Type I. $r \in (0,1), t \neq 0$. (This is the family of finite beta distribution.)
Type II. $r \in (0,1), t = 0$. (This is a symmetric beta distribution.)
Type III. $r = 1, t \neq 0$. (This is the family of gamma distribution.)
$r = 1, t = 0$. (This is the normal distribution.)
Type IV. $r \in \left(1 + \frac{t^2}{4w}, \infty\right), t \neq 0$.
Type V. $r = 1 + \frac{t^2}{4w}, t \neq 0$. (This is the family of inverse Gaussian distribution.)
Type VI. $r \in \left(1, 1 + \frac{t^2}{4w}\right), t \neq 0$. (This is the family of infinite beta distribution.)
Type VII. $r \in \left(1 + \frac{t^2}{4w}, \infty\right), t = 0$.

The following proposition is given in Glänzel and Hamedani [11]

Proposition A-4. *Let $X : \Omega \to H \subseteq \mathbb{R}$ be a continuous random variable and let $g(x) = \frac{\{(a_0+1)x^2+(a_1+c)x+a_2\}}{\{a_0x^2+a_1x+a_2\}}$, $h(x) = \frac{\{x+c\}}{\{a_0x^2+a_1x+a_2\}}$ for $x \in H$, where $c > 0$, a_0, a_1, and a_2 are real parameters such that the distribution function is well defined on H. The distribution function of X belongs to Pearson's system if and only if the function λ has the form $\lambda = x$, $x \in H$.*

The families of Pearson's system can be obtained from special choices of the parameters c, a_0, a_1, and a_2(see, e.g., [18]).

References

1. Ahsanullah M (1973) A characterization of the power function distribution. Commun Stat 2:259–262
2. Ahsanullah M (2004) On characterizations of the uniform distribution based on functions of order statistics. Commun Stat Theor Methods 33 :2921–2928
3. Ahsanullah M, Hamedani GG (2007) Certain characterizations of the power function and beta distribution based on order statistics. J Stat Theor Appl 6 :220–225
4. Asadi M, Ebrahimi N, Hamedani GG, Soofi E (2004) Maximum dynamic entropy models. J Appl Probab 41:379–390
5. Asadi M, Ebrahimi N, Hamedani GG, Soofi E (2005) Minimum dynamic discrimination information models. J Appl Probab 42:643–660
6. Burr IW (1942) Cumulative frequency functions. Ann Math Stat 13 :215–232
7. Crooks GE (2010) The Amoroso distribution. arXiv:1005.3274v1 [math.ST]
8. Galambos J, Kotz S (1978) Characterizations of probability distributions. In: A unified approach with an emphasis on exponential and related models, Lecture Notes in Mathematics, vol 675. Springer, Berlin
9. Glänzel W (1987) A characterization theorem based on truncated moments and its application to some distribution families. Math Stat Probab Theor (bad Tatzmannsdorf, 1986):75–84
10. Glänzel W (1988) A characterization of the normal distribution. Studia Sci Math Hungar 23:89–91
11. Glänzel W, Hamedani GG (2001) Characterizations of univariate continuous distributions. Studia Sci Math Hungar 37:83–118

12. Glänzel W, Telcs A, Schubert A (1984) Characterization by truncated moments and its application to Pearson-type distributions. Z Wahrsch Verw Gebiete 66:173–183
13. Hamedani GG (2002) Characterizations of univariate continuous distributions II. Studia Sci Math Hungar 39:407–424
14. Hamedani GG (2006) Characterizations of univariate continuous distributions III. Studia Sci Math Hungar 43:361–385
15. Hamedani GG (2010) Characterizations of continuous univariate distributions based on the truncated moments of functions of order statistics. Studia Sci Math Hungar 47:462–484
16. Hamedani GG (2011) Characterizations of the Shakil-Kibria-Singh distribution 40:201–207
17. Hamedani GG, Ahsanullah M, Sheng R (2008) Characterizations of continuous univariate distributions based on truncated moment of first order statistic. Aligarh J Stat 28:75–81
18. Johnson NI, Kotz S (1970) Distributions in statistics. Continuous univariate distributions, vols 1 and 2. Houghton Miffin Co., Boston, Mass
19. Kamps U (1995) A concept of generalized order statistics. In: Teubner BG, Stuttgart Teubner Skripten zur Mathematischen Stochastik. [Teubner Texts on Mathematical Stochastics]
20. Kotz S, Shanbhag DN (1980) Some new approaches to probability distributions. Adv Appl Probab 12:903–921
21. McDonald JB (1984) Some generalized functions for the size distribution of income. Econometrica 52:647–663
22. Pearson K, (1895) Contributions to the mathematical theory of evolution. Skew variation in homogeneous material. London Phil Trans Ser A 186:393–415; Lond R S Proc 57:257–260
23. Shakil M, Kibria BM (2010) On a family of life distributions based on generalized Pearson differential equation with applications in health statistics. J Stat Theor Appl 9:255–281
24. Shakil M, Kibria BM, Singh JN (2010) A new family of distributions based on the generalized Pearson differential equation with some applications. Austrian J Stat 39:259–278
25. Shakil M, Singh JN, Kibria BM (2010) On a family of product distributions based on the Wittaker functions and generalized Pearson differential equation. Pakistan J Stat 26:111–125

Chapter 14
Bayesian Wavelet Shrinkage Strategies: A Review

Norbert Reményi and Brani Vidakovic

Abstract In this chapter the authors overview recent developments and current status of use of Bayesian paradigm in wavelet shrinkage. The paradigmatic problem where wavelet shrinkage is employed is that of nonparametric regression where data are modeled as observations from an unknown signal contaminated with a Gaussian noise. Bayes rules as general shrinkers provide a formal mechanism to implement shrinkage in the wavelet domain that is model based and adaptive. New developments including dependence models, complex wavelets and MCMC strategies are described. Applications include inductance plethysmography data and curve classification procedure applied in botany. The chapter features an extensive set of references consisting of almost 100 entries.

14.1 Introduction

Wavelet-based tools became standard methodology in many areas of modern statistics, for example, in regression, density and function estimation, factor analysis, modeling and forecasting of time series, functional data analysis, and data mining and classification, with ranges of application areas in science and engineering. Wavelets owe their initial popularity in statistics to shrinkage, a simple and yet powerful procedure in nonparametric statistical modeling. Wavelet shrinkage is a

N. Reményi (✉)
School of Industrial and Systems Engineering, Georgia Institute of Technology, 765 Ferst Drive, NW Atlanta, GA 30332-0205, USA
e-mail: nremenyi@gatech.edu

B. Vidakovic
School of Biomedical Engineering, Georgia Institute of Technology, 765 Ferst Drive, NW Atlanta, GA 30332-0205, USA
e-mail: brani@gatech.edu

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_14,

three-step procedure: (1) data are transformed into a set of wavelet coefficients; (2) a shrinkage of the coefficients is performed; and (3) the processed wavelet coefficients are transformed back to the domain of the original data.

Wavelet domains are desirable modeling environments; several supporting arguments are listed below.

Discrete wavelet transformations tend to "disbalance" the data. Even though the orthogonal transforms preserve the ℓ_2 norm of the data (the square root of sum of squares of observations, or the "energy" as engineers like to say), most of the ℓ_2 norm in the transformed data is concentrated in only a few wavelet coefficients. This concentration narrows the class of plausible models and facilitates the thresholding. The disbalancing property also yields a variety of criteria for the selection of best basis.

Wavelets, as modeling building blocks, are well localized in both time and scale (frequency). Signals with rapid local changes (signals with discontinuities, cusps, sharp spikes, etc.) can be represented with only a few wavelet coefficients. This parsimony does not, in general, hold for other standard orthonormal bases which may require many "compensating" coefficients to describe discontinuity artifacts or local bursts.

Heisenberg's principle states that time-frequency models cannot be arbitrarily precise in the time and frequency domains simultaneously, rather this precision is bounded from the below by a universal constant. Wavelets adaptively distribute the time-frequency precision by their innate nature. The economy of wavelet transforms can be attributed to their ability to confront the limitations of Heisenberg's principle in a data-dependent manner.

An important feature of wavelet transforms is their whitening property. There is ample theoretical and empirical evidence that wavelet transforms simplify the dependence structure in the original data. For example, it is possible, for any given stationary dependence in the input signal, to construct a biorthogonal wavelet basis such that the corresponding in the transform are uncorrelated (a wavelet counterpart of Karhunen–Loève transform). For a discussion and examples see [91].

We conclude this incomplete list of features of wavelet transforms by pointing out their sensitivity to self-similar data. The scaling laws are distinctive features of self-similar data. Such laws are clearly visible in the wavelet domain in the so-called wavelet spectra, wavelet counterparts of the Fourier spectra.

More arguments can be given: computational speed of the wavelet transformation, easy incorporation of prior information about some features of the signal (smoothness, distribution of energy across scales), etc.

Prior to describing a formal setup for Bayesian wavelet shrinkage, we provide a brief review of discrete wavelet transforms and traditional wavelet shrinkage.

Basics on wavelets can be found in many texts, monographs, and papers at many different levels of exposition. The interested reader should consult monographs by [33, 68, 87, 91], among others. An introductory article is [88].

14.1.1 Discrete Wavelet Transformations and Wavelet Shrinkage

Let $\boldsymbol{y}$ be a data vector of dimension (size) n. For the simplicity we choose n to be a power of 2, say 2^J. We assume that measurements $\boldsymbol{y}$ belong to an interval and consider periodized wavelet bases. Generalizations to different sample sizes and general wavelet and wavelet-like transformations are straightforward.

Suppose that the vector $\boldsymbol{y}$ is wavelet transformed to a vector $\boldsymbol{d}$. This linear and orthogonal transform can be fully described by an $n \times n$ orthogonal matrix $\boldsymbol{W}$. The use of the matrix $\boldsymbol{W}$ is possible when n is not large (of order of a few thousand, at most), but for large n, fast filtering algorithms are employed. The filtering procedures are based on so-called quadrature mirror filters which are uniquely determined by the choice of wavelet and fast Mallat's algorithm [63]. The wavelet decomposition of the vector $\boldsymbol{y}$ can be written as

$$\boldsymbol{d} = (H^{\ell}\boldsymbol{y}, GH^{\ell-1}\boldsymbol{y}, \ldots, GH^2\boldsymbol{y}, GH\boldsymbol{y}, G\boldsymbol{y}). \tag{14.1}$$

Note that in (14.1), $\boldsymbol{d}$ has the same length as $\boldsymbol{y}$ and ℓ is any fixed number between 1 and $J = \log_2 n$. The operators G and H acting on data sequences are defined coordinate-wise via

$$(Ha)_k = \Sigma_{m\in\boldsymbol{Z}} h_{m-2k} a_m, \text{ and } (Ga)_k = \Sigma_{m\in\boldsymbol{Z}} g_{m-2k} a_m, \ k \in \boldsymbol{Z},$$

where g and h are high- and low-pass wavelet filters. Components of g and h are connected via the *quadrature mirror* relationship, $g_n = (-1)^n h_{1-n}$. For all commonly used wavelet bases, the taps of filters g and h are readily available in the literature or in standard software packages.

The elements of $\boldsymbol{d}$ are called "wavelet coefficients." The subvectors described in (14.1) correspond to detail levels. For instance, the vector $G\boldsymbol{y}$ contains $n/2 = 2^{J-1}$ coefficients representing the level of the finest detail. When $\ell = J$, the vectors $GH^{J-1}\boldsymbol{y} = \{d_{00}\}$ and $H^J\boldsymbol{y} = \{c_{00}\}$ contain a single coefficient each and represent the coarsest possible level of detail and the smooth part in wavelet decomposition, respectively.

In general, jth detail level in the wavelet decomposition (14.1) contains 2^j elements, and can be written as

$$GH^{J-j-1}\boldsymbol{y} = (d_{j,0}, d_{j,1}, \ldots, d_{j,2^j-1}). \tag{14.2}$$

Wavelet shrinkage methodology consists of shrinking the magnitudes of wavelet coefficients. The simplest wavelet shrinkage technique is thresholding. The components of $\boldsymbol{d}$ are replaced by 0 if their absolute value does not exceed a fixed threshold λ.

The two most common thresholding policies are *hard* and *soft* thresholding with corresponding rules given by:

$$\theta^h(d,\lambda) = d\,\mathbf{1}(|d| > \lambda),$$
$$\theta^s(d,\lambda) = (d - \text{sign}(d)\lambda)\,\mathbf{1}(|d| > \lambda),$$

where $\mathbf{1}(A)$ is the indicator of relation A, i.e., $\mathbf{1}(A) = 1$ if A is true and $\mathbf{1}(A) = 0$ if A is false.

In the next section we describe how the Bayes rules, resulting from the models on wavelet coefficient, can act as shrinkage/thresholding rules.

14.2 Wavelets and Bayes

Bayesian paradigm has become very popular in wavelet data processing since Bayes rules are shrinkers. This is true in general, although examples of Bayes rules that expand can be found, see [89]. The Bayes rules can be constructed to mimic the thresholding rules: to slightly shrink the large coefficients and heavily shrink the small coefficients. In addition, Bayes rules result from realistic statistical models on wavelet coefficients and such models allow for incorporation of prior information about the *true* signal. Furthermore, most Bayes rules can be easily either computed by simulation or expressed in a closed form. Reviews of early Bayesian approaches can be found in [3, 78, 86, 87]. An edited volume on Bayesian modeling in the wavelet domain appeared 12 years ago [65].

A paradigmatic task in which the wavelets are typically applied is recovery of an unknown signal $\boldsymbol{f}$ observed with noise $\boldsymbol{e}$. In statistical terms this would be a task of nonparametric regression. Wavelet transformations $\boldsymbol{W}$ are applied to noisy measurements $y_i = f_i + e_i,\ i = 1,\ldots,n$, or, in vector notation, $\boldsymbol{y} = \boldsymbol{f} + \boldsymbol{e}$. The linearity of $\boldsymbol{W}$ implies that the transformed vector $\boldsymbol{d} = \boldsymbol{W}(\boldsymbol{y})$ is the sum of the transformed signal $\boldsymbol{\theta} = \boldsymbol{W}(\boldsymbol{f})$ and the transformed noise $\boldsymbol{\varepsilon} = \boldsymbol{W}(\boldsymbol{e})$. Furthermore, the orthogonality of $\boldsymbol{W}$ and Gaussianity of $\boldsymbol{e}$ implies Gaussianity of $\boldsymbol{\varepsilon}$ as well.

Bayesian methods are applied in the wavelet domain, that is, after the data have been transformed. The wavelet coefficients can be modeled in totality, as a single vector, or one by one, due to decorrelating property of wavelet transforms. Block-modeling approaches are also possible.

When the model is on individual wavelet (detail) coefficients $d_i \sim N(\theta_i, \sigma^2),\ i = 1,\ldots,n$, the interest relies in the estimation of the θ_i. Usually we concentrate on typical wavelet coefficient and model: $d = \theta + \varepsilon$. Bayesian methods are applied to estimate the location parameter θ, which will be, in the sequel, argument in the inverse wavelet transform. A prior on θ, and possibly on other parameters of the distribution of ε, is elicited, and the corresponding Bayes estimators are back-transformed. Various choices of Bayesian models have been motivated by different,

often contrasting, interests. Some models were driven by empirical justifications, others by pure mathematical considerations; some models lead to simple closed-form rules, the other require extensive Markov Chain Monte Carlo (MCMC) simulations to produce the estimate. Bayes rules with respect to absolute or 0-1 loss functions are capable of producing bona fide thresholding rules.

14.2.1 An Illustrative Example

As an illustration of the Bayesian approach we present BAMS (Bayesian adaptive multiresolution shrinkage). The method, due to [90], is motivated by empirical considerations on the coefficients and leads to easily implementable Bayes estimates, available in closed form.

The BAMS originates from the observation that a realistic Bayes model should produce prior predictive distributions of the observations which "agree" with the observations. Other authors were previously interested in the empirical distribution of the wavelet coefficients, see, for example, [57, 58, 63, 77, 81, 86]. Their common argument can be summarized by the following statement:

> For most of the signals and images encountered in practice, the empirical distribution of a typical detail wavelet coefficient is notably centered about zero and peaked at it.

In accordance with the spirit of this statement, [63] suggested to fit empirical distributions of wavelet coefficients by the exponential power model

$$f(d) = C \cdot \mathrm{e}^{-(|d|/\alpha)^{\beta}},\ \alpha, \beta > 0,$$

where $C = \frac{\beta}{2\alpha\Gamma(1/\beta)}$.

Following the Bayesian paradigm, prior distributions should be elicited on the parameters of the model $d|\theta, \sigma^2 \sim N(\theta, \sigma^2)$ and Bayesian estimators (namely, posterior means under squared loss) computed. In BAMS, priors on θ and σ^2 are set such that the marginal (prior predictive) distribution of the wavelet coefficients is a double exponential distribution DE, that is, an exponential power one with $\beta = 1$. The double exponential distribution can be obtained by marginalizing the normal likelihood by adopting exponential prior on its variance σ^2. The choice of an exponential prior can be justified by its *maxent* property, that is, exponential distribution is the entropy maximizer in the class of all distributions supported on $(0, \infty)$ with a fixed first moment, and in that sense is noninformative.

Thus, BAMS uses the exponential prior $\sigma^2 \sim E(\mu), \quad \mu > 0$, which leads to the marginal likelihood

$$d|\theta \sim DE\left(\theta, \frac{1}{\sqrt{2\mu}}\right), \quad \text{with density } f(d|\theta) = \frac{1}{2}\sqrt{2\mu}\mathrm{e}^{-\sqrt{2\mu}|d-\theta|}.$$

Vidakovic [86] considered the previous marginal likelihood but with a t distribution as the prior on θ. The Bayes rules with respect to the squared error loss under general but symmetric priors $\pi(\theta)$ can be expressed using the Laplace transforms of $\pi(\theta)$.

In personal communication with the second author, Jim Berger and Peter Müller suggested in 1993 the use of ε-contamination priors in the wavelet context pointing out that such priors would lead to rules which are smooth approximations to a thresholding.

The choice

$$\pi(\theta) = \varepsilon\delta(0) + (1-\varepsilon)\xi(\theta) \tag{14.3}$$

also reflects prior belief that some locations (corresponding to the signal or function to be estimated) are 0 and that there is a nonzero spread component ξ describing "large" locations. In addition to this prior sparsity of the signal part, this prior leads to desirable shapes of the resulting Bayes rules. Note that here $0 \le \varepsilon \le 1$ denotes the mixing weight, not the random error component, and will be used throughout this chapter in contamination priors.

In BAMS, the spread part ξ is chosen as $\theta \sim DE(0,\tau)$. The Bayes rule under the squared error loss is

$$\delta_\pi(d) = \frac{(1-\varepsilon)\, m_\xi(d)\, \delta_\xi(d)}{(1-\varepsilon)\, m_\xi(d) + \varepsilon\, DE\left(0, \frac{1}{\sqrt{2\mu}}\right)}, \tag{14.4}$$

where

$$m_\xi(d) = \frac{\tau\mathrm{e}^{-|d|/\tau} - \frac{1}{\sqrt{2\mu}}\mathrm{e}^{-\sqrt{2\mu}|d|}}{2\tau^2 - 1/\mu}$$

and

$$\delta_\xi(d) = \frac{\tau(\tau^2 - 1/(2\mu))d\mathrm{e}^{-|d|/\tau} + \tau^2(\mathrm{e}^{-|d|\sqrt{2\mu}} - \mathrm{e}^{-|d|/\tau})/\mu}{(\tau^2 - 1/(2\mu))(\tau\mathrm{e}^{-|d|/\tau} - (1/\sqrt{2\mu})\mathrm{e}^{-|d|\sqrt{2\mu}})}$$

are the prior predictive distribution and the Bayes rule for the spread part of the prior, ξ. Rule (14.4) is the BAMS rule, which falls between comparable hard and soft thresholding rules.

Bayes rules under the squared error loss and regular models are never thresholding rules. To extend this motivating example, we consider the posterior median as an estimator for θ. It is well known that under the absolute error loss $L(\theta,d) = |\theta - d|$, the posterior risk is minimized by the posterior median. The posterior median was first considered by Abramovich et al. [7] in the context of wavelet shrinkage. It could be a thresholding rule, which is preferable to smooth shrinkage rules in many applications, like model selection, data compression, dimension reduction, and related statistical tasks in which it is desirable to replace by zero a majority of the processed coefficients.

For the model above the posterior distribution is $\pi^\star(\theta|d) = f(d|\theta)\pi(\theta)/m_\pi(d)$, where

$$m_\pi(d) = (1-\varepsilon)\, m_\xi(d) + \varepsilon\, DE\left(0, \frac{1}{\sqrt{2\mu}}\right).$$

In order to find the median of the posterior distribution, the solution of the following equation, with respect to u, is needed:

$$\int_{-\infty}^{u} \pi^\star(\theta|d)\mathrm{d}\theta = \frac{1}{2}. \tag{14.5}$$

It is easy to show with simple calculus that if $d \geq 0$,

$$\max \int_{-\infty}^{0^-} \pi^\star(\theta|d)\mathrm{d}\theta = \frac{1}{2}, \tag{14.6}$$

and in case $d < 0$,

$$\min \int_{-\infty}^{0} \pi^\star(\theta|d)\mathrm{d}\theta = \frac{1}{2}. \tag{14.7}$$

Because $\pi^\star(\theta|d)$ is a probability density, the integral in (14.5) is non-decreasing in u. Therefore, by using results (14.6) and (14.7), the posterior median is always greater than equal to zero, when $d \geq 0$, and less than equal to zero, when $d < 0$.

To find the posterior median, first consider the case $d \geq 0$. We know that the solution u satisfies $u \geq 0$. The equation in (14.5) becomes

$$\frac{\varepsilon \frac{\sqrt{2\mu}}{2} \mathrm{e}^{-\sqrt{2\mu}d} + (1-\varepsilon)\frac{\sqrt{2\mu}}{4\tau}\mathrm{e}^{-\sqrt{2\mu}d}\left\{\frac{1}{\sqrt{2\mu}+1/\tau} + \frac{1}{\sqrt{2\mu}-1/\tau}\left[\mathrm{e}^{(\sqrt{2\mu}-1/\tau)u} - 1\right]\right\}}{m_\pi(d)} = \frac{1}{2}.$$

Next, assume $d < 0$. Then the solution satisfies $u \leq 0$ and (14.5) becomes:

$$\frac{(1-\varepsilon)\frac{\sqrt{2\mu}}{4\tau}\left\{\frac{1}{\sqrt{2\mu}+1/\tau}\mathrm{e}^{d/\tau} + \frac{1}{\sqrt{2\mu}-1/\tau}\mathrm{e}^{d/\tau} - \frac{1}{\sqrt{2\mu}-1/\tau}\mathrm{e}^{-(\sqrt{2\mu}-1/\tau)u}\right\}}{m_\pi(x)} = \frac{1}{2}.$$

From the above, the algorithm for finding the posterior median $\delta_M(d)$ is:

For $d > 0$,

$$\text{if } \frac{\varepsilon \frac{\sqrt{2\mu}}{2} e^{-\sqrt{2\mu}d} + (1-\varepsilon)\frac{\sqrt{2\mu}}{4\tau} e^{-\sqrt{2\mu}d} \frac{1}{\sqrt{2\mu}+1/\tau}}{m_\pi(d)} > \frac{1}{2}, \quad \delta_M(d) = 0$$

$$\text{else } \delta_M(d) = \frac{1}{\sqrt{2\mu}-1/\tau} \log\left\{ \left[\frac{m_\pi(d)/2 - \varepsilon \frac{\sqrt{2\mu}}{2} e^{-\sqrt{2\mu}d}}{(1-\varepsilon)\frac{\sqrt{2\mu}}{4\tau} e^{-\sqrt{2\mu}d}} + \frac{2/\tau}{2\mu - 1/\tau^2} \right] (\sqrt{2\mu} - 1/\tau) \right\}.$$

For $d < 0$,

$$\text{if } \frac{(1-\varepsilon)\frac{\sqrt{2\mu}}{4\tau}\left[\frac{1}{\sqrt{2\mu}+1/\tau} e^{d/\tau} + \frac{1}{\sqrt{2\mu}-1/\tau} e^{d/\tau} - \frac{1}{\sqrt{2\mu}-1/\tau} e^{(\sqrt{2\mu}-1/\tau)d}\right]}{m_\pi(d)} < \frac{1}{2},$$

$$\delta_M(d) = 0$$

$$\text{else } \delta_M(d) = -\frac{1}{\sqrt{2\mu}-1/\tau} \log\left\{ -\left[\frac{\frac{m_\pi(d)/2}{(1-\varepsilon)\frac{\sqrt{2\mu}}{4\tau}} - \frac{1}{\sqrt{2\mu}+1/\tau} e^{d/\tau}}{\frac{1}{\sqrt{2\mu}-1/\tau} e^{\sqrt{2\mu}d}} - e^{-(\sqrt{2\mu}-1/\tau)d} \right] \right\}.$$

For $d = 0$,

$$\delta_M(d) = 0. \tag{14.8}$$

The rule $\delta_M(d)$ based on algorithm (14.8) is the BAMS-MED rule. As evident from Fig. 14.1, the BAMS-MED rule is a thresholding rule.

14.3 Bayesian Wavelet Regression

14.3.1 Term-by-Term Shrinkage

As we indicated in the introduction, the most popular application of wavelets is the nonparametric regression problem

$$y_i = f(x_i) + e_i, \quad i = 1, \ldots, n.$$

The usual assumptions are that x_i, $i = 1, \ldots, n$ are equispaced (e.g., time points), and the random errors e_i are i.i.d. normal, with zero mean and variance σ^2. The interest

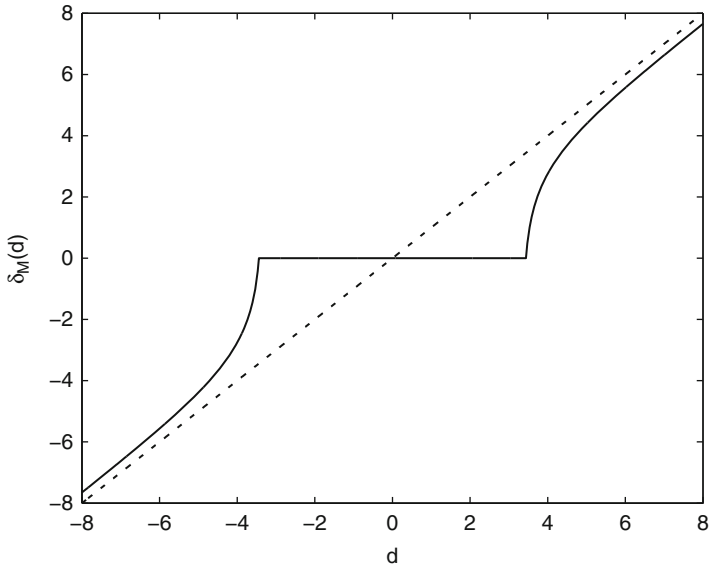

Fig. 14.1 BAMS-MED rule (14.8) for $\varepsilon = 0.9$, $\mu = 1$, and $\tau = 2$

is to estimate the function f using the observations y. After applying a linear and orthogonal wavelet transform, the problem becomes

$$d_{jk} = \theta_{jk} + \varepsilon_{jk},$$

where d_{jk}, θ_{jk}, and ε_{jk} are the wavelet coefficients (at resolution j and position k) corresponding to y, f, and e, respectively.

Due to the whitening property of wavelet transforms [39], many existing methods assume independence of the wavelet coefficients and model the wavelet coefficients one by one using notation for a generic wavelet coefficient, $d = \theta + \varepsilon$. Shrinkage is performed term by term, which is sometimes referred to as diagonal shrinkage.

An early example of the diagonal Bayesian approach to wavelet regression is the adaptive Bayesian wavelet shrinkage (ABWS) proposed by Chipman et al. [27]. Their approach is based on the stochastic search variable selection (SSVS) proposed by George and McCulloch [41], with the assumption that σ is known.

Chipman et al. [27] start with the model

$$d|\theta, \sigma^2 \sim N(\theta, \sigma^2).$$

The prior on θ is defined as a mixture of two normals

$$\theta|\gamma_j \sim \gamma_j N(0, (c_j \tau_j)^2) + (1 - \gamma_j) N(0, \tau_j^2),$$

where

$$\gamma_j \sim \text{Ber}(p_j).$$

Because the hyperparameters p_j, c_j, and τ_j depend on the level j to which the corresponding θ (or d) belongs, and can be level-wise different, the method is adaptive.

The Bayes rule under squared error loss for θ (from the level j) has an explicit form,

$$\delta(d) = \left[P(\gamma_j = 1|d) \frac{(c_j\tau_j)^2}{\sigma^2 + (c_j\tau_j)^2} + P(\gamma_j = 0|d) \frac{\tau_j^2}{\sigma^2 + \tau_j^2} \right] d, \quad (14.9)$$

where

$$P(\gamma_j = 1|d) = \frac{p_j \pi(d|\gamma_j = 1)}{(1-p_j)\pi(d|\gamma_j = 0)}$$

and

$$\pi(d|\gamma_j = 1) \sim N(0, \sigma^2 + (c_j\tau_j)^2) \text{ and } \pi(d|\gamma_j = 0) \sim N(0, \sigma^2 + \tau_j^2).$$

For other early examples of the Bayesian approach to wavelet regression see papers, for example, by Abramovich et al. [7, 28, 31, 85].

A more recent paper by Johnstone and Silverman [51] presents a class of empirical Bayes methods for wavelet shrinkage. The hyperparameters of the model are estimated by marginal maximum likelihood; therefore, the threshold is estimated from the data. The authors consider different level-dependent priors, all of which are a mixture of point mass at zero and a heavy-tailed density. One of the choices for the heavy-tailed density is the double exponential (Laplace) prior, for which we present the posterior mean to exemplify their methodology.

At level j of the wavelet decomposition, define the sequence $z_k = d_{jk}/\sigma_j$, where σ_j is the standard deviation of the noise at level j, which is estimated from the data. Therefore, $z_k = \mu_k + \varepsilon_k$, where the ε_k are i.i.d. $N(0,1)$ random variables. The authors model parameters μ_k with independent mixture prior distributions

$$\pi(\mu) = (1-w)\delta_0(\mu) + w\gamma(\mu),$$

where $\delta_0(\mu)$ denotes a point mass at zero. Using the double exponential distribution $\gamma_a(\mu) = \frac{1}{2}\exp\{-a|\mu|\}$, with scale parameter $a > 0$, the marginal distribution of z becomes

$$m(z) = (1-w)\varphi(z) + wg(z),$$

where φ denotes the standard normal density and

$$g(z) = \tfrac{1}{2} a \exp\left\{\tfrac{1}{2}a^2\right\}\left[\mathrm{e}^{-az}\Phi(z-a) + \mathrm{e}^{az}\tilde{\Phi}(z+a)\right].$$

In the above equation Φ denotes the cumulative distribution of the standard normal and $\tilde{\Phi} = 1 - \Phi$. The posterior distribution of μ becomes

$$\pi^{\star}(\mu|z) = (1 - w_{\text{post}})\delta_0(\mu) + w_{\text{post}} f_1(\mu|z),$$

where the posterior probability w_{post} is

$$w_{\text{post}}(z) = w g(z) \big/ \left[w g(z) + (1-w)\varphi(z)\right]$$

and

$$f_1(\mu|z) = \begin{cases} \mathrm{e}^{az}\varphi(\mu - z - a) \big/ \left[\mathrm{e}^{-az}\Phi(z-a) + \mathrm{e}^{az}\tilde{\Phi}(z+a)\right], & \mu \le 0 \\ \mathrm{e}^{-az}\varphi(\mu - z + a) \big/ \left[\mathrm{e}^{-az}\Phi(z-a) + \mathrm{e}^{az}\tilde{\Phi}(z+a)\right], & \mu > 0, \end{cases}$$

which is a weighted sum of truncated normal distributions. Detailed derivations of $g(z)$ and $f_1(\mu|z)$ are provided by Pericchi and Smith [72]. It can be shown that the posterior mean is

$$\mathbb{E}(\mu|z) = w_{\text{post}}(z)\left[z - \frac{a\left[\mathrm{e}^{-az}\Phi(z-a) - \mathrm{e}^{az}\tilde{\Phi}(z+a)\right]}{\mathrm{e}^{-az}\Phi(z-a) + \mathrm{e}^{az}\tilde{\Phi}(z+a)}\right]. \qquad (14.10)$$

A schematic picture of the posterior mean (14.10) is presented in Fig. 14.2 for $w = 0.1$ and $a = 0.5$. It exhibits a desirable shrinkage pattern slightly shrinking large and heavily shrinking small coefficients in magnitude.

The mixing weight w and scale parameter a are estimated by marginal maximum likelihood for each dyadic level j. The authors also provide the posterior median for the above model and closed-form equations for the posterior mean and median in case $\gamma(\mu)$ is a quasi-Cauchy distribution. For more details and related theoretical results the reader is referred to [51], and for more examples using the method, see [52].

Several more recent papers have considered term-by-term Bayesian wavelet shrinkage. Angelini and Sapatinas [10] consider an empirical Bayes approach to wavelet regression by eliciting the ε-contamination class of prior distributions and using type II maximum likelihood approach to prior selection. Angelini and Vidakovic [11] show that Γ-minimax shrinkage rules are Bayes with respect to a least favorable contamination prior with a uniform spread distribution. Their method allows for incorporation of information about the energy in the signal of interest. Cutillo et al. [32] consider thresholding rules induced by a variation of the Bayesian

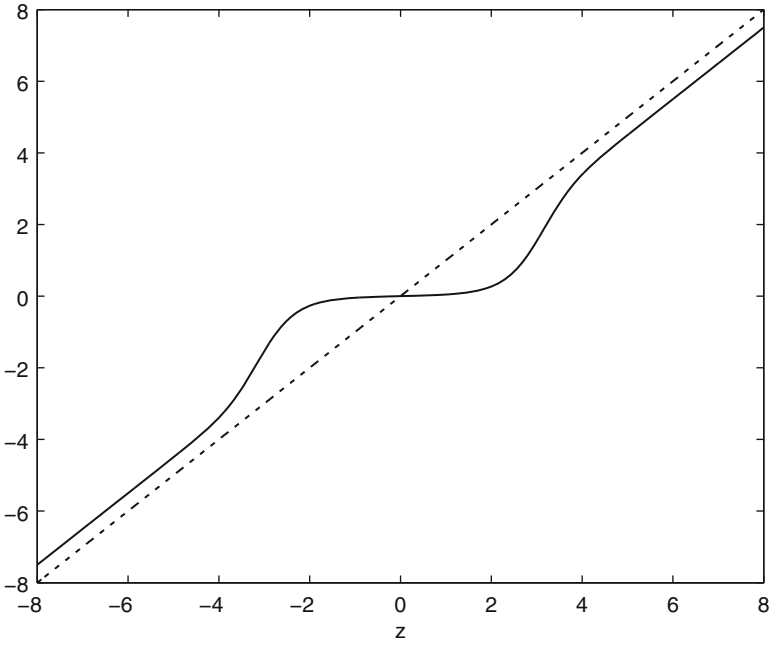

Fig. 14.2 Posterior mean rule (14.10) for $w = 0.1$ and $a = 0.5$

MAP principle in a properly set Bayesian model. The rule proposed is called larger posterior mode (LPM) because it always picks the mode of the posterior larger in absolute value. Ter Braak (2006) extends the normal Bayesian linear model by specifying a flat prior on the δth power of the variance components of the regression coefficients. In the orthonormal case, easy-to-compute analytic expressions are derived, and the procedure is applied in a simulation study of wavelet denoising.

14.3.2 Bayesian Block Shrinkage

Methods considered above are called diagonal, since the wavelet coefficients are assumed independent. In reality the wavelet coefficients are dependent, but this dependence is weak and decreases with increasing the separation distance between them and the number of vanishing moments of the decomposing wavelet. Many authors argued that shrinkage performance can be improved by considering the neighborhoods of wavelet coefficients (blocks, parent-child relations, cones of influence, etc.) and report improvements over the diagonal methods. Examples include classical block thresholding methods by Hall et al. [19–21, 44–46] where wavelet coefficients are thresholded based on block sums of squares.

Abramovich et al. [4] considered an empirical Bayes approach to incorporating information on neighboring wavelet coefficients into function estimation. The authors group wavelet coefficients d_{jk} into m_j nonoverlapping blocks b_{jK} $(K = 1,\ldots,m_j)$ of length l_j at each resolution level j. The block of observed wavelet coefficients will be denoted as $\hat{b}_{jK}$. They consider the following prior model for blocks b_{jK}:

$$b_{jK}|\gamma_{jK} \sim N(0, \gamma_{jK} V_j),$$
$$\gamma_{jK} \sim \text{Ber}(\pi_j).$$

Independence of blocks across different resolution levels is assumed. This prior model allows for a covariance structure between neighboring coefficients in the same block, supporting the fact that wavelet coefficients are more likely to contain signal if this is true for their neighbors as well. The covariance matrix V_j is specified at each level j by two hyperparameters τ_j and ρ_j, where the correlation between the coefficients, ρ_j, decreases as the distance between the coefficients increases. Combining the prior model with the likelihood $\hat{b}_{jK} \sim N(b_{jK}, \sigma^2 I)$ leads to the posterior mean of b_{jK} as

$$\mathbb{E}(b_{jK}|\hat{b}_{jK}) = \frac{1}{1+O_{jK}} A_j \hat{b}_{jK}, \tag{14.11}$$

where

$$O_{jK} = \frac{1-\pi_j}{\pi_j} \left(\frac{\det(V_j)}{\sigma^{2l_j} \det(A_j)} \right)^{1/2} \exp\left\{ -\frac{\hat{b}'_{jK} A_j \hat{b}_{jK}}{2\sigma^2} \right\},$$
$$A_j = (\sigma^2 V_j^{-1} + I)^{-1}.$$

Rule (14.11) is a nonlinear block shrinkage rule, by which the observed wavelet coefficients in block jK are shrunk by the same factor determined by all the coefficients within the block. The authors also provide details for the posterior median and the Bayes factor procedure, which are individual and block thresholding rules, respectively.

Hyperparameters π_j, τ_j, and ρ_j are estimated by marginal maximum likelihood method for each level j, and hyperparameter σ is estimated by the standard median absolute deviation suggested by Donoho and Johnstone [35]. After plugging in the estimate $\hat{\sigma}$ and some reparametrization, the negative log-likelihood function $-l_j(\pi_j, \tau_j, \rho_j, \hat{\sigma})$ was minimized by the Nelder–Mead simplex search method.

The authors present detailed simulation study of the method and an application to inductance plethysmography data. For details the reader is referred to [4].

A paper by De Canditiis and Vidakovic [34] proposed the BBS (Bayesian block shrinkage) method, which also allows for dependence between the

wavelet coefficients. The modeling is accomplished by using a mixture of two normal-inverse-gamma (NIG) distributions as a joint prior on wavelet coefficients and noise variance within each block. In this sense it is a generalization of the ABWS method by Chipman et al. [27]. The authors group the wavelet coefficients into nonoverlapping, mutually independent blocks $\boldsymbol{d}_{jH}$ of size l_j. Assuming a normal likelihood $\boldsymbol{d}_{jH} \sim N(\boldsymbol{\theta}_{jH}, \sigma^2 I)$, the prior model is specified as

$$\begin{aligned} \boldsymbol{\theta}_{jH}, \sigma^2 | \gamma_j &\sim \gamma_j NIG(\alpha, \delta, \mathbf{0}, \Sigma_j) + (1-\gamma_j) NIG(\alpha, \delta, \mathbf{0}, \Delta_j), \\ \gamma_j &\sim Ber(p_j), \end{aligned}$$

where the covariance matrices are specified as $\Sigma[s,t] = c_j^2 \rho^{|s-t|}$ and $\Delta[s,t] = \tau_j^2 \rho^{|s-t|}$, which is in the same fashion as in [4]. The first part of the above mixture prior models wavelet coefficients with large magnitude ($c_j \gg 1$) and the second part captures small coefficients (τ_j is small), similarly to the ABWS method. The posterior distribution for the model above remains a mixture of NIG distribution with mixing weights updated by the observed wavelet coefficients. The posterior and marginal distributions are derived in the paper. The posterior mean of $\boldsymbol{\theta}_{jH}$ becomes

$$\mathbb{E}(\boldsymbol{\theta}_{jH} | \boldsymbol{d}_{jH}) = A_{jH}(\boldsymbol{d}_{jH}) \boldsymbol{m}^{\star}_{jH} + (1 - A_{jH}(\boldsymbol{d}_{jH})) \boldsymbol{m}^{\star\star}_{jH}, \tag{14.12}$$

where

$$A_{jH}(\boldsymbol{d}_{jH}) = \frac{p_j \frac{|\Sigma_j^{\star}|^{1/2}}{|\Sigma_j|^{1/2}}}{p_j \frac{|\Sigma_j^{\star}|^{1/2}}{|\Sigma_j|^{1/2}} + (1-p_j) \frac{|\Delta_j^{\star\star}|^{1/2}}{|\Delta_j|^{1/2}} + \left[\frac{\alpha + \boldsymbol{d}_{jH}^T (I - \Delta_j^{\star\star}) \boldsymbol{d}_{jH}}{\alpha + \boldsymbol{d}_{jH}^T (I - \Sigma_j^{\star}) \boldsymbol{d}_{jH}} \right]^{-(\delta + l_j)/2}}$$

and

$$\begin{aligned} \Sigma_j^{\star} &= (\Sigma_j^{-1} + I)^{-1}, \\ \Delta_j^{\star\star} &= (\Delta_j^{-1} + I)^{-1}, \\ \boldsymbol{m}^{\star}_{jH} &= \Sigma_j^{\star} \boldsymbol{d}_{jH}, \\ \boldsymbol{m}^{\star\star}_{jH} &= \Delta_j^{\star\star} \boldsymbol{d}_{jH}. \end{aligned}$$

The posterior mean (14.12) is a linear combination of two affine shrinkage estimators $\boldsymbol{m}^{\star}_{jH}$ and $\boldsymbol{m}^{\star\star}_{jH}$, which preserve the smooth part and remove the noise, respectively. The weight $A_{jH}(\boldsymbol{d}_{jH})$ depends on the observed wavelet coefficients in a nonlinear fashion. For more details on hyperparameter selection, simulations, and performance the reader is referred to [34].

Huerta [47] proposed a multivariate Bayes wavelet shrinkage method which allows for correlations between wavelet coefficients corresponding to the same level of detail. The paper assumes the multivariate normal likelihood for the observed wavelet coefficients, that is,

$$d|\theta,\sigma^2 \sim N(\theta,\sigma^2 I_n).$$

Note that the wavelet coefficients are not grouped into blocks, as opposed to the methods discussed before. The prior structure is specified as

$$\begin{aligned}\theta|\tau^2 &\sim N(0,\tau^2\Sigma),\\ \sigma^2 &\sim \mathrm{IG}(\alpha_1,\delta_1),\\ \tau^2 &\sim \mathrm{IG}(\alpha_2,\delta_2),\end{aligned}$$

where Σ is an $n\times n$ matrix defining the prior correlation structure among wavelet coefficients. The matrix is specified as a block diagonal matrix, where each block defines the correlation structure for different wavelet decomposition level. The building blocks of matrix Σ are defined in the same way as in the methods discussed above.

Since there is no closed-form expression for the marginal posterior $\pi^\star(\theta|d)$, a standard Gibbs sampling procedure is adopted to obtain posterior inferences on the vector of wavelet coefficients d. For further details and applications of the method the reader is referred to [47].

Wang and Wood [93] considered a different approach for Bayesian block shrinkage, based directly on the block sum of squares. The sum of squares of the coefficients in the block forms a noncentral chi-square random variable, on which the Bayesian model is formulated. Let $\hat{c}_B$ denote the block of empirical wavelet coefficients, B representing the labels and $n(B)$ the number of labels, in general. Then the assumed likelihood function is $\hat{c}_B \sim N_{n(B)}(c_B,\sigma^2 I_{n(B)})$. Define $z=\|\hat{c}_B\|^2=\Sigma_{i\in B}\hat{c}_i^2$, the sum of squares of the coefficients in the block. It follows that $z\sim\chi^2_m(z|\rho,\sigma^2)$, that is, z has noncentral χ^2 distribution with $m=n(B)$ degrees of freedom, noncentrality parameter $\rho=\|c_B\|^2$, and scale parameter σ^2. The authors formulate the prior model on the noncentrality parameter as

$$\begin{aligned}\rho|\beta &\sim \chi^2_m(\rho|0,\beta^{-1}),\\ \beta|\sigma^2,\theta &\sim F(\beta|\sigma^2,\theta).\end{aligned}$$

In other words this specifies a central χ^2 density with m degrees of freedom and scale parameter β^{-1} as a prior for ρ and specifies a prior for β with cumulative distribution function $F(\beta|\sigma^2,\theta)$. Their article focuses on a mixture structure

$$F(\beta|\sigma^2,\theta)=pF(\beta|\sigma^2,\lambda,J=1)+(1-p)F(\beta|\sigma^2,\lambda,J=0),$$

where

$$F(\beta|\sigma^2,\lambda,J=1)=I_{\{\beta=\infty\}}(\beta).$$

Here J is a Bernoulli random variable, with $J=0$ corresponding to a distribution on the right side of the mixture, and $J=1$ referring to a point mass at infinity distribution. Using an identity satisfied by the noncentral χ^2 density the authors provide closed-form equations for the marginal distribution and the posterior mean of ρ for the model setup above. The equations are the function of $F(\beta|\sigma^2,\lambda,J=0)$, which is to be specified. The authors consider four particular cases of this prior, the point mass prior, the power prior, the exponential prior, and general discrete prior. For the power prior—on which the paper focuses on—the marginal distribution and posterior mean of ρ is derived as

$$f\left(z|\sigma^2,\theta\right)=p\chi^2_m\left(\rho|0,\sigma^2\right)+(1-p)\frac{(\lambda+1)\left(2\sigma^2\right)^{\lambda+1}}{\Gamma\left(\frac{1}{2}m\right)z^{\lambda+2}}\gamma\left(\eta,\frac{z}{2\sigma^2}\right),$$

$$\mathbb{E}\left(\rho|z,\sigma^2,\theta\right)=(1-\pi)\left\{m\sigma^2+z-\frac{m\sigma^2+2z}{z/\left(2\sigma^2\right)}\mathscr{C}_{\eta,1}\left(\frac{z}{2\sigma^2}\right)+\frac{4\sigma^4}{z}\mathscr{C}_{\eta,2}\left(\frac{z}{2\sigma^2}\right)\right\},$$

where

$$\begin{aligned}\pi&=\frac{p\chi^2_m(\rho|0,\sigma^2)}{f(z|\sigma^2,\theta)},\\ \mathscr{C}_{\eta,j}(x)&=\gamma(\eta+j,x)/\gamma(\eta,x),\\ \eta&=1+\lambda+\frac{1}{2}m,\\ \gamma(a,x)&=\int_0^x t^{a-1}\mathrm{e}^{-t}\mathrm{d}t.\end{aligned}$$

Hyperparameter σ^2 is estimated analogously to the median absolute deviation estimator suggested by Donoho and Johnstone [35], hyperparameter λ is estimated by a "quick-and-dirty" heuristics, and finally hyperparameter p is estimated by marginal maximum likelihood. Given values of hyperparameters σ^2 and $\theta=(p,\lambda)$, the authors propose to estimate wavelet coefficients c_B by the shrinkage procedure

$$c_B=\hat{c}_B\{\mathscr{B}_{\sigma^2,\theta}(z)/z\}^{\frac{1}{2}}, \tag{14.13}$$

where $\mathscr{B}_{\sigma^2,\theta}(z)$ denotes the posterior mean or posterior median of ρ. The authors report good MSE results based on simulations on well-known test functions. For more details the reader is referred to [93].

There is a wide range of other articles considering Bayesian modeling of neighboring wavelet coefficients. To name a few, [76] use a Bayesian hidden Markov tree (HMT) to model the structure of wavelet coefficients in images. Jansen and Bultheel [48] introduce a geometrical prior model for configurations of wavelet coefficients and combine this with local characterization of a classical thresholding into a Bayesian framework. Sendur and Selesnick [80] use parent–child neighboring relation and Laplacian bivariate prior to derive MAP estimators for wavelet coefficients. Pižurica et al. [73] use a Markov random field (MRF) prior model to incorporate inter- and intrascale dependencies of wavelet coefficients. Portilla et al. [74] models neighborhoods of image wavelet coefficients at adjacent positions and scales using scale mixtures of Gaussians.

A recent non-Bayesian development was proposed by Fryzlewitz [40] in a form of fast, hard-thresholding algorithm based on coupling parents and children in the wavelet coefficient tree.

14.3.3 Complex Wavelet Shrinkage

Wavelet shrinkage methods using complex-valued wavelets provide additional insights to shrinkage process. Lina and Mayrand [61] describes the complex-valued Daubechies' wavelets in detail. Both complex- and real-valued Daubechies' wavelets are indexed by the number of vanishing moments, N. For a given N, there are 2^{N-1} solutions to the defining equations of Daubechies' wavelets, of which not all are distinct. For example, in case $N = 3$, there are four possible solutions to the defining equations, but only two are distinct. Two solutions give the real-valued extremal-phase wavelet and the other two are a complex-valued conjugate pair, giving equivalent complex-valued wavelets. This complex wavelet was also derived by Lawton [56] through "zero-flipping"; he notes that apart from the Haar wavelet, complex wavelets with an odd number of vanishing moments are the only compactly supported wavelets which are symmetric. The complex-valued wavelet transformation can also be represented by a complex-valued matrix W, which is unitary; therefore, $\bar{W}^T W = W\bar{W}^T = I$. Here $\bar{W}$ denotes the complex conjugate of W.

After taking complex wavelet transformation of a real-valued signal, our model becomes

$$d_{jk} = \theta_{jk} + \varepsilon_{jk},$$

where the observed wavelet coefficients d_{jk} are complex numbers at resolution j and location k.

Several papers considering Bayesian wavelet shrinkage with complex wavelets are available. For example, [59, 60, 62] focus on image denoising, in which the phase of the observed wavelet coefficients is preserved, but the modulus of the coefficients is shrunk by the Bayes rule.

Here we summarize the complex empirical Bayes (CEB) procedure proposed by Barber and Nason [14], which modifies both the phase and modulus of wavelet

coefficients by a bivariate shrinkage rule. The authors assume a common i.i.d. normal noise model $\boldsymbol{e} \sim N_n(\mathbf{0}, \sigma^2 I_n)$; however, after taking complex wavelet transform, the real and imaginary parts of the transformed noise $\boldsymbol{\varepsilon} = W\boldsymbol{e}$ become correlated. The authors demonstrate that

$$\begin{aligned}
\mathrm{cov}\{\mathrm{Re}(\boldsymbol{\varepsilon}), \mathrm{Im}(\boldsymbol{\varepsilon})\} &= -\sigma^2 \mathrm{Im}(WW^T)/2, \\
\mathrm{cov}\{\mathrm{Re}(\boldsymbol{\varepsilon}), \mathrm{Re}(\boldsymbol{\varepsilon})\} &= \sigma^2 \{I_n + \mathrm{Re}(WW^T)\}/2, \\
\mathrm{cov}\{\mathrm{Im}(\boldsymbol{\varepsilon}), \mathrm{Im}(\boldsymbol{\varepsilon})\} &= \sigma^2 \{I_n - \mathrm{Re}(WW^T)\}/2. \qquad (14.14)
\end{aligned}$$

Representing the complex-valued wavelet coefficients as a bivariate real-valued random variables, the model for the observed wavelet coefficients becomes

$$d_{jk}|\theta_{jk} \sim N_2(\theta_{jk}, \Sigma_j),$$

where Σ_j is determined by (14.14) for each dyadic level j. Noise variance σ^2 is estimated by the usual median absolute deviation by Donoho and Johnstone [35].

The authors consider a bivariate mixture prior of the form

$$\theta_{jk} \sim p_j N_2(\mathbf{0}, V_j) + (1 - p_j)\delta_0,$$

where δ_0 is the usual point mass probability at $(0,0)^T$. This prior is the bivariate extension of the prior considered by Abramovich et al. [7]. Conjugacy of the normal distribution results in the posterior distribution

$$\theta_{jk}|d_{jk} \sim \tilde{p}_{jk} N_2(\mu_{jk}, \tilde{V}_j) + (1 - \tilde{p}_{jk})\delta_0,$$

where

$$\begin{aligned}
\tilde{p}_{jk} &= \frac{p_j f(d_{jk}|p_j = 1)}{p_j f(d_{jk}|p_j = 1) + (1 - p_j) f(d_{jk}|p_j = 0)}, \\
f(d_{jk}|p_j = 1) &= \frac{1}{2\pi\sqrt{|V_j + \Sigma_j|}} \exp\left\{-\frac{1}{2} d_{jk}^T (V_j + \Sigma_j)^{-1} d_{jk}\right\}, \\
f(d_{jk}|p_j = 0) &= \frac{1}{2\pi\sqrt{|\Sigma_j|}} \exp\left\{-\frac{1}{2} d_{jk}^T \Sigma_j^{-1} d_{jk}\right\}, \\
\tilde{V}_j &= \left(V_j^{-1} + \Sigma_j^{-1}\right)^{-1} \text{ and } \mu_{jk} = \tilde{V}_j \Sigma_j^{-1} d_{jk}.
\end{aligned}$$

The posterior mean of θ_{jk} becomes

$$\mathbb{E}(\theta_{jk}) = \tilde{p}_{jk}\mu_{jk}, \qquad (14.15)$$

which is denoted as "CEB-Posterior mean." The authors consider two additional estimation rules, the phase-preserving "CEB-Keep or kill" and the hybrid "CEB-MeanKill" procedure.

Estimation of the prior parameters p_j and V_j is employed by the data-driven empirical Bayes approach maximizing the logarithm of the marginal likelihood. However, optimizing the bivariate likelihood is more involved because we have more parameters compared to the real-valued case.

Barber and Nason [14] present an extensive simulation study of the CEB method alongside with the phase-preserving CMWS hard-thresholding method also developed in their paper. Simulations show that complex-valued denoising is very effective and dominates existing real-valued wavelet shrinkage methods.

14.3.4 Complex Wavelet Shrinkage via Gibbs Sampling

In this section, we describe a new adaptive wavelet denoising methodology using complex wavelets. The method is based on a fully Bayesian hierarchical model that uses a bivariate mixture prior. The crux of the procedure is computational in which the posterior mean is computed through MCMC simulations.

We build on the results of [14] and formulate a bivariate model in the complex wavelet domain, representing the wavelet coefficients as bivariate real-valued random variables. As standardly done in Bayesian modeling, we formulate a hierarchical model which accounts for the uncertainty of the prior parameters by adopting hyperpriors on them. Since a closed-form solution to the Bayes estimator does not exist, MCMC methodology is applied and an approximate estimator (posterior mean) from the output of simulational runs is computed. Although the simplicity of a closed-form solution is lost, the procedure is fully Bayesian, adaptive to the underlying signal and the estimation of the hyperparameters is automatic via the MCMC sampling algorithm. The estimation is governed by the data and hyperprior distributions on the parameters.

We start with the following hierarchical bivariate Bayesian model on the observed complex-valued wavelet coefficients d_{jk}:

$$\begin{aligned} d_{jk}|\theta_{jk},\sigma^2 &\sim N_2(\theta_{jk},\sigma^2\Sigma_j), \\ \theta_{jk}|\varepsilon_j,C_j &\sim (1-\varepsilon_j)\delta_0+\varepsilon_j EP_2(\mu,C_j,\beta), \end{aligned} \tag{14.16}$$

where EP_2 denotes the bivariate exponential power distribution. The multivariate exponential power distribution is an extension of the class of normal distributions in which the heaviness of tails can be controlled. Its definition and properties can be found in [42]. The prior on the location θ_{jk} is a bivariate extension of the standard mixture prior in the Bayesian wavelet shrinkage literature, consisting of a point mass at zero and a heavy-tailed distribution. As a prior, [14] considered a mixture of point

mass and bivariate normal distribution. A heavy-tailed mixture prior can probably better capture the sparsity of wavelet coefficients; however, in the bivariate case, a closed-form solution is infeasible, and we rely on MCMC simulation.

To specify the general case exponential power prior in (14.16), we use $\mu = 0$, because the wavelet coefficients are centered around zero by their definition. We also fix $\beta = 1/2$, which gives our prior the following form:

$$\pi(\theta|C) = \frac{1}{8\pi|C|^{1/2}} \exp\left\{-\frac{1}{2}\left(\theta' C^{-1}\theta\right)^{1/2}\right\}. \tag{14.17}$$

The prior in (14.17) is equivalent to the bivariate double exponential distribution. The univariate double exponential prior was extensively used in the real-valued wavelet context, hence it is natural to extend it to the bivariate case.

From model (14.16) it is apparent that the mixture prior on θ_{jk} is set level-wise, for each dyadic level j, which ensures the adaptivity of the method. Quantity $\sigma^2\Sigma_j$ represents the scaled covariance matrix of the noise for each decomposition level, and C_j represents the level-wise scale matrix in the exponential power prior. Explicit expression for the covariance (Σ_j) induced by white noise in complex wavelet shrinkage can be found in [14] and mentioned above in (14.14). We adopt the approach described in their paper to model the covariance structure of the noise.

Instead of estimating hyperparameters σ^2, ε_j, and C_j, we specify hyperprior distributions on them in a fully Bayesian manner. We specify a conjugate inverse gamma prior on the noise variance σ^2 and an inverse-Wishart prior on the matrix C_j describing the covariance structure of the spread prior of θ_{jk}. Mixing weight ε_j regulates the strength of shrinkage of a wavelet coefficient to zero. We specify a "noninformative" uniform prior on this parameter, allowing the estimation to be fully governed by the data.

For computational purposes, we represent our exponential power prior as a scale mixtures of multivariate normal distributions, which is an essential step for efficient Monte Carlo simulation. From [43], the bivariate exponential power distribution with $\mu = 0$ and $\beta = 1/2$ can be represented as

$$EP_2(\mu = 0, C_j, \beta = 1/2) = \int_0^\infty N_2(0, vC_j) \frac{1}{\Gamma(3/2)8^{3/2}} v^{1/2} \mathrm{e}^{-v/8} \mathrm{d}v,$$

which is a scale mixtures of bivariate normal distributions with mixing distribution gamma. Using the specified hyperpriors and the mixture representation, the model in (14.16) extends to

$$\begin{aligned} d_{jk}|\theta_{jk}, \sigma^2 &\sim N_2(\theta_{jk}, \sigma^2\Sigma_j), \\ \sigma^2 &\sim \mathrm{IG}(a, b), \\ \theta_{jk}|z_{jk}, v_{jk}, C_j &\sim (1 - z_{jk})\delta_0 + z_{jk} N_2(0, v_{jk}C_j), \end{aligned}$$

$$\begin{aligned} z_{jk}|\varepsilon_j &\sim \text{Ber}(\varepsilon_j), \\ \varepsilon_j &\sim U(0,1), \\ v_{jk} &\sim \text{Ga}(3/2,8), \\ C_j &\sim \text{IW}(A_j,w). \end{aligned} \tag{14.18}$$

Note that, for computational purposes, we also introduced a latent variable z_{jk} in the above model. Variable z_{jk} is a Bernoulli variable indicating whether our parameter θ_{jk} comes from a point mass at zero ($z_{jk} = 0$) or from a bivariate normal distribution ($z_{jk} = 1$). By representing the exponential power prior as a scale mixtures of normals, the hierarchical model in (14.18) becomes tractable, because the full conditional distributions of all the parameters become explicit. Therefore, we can develop a Gibbs sampling algorithm to update all the necessary parameters. We used the sample average $\hat{\theta}_{jk} = \Sigma_i \theta_{jk}^{(i)}/N$ of the simulational runs, as the standard estimator for the posterior mean. To apply the Gibbs sampling algorithm we only need to specify hyperparameters a, b, A_j, and w, which influence lower level of the hierarchical model. The rest of the parameters are updated via the Gibbs sampling procedure. The method is called complex Gibbs sampling wavelet smoother (CGSWS). For more details about the implementation, contact the authors.

Application to Inductance Plethysmography Data

For illustration we apply the described CGSWS method to a real-world data set from anesthesiology collected by inductance plethysmography. The recordings were made by the Department of Anaesthesia at the Bristol Royal Infirmary and represent measure of flow of air during breathing. The data set was analyzed by several authors, for example, [4, 7, 66]. For more information about the data, refer to these papers.

The top part of Fig. 14.3 shows a section of plethysmograph recording lasting approximately 80 s ($n = 4{,}096$ observations), while the bottom part shows the reconstruction of the signal with the CGSWS method. In the reconstruction process we applied $N = 5{,}000$ iterations of the Gibbs sampler of which the first 2,000 was burn-in. The aim of smoothing is to preserve features such as peak heights while eliminating spurious rapid variation. The result provided by the proposed method satisfies these requirements providing a very smooth result. Abramovich et al. [4] report the heights of the first peak while analyzing this data set. In our case the height is 0.8389, which is quite close to the result 0.8433, obtained by Abramovich et al. [4], and better compared to the results obtained by other established methods analyzed in their paper.

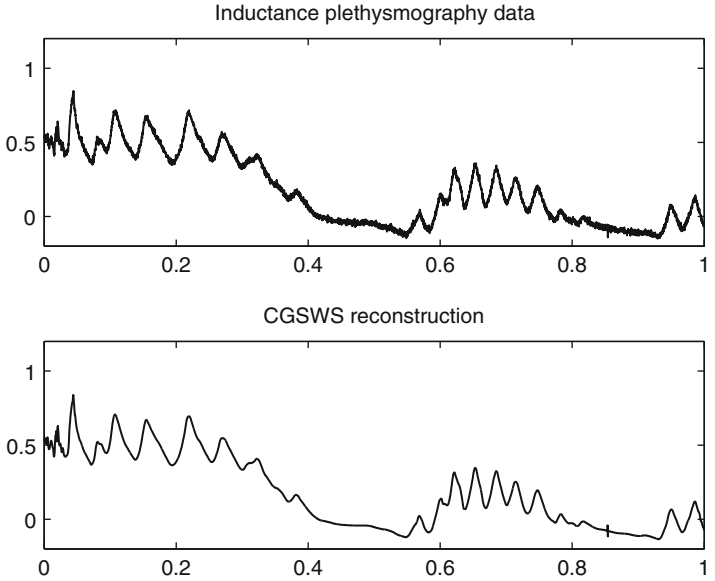

Fig. 14.3 Reconstruction of the (*IPD*) inductance plethysmography data by CGSWS

14.3.5 Bayesian Wavelet Shrinkage in Curve Classification

We consider the paper by Wang et al. [92] to give an application of Bayesian wavelet shrinkage in curve classification. The authors consider Bayesian wavelet-based classification models for binary and multicategory data where the predictor is a random function.

Functional data analysis deals with the analysis of data sets where the units are curves that are ordered measurements on a regular grid. Functional data is frequently encountered in scientific research. Classification of functional data is a relatively new problem, and there are several approaches, from using simple summary quantiles to nonparametric methods using splines. Wang et al. [92] propose a Bayesian wavelet-based classification method, because wavelets are known to have nice properties for representing a wide range of functional spaces including functions with sharp-localized changes. The proposed method unifies wavelet-based regression with logistic classification models, representing functional data using wavelet basis functions and using the wavelet coefficients for classification within a logistic model.

Consider data set $\{\boldsymbol{Y_i}, z_i\},\ i = 1, \ldots, n$, where $\boldsymbol{Y_i}$ is a vector of m measurements and z_i is a binary classification variable. We represent the vector of measurements as $\boldsymbol{Y_i} = \boldsymbol{f_i} + \boldsymbol{\varepsilon_i}$, where $\boldsymbol{f_i}$ is an underlying nonparametric function and $\boldsymbol{\varepsilon_i} \sim N(0, \sigma^2 I)$. Representing functions $\boldsymbol{f_i}$ in wavelet basis we get $\boldsymbol{Y_i} = \boldsymbol{X}\boldsymbol{\beta_i} + \boldsymbol{\varepsilon_i}$, where $\boldsymbol{X}$ is the discrete wavelet transformation matrix and $\boldsymbol{\beta_i}$ is the vector of wavelet coefficients.

The authors consider the following unified hierarchical Bayesian model for wavelet regression and classification:

$$\begin{aligned}
\text{Random function } \boldsymbol{Y_i} &\sim N(\boldsymbol{X\beta_i}, \sigma^2 I),\\
\boldsymbol{\beta_i}, \sigma^2 | \boldsymbol{\eta_i}, \boldsymbol{g} &\sim \text{NIG}(0, \text{diag}(\boldsymbol{\eta_i})\text{diag}(\boldsymbol{g}), a_\sigma, b_\sigma),\\
g_j &\sim \text{IG}(u_j, v_j),\\
\eta_{ijk} &\sim \text{Ber}(\rho_j).\\
\text{Binary outcome } z_i &\sim \text{Ber}(p_i),\\
T_i &\sim N(\boldsymbol{\beta_i^t\theta}, \tau^2), \quad \text{where } T_i = \text{logit}(p_i),\\
\boldsymbol{\theta}, \tau^2 | \boldsymbol{\gamma}, \boldsymbol{h} &\sim \text{NIG}(0, \text{diag}(\boldsymbol{\gamma})\text{diag}(\boldsymbol{h}), a_\tau, b_\tau),\\
h_j &\sim \text{IG}(c_j, d_j),\\
\gamma_{jk} &\sim \text{Ber}(\pi_j),
\end{aligned} \tag{14.19}$$

for $i = 1, \ldots, n$, $j = 1, \ldots, \log_2 m$, and $k = 0, \ldots, 2^j - 1$.

The first part in (14.19) is a model for the observed random functions $\boldsymbol{Y_i}$, where variable selection priors for the wavelet coefficients are adopted from the Bayesian wavelet modeling literature similar to [34]. Parameter g_j is a scaling parameter, and parameter η_{ijk} is the usual latent indicator variable to model the sparsity of the wavelet representation. The second part in (14.19) is a classification model for variable $z_i \in \{0, 1\}$ taking unit value with unknown probability p_i. The logistic classification model relates the wavelet coefficients $\boldsymbol{\beta_i}$ to the latent variable $T_i = \text{logit}(p_i)$ through a linear model $T_i = \boldsymbol{\beta_i^t\theta} + \delta_i$, where $\delta_i \sim N(0, \tau^2)$ and where $\boldsymbol{\theta}$ is a vector of regression coefficients. Similar variable selection prior for $\boldsymbol{\theta}$ is assumed as for $\boldsymbol{\beta_i}$ to reduce the dimensionality of the problem.

For functional data with binary outcomes the model in (14.19) is an extension of a standard classification model with an additional layer of functional regression model. Because the posterior distribution of the parameters is not available in a standard form, posterior inference has to rely on MCMC methods. Wang et al. [92] derive the full conditional distributions for the parameters, which allow for implementation of a Gibbs sampling algorithm. The model in (14.19) is also extended to multicategory classification by the authors.

14.3.5.1 Application to Leaf Data

Wang et al. [92] analyzed a data set from [53] that contains leaf images of six different species. The data was converted into a pseudo-time series by measuring local angle and trace of the leaf images. For a purpose of binary classification analysis one maple (*Circinatum*) and one oak (*Garryana*) species were selected with 150 instances. Example curves adopted from [92] can be seen in Fig. 14.4.

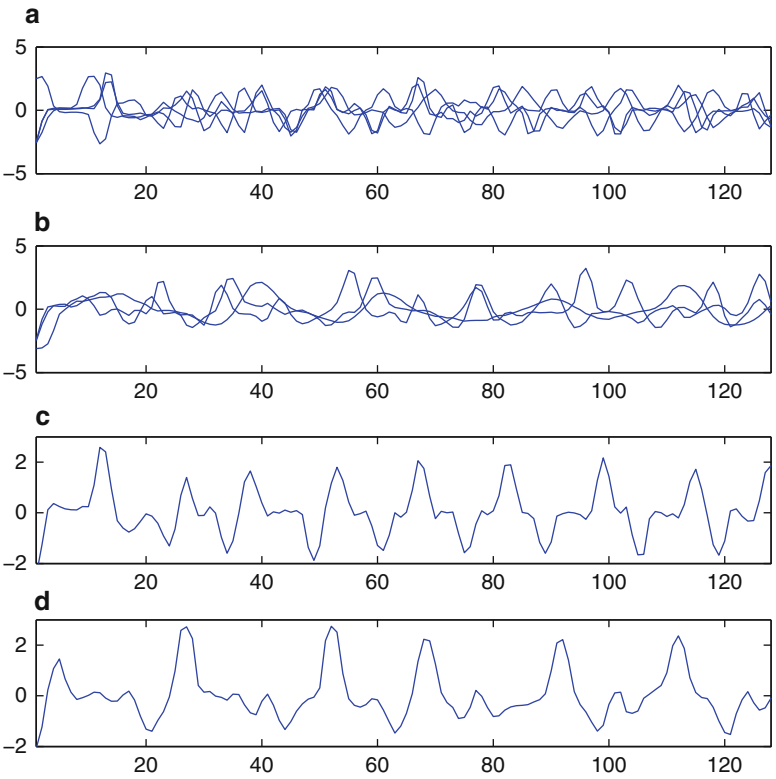

Fig. 14.4 Adopted from [92]: "Pseudo-time series curves from leaf images. (**a**) and (**b**) Every other curve in two species in the data set, 33 of *Circinatum* and 42 of *Garryana*. (**c**) and (**d**) Example of single curve from two species, *Circinatum* and *Garryana*"

The classification was carried out by randomly selecting 140 curves from the training and ten curves from the testing set. This was repeated 20 times, and the correct classification rate (CCR) was reported. The proposed wavelet-based classification method had CCR=94% and outperformed all other methods considered, including empirical Bayes thresholding plugged into a support vector machine (SVM) classifier. The authors carried out analysis for other existing and simulated data sets, including nonequispaced and multicategory data, and reported good performance. For more details the reader is referred to [92].

14.3.6 Related Work

There are numerous papers related to wavelet shrinkage and wavelet regression. Here we list some additional references related to the topics discussed in this chapter, as a repository for researchers interested in the area.

For related overview summaries about wavelet methods see [3, 12, 67], for example. An excellent critical overview and simulation study comparing different wavelet shrinkage methods can be found in [13]. Articles focusing only on Bayesian wavelet-based modeling include [65, 78, 86].

Some recent results about theoretical properties and optimality of Bayesian wavelet estimators can be found in [1, 2, 16, 17, 51, 70, 71].

There are several papers on Bayesian wavelet estimation in the signal and image processing community. These papers usually specify a single, nonmixture prior on the wavelet coefficients and compute a Bayes estimator. Posterior mode is a popular choice, which is used, for example, by Figueiredo and Nowak [38, 64], who use generalized Gaussian and complexity priors to model wavelet coefficients. Other articles in this group include [18] using approximate α-stable prior, [23] using generalized Gaussian distribution (GCD) as a prior, [37] using Bessel K forms (BKF) densities, and [58] using Besov norm priors for modeling wavelet coefficients. Achim and Kuruoğlu [8] develop a bivariate maximum a posteriori estimator using a bivariate α-stable distribution to model wavelet coefficients in the complex wavelet domain.

Some non-Bayesian improvements related to block thresholding include [20, 22, 24–26, 36], to name a few. More general theoretical results about block empirical Bayes estimation appear in [95].

All Bayesian estimators depend on hyperparameters that have to be specified. Purely subjective elicitation is only possible when considerable knowledge about the underlying signal is available. The empirical Bayes method is an efficient, completely data-driven procedure to estimate the hyperparameters based on marginal maximum likelihood method. Several papers in the literature used this method to estimate hyperparameters of the model. For more information about the method see, for example, papers by Clyde and George [29, 30, 50, 51].

The usual assumptions for wavelet regression are equispaced sampling points with a sample size being a power of 2, i.i.d. normal random errors with zero mean and constant variance. Extension of these assumptions has been considered in several articles. To name a few non-Bayesian procedures, [49] consider wavelet thresholding with stationary correlated noise, and [55] extend wavelet thresholding to irregularly spaced data, to equally spaced data sets of arbitrary size, to heteroscedastic and correlated data, and to data which contains outliers. An early example of a Bayesian wavelet shrinkage method incorporating theoretical results on the covariance structure of wavelet coefficients is by Vannucci and Corradi [84]. Ambler and Silverman [9] allow for the possibility that the wavelet coefficients are locally correlated in both location (time) and scale (frequency). This leads to an analytically intractable prior structure; however, they show that it is possible to draw independent samples from a close approximation to the posterior distribution by an approach based on *coupling from the past*, making it possible to take a simulation-based approach to wavelet shrinkage. Wang and Wood [94] consider a Bayesian wavelet shrinkage method which includes both time and wavelet domain methods to estimate the correlation structure of the noise and a Bayesian block shrinkage procedure based on [93]. Ray and Mallick [75] develop a Bayesian

wavelet shrinkage method to accommodate broad class of noise models for image processing applications. The method is based on the Box-Cox family of power transformations.

Kohn et al. [54] develop a wavelet shrinkage method which incorporates a Bayesian approach for automatically choosing among wavelet bases and averaging of the regression function estimates over different bases.

Barber et al. [15, 79] derive Bayesian credible intervals for Bayesian wavelet regression estimates based on cumulants and saddlepoint approximation, respectively.

Olhede and Walden [69] discuss an "analytic" wavelet thresholding which incorporates information from the discrete Hilbert transform of the signal, creating a complex-valued "analytic" vector. A recent paper describing a data-adaptive thresholding by controlling the false discovery rate (FDR) is by Abramovich et al. [5]. A Bayesian interpretation of the FDR procedure and application to wavelet thresholding can be found in [82].

Application of the Bayesian maximum a posteriori multiple testing (testimation) procedure to wavelet thresholding can be found in [6].

References

1. Abramovich F, Amato U, Angelini C (2004) On optimality of Bayesian wavelet estimators. Scand J Stat 31:217–234
2. Abramovich F, Angelini C, De Canditiis D (2007) Pointwise optimality of Bayesian wavelet estimators. Ann Inst Stat Math 59:425–434
3. Abramovich F, Bailey TC, Sapatinas T (2000) Wavelet analysis and its statistical applications. Statist 49:1–29
4. Abramovich F, Besbeas P, Sapatinas T (2002) Empirical Bayes approach to block wavelet function estimation. Comput Stat Data Anal 39:435–451
5. Abramovich F, Benjamini Y, Donoho DL, Johnstone IM (2006) Adapting to unknown sparsity by controlling the false discovery rate. Ann Stat 34:584–653
6. Abramovich F, Grinshtein V, Petsa A, Sapatinas T (2010) On Bayesian "testimation" and its application to wavelet thresholding. Biometrika 97:181–198
7. Abramovich F, Sapatinas T, Silverman BW (1998) Wavelet thresholding via Bayesian approach. J Roy Stat Soc Ser B 60:725–749
8. Achim A, Kuruoğlu EE (2005) Image denoising using bivariate α-stable distributions in the complex wavelet domain. IEEE Signal Process Lett 12:17–20
9. Ambler GK, Silverman BW (2004) Perfect simulation for wavelet thresholding with correlated coefficients. In: Technical Report 04:01. Department of Mathematics, University of Bristol
10. Angelini C, Sapatinas T (2004) Empirical Bayes approach to wavelet regression using ε-contaminated priors. J Stat Comput Simul 74:741–764
11. Angelini C, Vidakovic B (2004) Γ-minimax wavelet shrinkage: a robust incorporation of information about energy of a signal in denoising applications. Stat Sinica 14:103–125
12. Antoniadis A (2007) Wavelet methods in statistics: some recent developments and their applications. Stat Surv 1:16–55
13. Antoniadis A, Bigot J, Sapatinas T (2001) Wavelet estimators in nonparametric regression: a comparative simulation study. J Stat Softw 6:1–83
14. Barber S, Nason GP (2004) Real nonparametric regression using complex wavelets. J Roy Stat Soc Ser B 66:927–939

15. Barber S, Nason GP, Silverman BW (2002) Posterior probability intervals for wavelet thresholding. J Roy Stat Soc Ser B 64:189–205
16. Bochkina N, Sapatinas T (2006) On pointwise optimality of Bayes factor wavelet regression estimators. Sankhyā 68:513–541
17. Bochkina N, Sapatinas T (2009) Minimax rates of convergence and optimality of Bayes factor wavelet regression estimators under pointwise risks. Stat Sinica 19:1389–1406
18. Boubchir L, Fadili JM (2006) A closed-form nonparametric Bayesian estimator in the wavelet domain of images using an approximate α-stable prior. Pattern Recogn Lett 27:1370–1382
19. Cai T (1999) Adaptive wavelet estimation: a block thresholding and oracle inequality approach. Ann Stat 27:898–924
20. Cai T (2002) On block thresholding in wavelet regression: adaptivity, block size, and threshold level. Stat Sinica 12:1241–1273
21. Cai T, Silverman BW (2001) Incorporating information on neighboring coefficients into wavelet estimation. Sankhyā, Ser B 63:127–148
22. Cai T, Zhou H (2009) A data-driven block thresholding approach to wavelet estimation. Ann Stat 37:569–595
23. Chang SG, Yu B, Vetterli M (2000) Adaptive wavelet thresholding for image denoising and compression. IEEE Trans Image Process 9:1532–1546
24. Chicken E (2003) Block thresholding and wavelet estimation for nonequispaced samples. J Stat Plan Inf 116:113–129
25. Chicken E (2005) Block-dependent thresholding in wavelet regression. J Nonparametr Stat 17:467–491
26. Chicken E (2007) Nonparametric regression with sample design following a random process. Commun Stat Theor Meth 36:1915–1934
27. Chipman H, McCulloch R, Kolaczyk E (1997) Adaptive Bayesian wavelet shrinkage. J Am Stat Assoc 92:1413–1421
28. Clyde M, George E (1998) Robust empirical Bayes estimation in wavelets. ISDS discussion paper, Duke University, Institute of Statistics and Decision Sciences
29. Clyde M, George E (1999) Empirical Bayes estimation in wavelet nonparametric regression. In: Müller P, Vidakovic B (eds.) Bayesian inference in wavelet based models. Lecture notes in statistics, vol 141. Springer-Verlag, New York, pp 309–322
30. Clyde M, George E (2000) Flexible empirical Bayes estimation for wavelets. J Roy Stat Soc Ser B 62:681–698
31. Clyde M, Parmigiani G, Vidakovic B (1998) Multiple shrinkage and subset selection in wavelets. Biometrika 85:391–402
32. Cutillo L, Jung YY, Ruggeri F, Vidakovic B (2008) Larger posterior mode wavelet thresholding and applications. J Stat Plan Inference 138:3758–3773
33. Daubechies I (1992) Ten lectures on wavelets. SIAM, Philadelphia
34. De Canditiis D, Vidakovic B (2004) Wavelet Bayesian block shrinkage via mixtures of normal-inverse gamma priors. J Comput Graph Stat 13:383–398
35. Donoho D, Johnstone I (1994) Ideal spatial adaptation by wavelet shrinkage. Biometrika 81:425–455
36. Efromovich S (2004) Analysis of blockwise shrinkage wavelet estimates via lower bounds for no-signal setting. Ann Inst Stat Math 56:205–223
37. Fadili J, Boubchir L (2005) Analytical form for a Bayesian wavelet estimator of images using the Bessel K form densities. IEEE Trans Image Process 14:231–240
38. Figueiredo M, Nowak R (2001) Wavelet-based image estimation: an empirical bayes approach using Jeffreys' noninformative prior. IEEE Trans Image Process 10:1322–1331
39. Flandrin P (1992) Wavelet analysis and synthesis of fractional Brownian motion. IEEE Trans Inf Theor 38:910–917
40. Fryzlewicz P (2007) Bivariate hard thresholding in wavelet function estimation. Stat Sinica 17:1457–1481

41. George EI, McCulloch R (1997) Approaches to Bayesian variable selection. Stat Sinica 7:339–373
42. Gomez E, Gomez-Villegas MA, Marin JM (1998) A multivariate generalization of the power exponential family of distributions. Commun Stat Theor Meth 27:589–600
43. Gomez E, Gomez-Villegasa MA, Marnb JM (2008) Multivariate exponential power distributions as mixtures of normal distributions with Bayesian applications. Commun Stat Theor Meth 37:972–985
44. Hall P, Kerkyacharian G, Picard D (1998) Block threshold rules for curve estimation using kernel and wavelet methods. Ann Stat 26:922–942
45. Hall P, Kerkyacharian G, Picard D (1999) On the minimax optimality of block thresholded wavelet estimators. Stat Sinica 9:33–50
46. Hall P, Penev S, Kerkyacharian G, Picard D (1997) Numerical performance of block thresholded wavelet estimators. Stat Comput 7:115–124
47. Huerta G (2005) Multivariate Bayes wavelet shrinkage and applications. J Appl Stat 32:529–542
48. Jansen M, Bultheel A (2001) Empirical Bayes approach to improve wavelet thresholding for image noise reduction. J Am Stat Assoc 96:629–639
49. Johnstone I, Silverman BW (1997) Wavelet threshold estimators for data with correlated noise. J Roy Stat Soc Ser B 59:319–351
50. Johnstone I, Silverman BW (1998) Empirical Bayes approaches to mixture problems and wavelet regression. In: Technical report. Department of Statistics, Stanford University
51. Johnstone IM, Silverman BW (2005a) Empirical Bayes selection of wavelet thresholds. Ann Stat 33:1700–1752
52. Johnstone IM, Silverman BW (2005b) Ebayesthresh: R programs for empirical Bayes thresholding. J Stat Softw 12:1–38 (With accompanying software and manual.)
53. Keogh E, Folias T (2002) UCR time series classification/clustering page. Computer Science and Engineering Department, University of California Riverside Available at http://www.cs.ucr.edu/eamonn/time_series_data
54. Kohn R, Marron JS, Yau P (2000) Wavelet estimation using Bayesian basis selection and basis averaging. Stat Sinica 10:109–128
55. Kovac A, Silverman BW (2000) Extending the scope of wavelet regression methods by coefficient-dependent thresholding. J Am Stat Assoc 95:172–183
56. Lawton W (1993) Applications of complex valued wavelet transforms to subband decomposition. IEEE Trans Signal Process 41:3566–3569
57. Leporini D, Pesquet J-C (1998) Wavelet thresholding for a wide class of noise distributions. EUSIPCO'98, Rhodes, Greece, pp 993–996
58. Leporini D, Pesquet J-C (2001) Bayesian wavelet denoising: Besov priors and non-gaussian noises. Signal Process 81:55–67
59. Lina J-M (1997) Image processing with complex Daubechies wavelets. J Math Imag Vis 7:211–223
60. Lina J-M, MacGibbon B (1997) Non-linear shrinkage estimation with complex Daubechies wavelets. Proc SPIE, Wavelet Appl Signal Image Process V 3169:67–79
61. Lina J-M, Mayrand M (1995) Complex Daubechies wavelets. Appl Comput Harmon Anal 2:219–229
62. Lina J-M, Turcotte P, Goulard B (1999) Complex dyadic multiresolution analysis. In: Advances in imaging and electron physics, vol 109. Academic Press
63. Mallat S (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 11:674–693
64. Moulin P, Liu J (1999) Analysis of multiresolution image denoising schemes using a generalized Gaussian and complexity priors. IEEE Trans Inf Theor 45:909–919
65. Müller P, Vidakovic B (eds.) (1999c) Bayesian inference in wavelet based models. Lecture notes in statistics, vol 141. Springer-Verlag, New York
66. Nason GP (1996) Wavelet shrinkage using cross-validation. J Roy Stat Soc Ser B 58:463–479

67. Nason GP (2008) Wavelet methods in statistis with R. Springer-Verlag, New York
68. Ogden T (1997) Essential wavelets for statistical applications and data analysis. Birkhäuser, Boston
69. Olhede S, Walden A (2004) 'Analytic' wavelet thresholding. Biometrika 91:955–973
70. Pensky M (2006) Frequentist optimality of Bayesian wavelet shrinkage rules for Gaussian and non-Gaussian noise. Ann Stat 34:769–807
71. Pensky M, Sapatinas T (2007) Frequentist optimality of Bayes factor thresholding estimators in wavelet regression models. Stat Sinica 17:599–633
72. Pericchi LR, Smith AFM (1992) Exact and approximate posterior moments for a normal location parameter. J Roy Stat Soc Ser B 54:793–804
73. Pižurica A, Philips W, Lemahieu I, Acheroy M (2002) A joint inter- and intrascale statistical model for Bayesian wavelet based image denoising. IEEE Trans Image Process 11:545–557
74. Portilla J, Strela V, Wainwright M, Simoncelli E (2003) Image denoising using scale mixtures of Gaussians in the wavelet domain. IEEE Trans Image Process 12:1338–1351
75. Ray S, Mallick BK (2003) A Bayesian transformation model for wavelet shrinkage. IEEE Trans Image Process 12:1512–1521
76. Romberg JK, Choi H, Baraniuk RG (2001) Bayesian tree structured image modeling using wavelet-domain hidden Markov models. IEEE Trans Image Process 10:1056–1068
77. Ruggeri F (1999) Robust Bayesian and Bayesian decision theoretic wavelet shrinkage. In: Müller P, Vidakovic B (eds.) Bayesian inference in wavelet based models. Lecture Notes in Statistics, vol 141. Springer-Verlag, New York, pp 139–154
78. Ruggeri F, Vidakovic B (2005) Bayesian modeling in the wavelet domain. In: Dey DK, Rao CR (eds.) Bayesian thinking: modeling and computation. Handbook of Statistics, vol 25. North-Holland, Amsterdam, pp 315–338
79. Semadeni C, Davison AC, Hinkley DV (2004) Posterior probability intervals in Bayesian wavelet estimation. Biometrika 91:497–505
80. Sendur L, Selesnick IW (2002) Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. IEEE Trans Signal Process 50:2744–2756
81. Simoncelli E (1999) Bayesian denoising of visual images in the wavelet domain. In: Müller P, Vidakovic B (eds.) Bayesian inference in wavelet based models. Lecture Notes in Statistics, vol 141. Springer-Verlag, New York, pp 291–308
82. Tadesse MG, Ibrahim JG, Vannucci M, Gentleman R (2005) Wavelet thresholding with Bayesian false discovery rate control. Biometrics 61:25–35
83. Ter Braak CJF (2010) Bayesian sigmoid shrinkage with improper variance priors and an application to wavelet denoising. Comput Stat Data Anal 51: 1232–1242
84. Vannucci M, Corradi F (1999) Covariance structure of wavelet coefficients: theory and models in a Bayesian perspective. J Roy Stat Soc Ser B 61: 971–986
85. Vidakovic B (1998a) Nonlinear wavelet shrinkage with Bayes rules and Bayes factors. J Am Stat Assoc 93:173–179
86. Vidakovic B (1998b) Wavelet-based nonparametric Bayes methods. In: Dey D, Müller P, Sinha D (eds.) Practical nonparametric and semiparametric bayesian statistics. Lecture Notes in Statistics, vol 133. Springer-Verlag, New York, pp 133–155
87. Vidakovic B (1999) Statistical modeling by wavelets. John Wiley & Sons, Inc., New York
88. Vidakovic B, Müller P (1999) An introduction to wavelets. In: Müller P, Vidakovic B (eds.) Bayesian inference in wavelet based models. Lecture Notes in Statistics, vol 141. Springer-Verlag, New York, pp 1–18
89. Vidakovic B, Ruggeri F (1999) Expansion estimation by Bayes rules. J Stat Plan Inf 79:223–235
90. Vidakovic B, Ruggeri F (2001) BAMS method: theory and simulations. Sankhyā, Ser B 63:234–249
91. Walter GG, Shen X (2001) Wavelets and other orthogonal systems, 2ns edn. Chapman & Hall/CRC, Boca Raton

92. Wang X, Ray S, Mallick BK (2007) Bayesian curve classification using wavelets. J Am Stat Assoc 102:962–973
93. Wang X, Wood ATA (2006) Empirical Bayes block shrinkage of wavelet coefficients via the non-central χ^2 distribution. Biometrika 93:705–722
94. Wang X, Wood ATA (2010) Wavelet estimation of an unknown function observed with correlated noise. Commun Stat Simul Comput 39:287–304
95. Zhang C-H (2005) General empirical Bayes wavelet methods and exactly adaptive minimax estimation. Ann Stat 33:54–100

Chapter 15
Multiparameter Regularization for Construction of Extrapolating Estimators in Statistical Learning Theory

Shuai Lu, Sergiy Pereverzyev Jr., and Sivananthan Sampath

Abstract One-parameter regularization methods, such as the Tikhonov regularization, are used to solve the operator equation for the estimator in the statistical learning theory. Recently, there has been a lot of interest in the construction of the so called extrapolating estimators, which approximate the input–output relationship beyond the scope of the empirical data. The standard Tikhonov regularization produces rather poor extrapolating estimators. In this paper, we propose a novel view on the operator equation for the estimator where this equation is seen as a perturbed version of the operator equation for the ideal estimator. This view suggests the dual regularized total least squares (DRTLS) and multi-penalty regularization (MPR), which are multi-parameter regularization methods, as methods of choice for constructing better extrapolating estimators. We propose and test several realizations of DRTLS and MPR for constructing extrapolating estimators. It will be seen that, among the considered realizations, a realization of MPR gives best extrapolating estimators. For this realization, we propose a rule for the choice of the used regularization parameters that allows an automatic selection of the suitable extrapolating estimator.

S. Lu (✉)
School of Mathematical Science, Fudan University, 200433 Shanghai, China
e-mail: slu@fudan.edu.cn

S. Pereverzyev Jr.
Industrial Mathematics Institute, Johannes Kepler University of Linz, Altenberger Strasse 69, 4040 Linz, Austria
e-mail: pereverzyev@indmath.uni-linz.ac.at

S. Sampath
Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenberger Strasse 69, 4040 Linz, Austria
e-mail: sivananthan.sampath@oeaw.ac.at

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8_15,

15.1 Introduction

Let us consider a system as a functioning entity that takes an input and gives the output. In many scientific studies, one would like to understand how a specific system performs, i.e., given an input how the system produces the output. In particular, one would like to be able to predict the system output. It is very difficult to access the internal structure of many systems, and this complicates the discovery of the system internal functioning mechanisms. In this case, the available information about the system are input–output pairs, which are obtained from the system observations and often called the *empirical data*.

In machine learning [1, 7, 27], a part of computer science, one is concerned with the design and development of algorithms, called (machine) learning algorithms, that allow computers (machines) to predict (to make a decision about) the system output based on the empirical data from the system observations.

The analysis of learning algorithms is done in the framework of (computational) learning theories. One of such theories is the so-called statistical learning theory [30, 33]. According to this theory, the learning algorithm should construct a function, called an *estimator*, that approximates well the relationship between system input and system output. The theory defines the measure of the approximation quality of an estimator and, according to this measure, an *ideal estimator* that has the best approximation quality over a specified function space.

Usually, the ideal estimator cannot be constructed. So, the task of the learning algorithm is to use the empirical data for constructing an estimator that converges to the ideal estimator when the number of observations goes to infinity. The theory suggests a natural approach for constructing an estimator based on the empirical data. This approach leads to an operator equation for the estimator.

As it was observed in [14, 20], there is a similarity between the construction of an estimator and the solution of inverse problems, which are usually formulated as operator equations [15, 17, 18, 31]. Many inverse problems are ill-posed, and for their stable solution, one uses the so-called regularization methods. The operator equation for the estimator in the statistical learning theory is also ill-posed: it does not have a unique solution, and many solutions of this equation are far away from the desired ideal estimator. So, this suggests to apply the regularization methods from the theory of inverse problems. In [16, 29], it was proposed to use the Tikhonov regularization method for solving the operator equation for the estimator. Application of general regularization methods, which are used for solving ill-posed inverse problems, is considered in [4].

One can distinguish between two types of estimators: interpolating and extrapolating. In the case of the interpolating estimator, the inputs in the empirical data are coming from some specified set, and further inputs are also expected to come from this set. One can also say that the interpolating estimator provides a prediction at the unknown inputs within the set that is defined by the existing observations. Whereas, the extrapolating estimator provides a prediction outside this set.

It has been observed that Tikhonov regularization could give good interpolating estimators. On the contrary, the extrapolating estimators that are constructed by the Tikhonov regularization have a rather poor approximation quality. Thus, alternative methods for constructing extrapolating estimators are needed.

Our analysis of the operator equation for the estimator suggests that it can be viewed as a perturbed version of the operator equation for the ideal estimator where both the operator and the right-hand side are modified (perturbed). Recently, in the regularization theory, there has been developed a method, called the dual regularized total least squares (DRTLS) [23–25], which is designed for perturbed operator equations. Therefore, this method can be suggested to solve the operator equation for the estimator. For each realization of DRTLS one can construct a corresponding realization of the so-called multi-penalty regularization (MPR) [8,21] method. This method can be also suggested to solve the operator equation for the estimator.

Tikhonov regularization belongs to a family of the so-called one-parameter regularization methods. On the contrary, DRTLS and MPR are multiparameter regularization methods. This gives them a bigger flexibility for the solution of the perturbed operator equations. And so, one could expect that they could construct better extrapolating estimators.

In this chapter, for solving the operator equation for the estimator, we propose several realizations of DRTLS and MPR. The quality of the extrapolating estimators that are constructed by these realizations will be compared. It will turn out that, among the considered realizations, a realization of MPR gives best extrapolating estimators.

Each realization of a regularization method requires a rule for the choice of the regularization parameters that are used in the method. We will propose such a rule for the mentioned realization of MPR that constructs best extrapolating estimators.

This chapter is organized as follows. In Sect. 15.2, we review the main concepts of the statistical learning theory and derive the operator equation for the estimator. DRTLS and MPR are presented in Sect. 15.3. The perturbation levels in the operator equation for the estimator, which can be used in the application of regularization methods, are estimated in Sect. 15.4. We present several realizations of DRTLS and MPR as well as the comparison of extrapolating estimators that are obtained by these realizations in Sect. 15.5. For the realization that gives the best extrapolating estimators, we propose a rule for the choice of the used regularization parameters in Sect. 15.6. This chapter is finished with conclusions and outlook in Sect. 15.7.

15.2 The Problem of the Construction of an Estimator in the Statistical Learning Theory

In the statistical learning theory, the empirical data $\mathbf{z} = \{(x_i, y_i),\ i = 1, \ldots, n\}$ are seen as the realizations of random variables $(x, y) \in X \times Y$ with a probability density $p(x, y)$. Specifically, we consider the situation when both X and Y are subsets of $\mathbb{R}$.

One of the central problems in the statistical learning theory is the construction of the estimator $f : X \to Y$ that approximates well the relationship between x and y, i.e. $y \approx f(x)$. The common way of measuring the approximation quality of f is the consideration of the expected error:

$$\mathscr{E}^2(f) = \int\limits_{X\times Y} (y - f(x))^2 p(x,y)\mathrm{d}x\mathrm{d}y.$$

Minimization of $\mathscr{E}(f)$ over an appropriate function space leads to an *ideal estimator*. A rather broad function space, for which it is also possible to give an explicit form of the corresponding ideal estimator, is obtained from the following splitting of the density p:

$$p(x,y) = p_x(x)\, p_{y|x}(y|x), \tag{15.1}$$

where $p_x(x) = \int_Y p(x,y)\mathrm{d}y$ is the so-called marginal probability density, and $p_{y|x}(y|x)$ is the so-called conditional probability density for y given x. Then, the minimizer of $\mathscr{E}(f)$ over the function space

$$L^2(X, p_x) = \left\{ f : X \to Y \;\middle|\; \|f\|_p^2 := \int\limits_X f^2(x) p_x(x)\mathrm{d}x < +\infty \right\}$$

is given by

$$f_p(x) = \int\limits_Y y\, p_{y|x}(y|x)\mathrm{d}y.$$

Minimization of the expected error $\mathscr{E}(f)$ over a subspace $\mathscr{H} \subset L^2(X, p_x)$, i.e.,

$$\mathscr{E}(f) \to \min_{f\in\mathscr{H}}, \tag{15.2}$$

means in fact finding a function $f \in \mathscr{H}$ that best approximates $f_p(x)$ in $L^2(X, p_x)$, i.e., a function for which the norm $\|f - f_p\|_p^2$ is minimal. This follows from the following property of the expected error:

$$\mathscr{E}^2(f) = \|f - f_p\|_p^2 + \mathscr{E}^2(f_p).$$

This fact allows the formulation of the operator equation for the solution of (15.2). Let $\mathscr{J} : \mathscr{H} \to L^2(X, p_x)$ be the inclusion operator. Then, solution of (15.2) satisfies the operator equation

$$\mathscr{J}^*\mathscr{J} f = \mathscr{J}^* f_p. \tag{15.3}$$

It is common to assume that this equation is uniquely solvable and define its solution as $f^\dagger$ [13, 14].

Since the probability density p is usually unknown in practice, function $f^\dagger$ provides an ideal estimator that one cannot have but that one tries to approximate using the empirical data $\mathbf{z}$. In the construction of this approximating estimator $f_{\mathbf{z}}$, an important role is played by the so-called *empirical error*

$$\mathscr{E}^2_{\text{emp}}(f) = \frac{1}{n}\sum_{i=1}^{n}(y_i - f(x_i))^2,$$

which is the statistical approximation of the expected error. The first idea for the construction of the estimator $f_{\mathbf{z}} \in \mathscr{H}$ could be to find such $f_{\mathbf{z}}$ that minimizes the empirical error over the function space $\mathscr{H}$, i.e., to solve the following minimization problem:

$$\mathscr{E}_{\text{emp}}(f) \to \min_{f\in\mathscr{H}}. \tag{15.4}$$

However, usually there are many minimizers of $\mathscr{E}_{\text{emp}}(f)$, even such that $\mathscr{E}_{\text{emp}}(f) = 0$, but among them, there are many that are far away from the desired $f^\dagger$.

Before discussing the further steps, let us formulate an operator equation for the minimizer of (15.4). For this purpose, let us define the so-called *sampling operator* $S_{\mathbf{x}} : \mathscr{H} \to \mathbb{R}^n$ that acts as follows $S_{\mathbf{x}} : f \mapsto (f(x_1), f(x_2), \ldots, f(x_n))$ and take the following weighted euclidean norm in $\mathbb{R}^n$: $\|\mathbf{x}\|^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ for $\mathbf{x} \in \mathbb{R}^n$. Then, the empirical error can be written as

$$\mathscr{E}_{\text{emp}}(f) = \|\mathbf{y} - S_{\mathbf{x}}f\|,$$

where $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. And so, the minimization problem (15.4) is equivalent to solving the operator equation

$$S^*_{\mathbf{x}}S_{\mathbf{x}}f = S^*_{\mathbf{x}}\mathbf{y}. \tag{15.5}$$

As the minimization problem (15.4), also the operator equation (15.5) does not have a unique solution, and there are many solutions of (15.5) that are far away from $f^\dagger$.

In [16, 29], it was proposed to modify (15.4) using the Tikhonov regularization:

$$\mathscr{E}^2_{\text{emp}}(f) + \beta\|f\|^2_{\mathscr{H}} \to \min_{f\in\mathscr{H}}, \tag{15.6}$$

where $\beta > 0$ is the so-called regularization parameter. The minimization problem (15.6) is equivalent to solving the following operator equation:

$$(S^*_{\mathbf{x}}S_{\mathbf{x}} + \beta I)f = S^*_{\mathbf{x}}\mathbf{y}, \tag{15.7}$$

where $I : \mathscr{H} \to \mathscr{H}$ is the identity operator. The regularization parameter β has to be chosen such that the corresponding estimator, i.e., the solution of (15.6) or (15.7), approximates well the ideal estimator $f^\dagger$.

As it was mentioned in the Introduction, two situations can be distinguished. In the first situation, the further inputs x are expected to come from a set X_{e} that is

defined by the existing inputs $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. This set X_e is usually $\mathrm{conv}\{x_i,\ i = 1, \ldots, n\}$. The estimators that correspond to this situation are called interpolating estimators. It is quite well-known that the regularization parameter β in (15.7) can be chosen such that the solution of (15.7) is a good interpolating estimator, i.e., it approximates well the ideal estimator $f^{\dagger}$.

On the contrary, in the second situation, the further inputs x are expected to come also outside X_e. The estimators in this situation are called extrapolating estimators. As it will be seen in Sect. 15.5, estimators that are constructed by Tikhonov regularization, i.e., the solutions of (15.7) for various values of β, have bad extrapolating properties. Thus, for the construction of the extrapolating estimators, other methods are needed.

To our best knowledge, it has not been yet observed that equation (15.5) can be viewed as a perturbed version of the operator equation (15.3). As it will be seen in Sect. 15.4, as the number of observations n increases, the operator and the right-hand side in (15.5) approach the operator and the right-hand side in (15.3). More precisely, in corresponding norms, it holds that

$$\lim_{n\to\infty} \| \mathscr{J}^* \mathscr{J} - S_{\mathbf{x}}^* S_{\mathbf{x}} \| = 0,$$
$$\lim_{n\to\infty} \| \mathscr{J}^* f_p - S_{\mathbf{x}}^* \mathbf{y} \| = 0.$$

Such a view suggests that the operator equation (15.5) can be treated by the recently developed DRTLS method [23–25] and the corresponding MPR [8, 21] method.

15.3 Dual Regularized Total Least Squares and Multi-penalty Regularization

Let us assume that there is an operator $A_0 : \mathscr{F} \to \mathscr{G}$, which acts between Hilbert spaces $\mathscr{F}$, space of solutions, and $\mathscr{G}$, space of data. Assume further, that for some perfect data $g_0 \in R(A_0) \subset \mathscr{G}$, there is a unique solution $f_0 \in \mathscr{F}$ to the problem

$$A_0 f = g_0. \tag{15.8}$$

Now, consider the situation when the pair (A_0, g_0) is not known, but instead, we are given an operator $A_h : \mathscr{F} \to \mathscr{G}$ and data $g_\delta \in \mathscr{G}$ that can be seen as noisy versions of the operator A_0 and data g_0 such that

$$\|g_0 - g_\delta\| \le \delta,$$
$$\|A_0 - A_h\| \le h,$$

with some known noise levels $\{\delta, h\} \subset (0, +\infty)$.

For the ill-posed problem (15.8), the operator equation

$$A_h f = g_\delta \tag{15.9}$$

may have no solution, or its solution may be arbitrarily far away from f_0. In this case, the so-called regularization methods [15, 17, 18, 31] are used. Many regularization methods consider the situation when only the data has some noise, and the involved operator A_0 is known exactly. A method, called DRTLS, that takes into account also the noise in the operator has been recently proposed in [23–25]. We review this method below.

Let us fix an operator B that is defined on $\mathscr{F}$ and acts to some other Hilbert space. The idea of DRTLS is to approximate f_0 by the solution of the following minimization problem:

$$\|Bf\| \to \min_{f,A} \quad \text{subject to} \quad \|A - A_h\| \le h,\ \|Af - g_\delta\| \le \delta. \tag{15.10}$$

The solution of this minimization problem for which its constrains are active solves the following operator equation:

$$(A_h^* A_h + \alpha B^* B + \beta I) f = A_h^* g_\delta, \tag{15.11}$$

where $I : \mathscr{F} \to \mathscr{F}$ is the identity operator and α, β satisfy the following conditions:

$$\begin{aligned} \|A_h f_{\alpha,\beta} - g_\delta\| &= \delta + h\|f_{\alpha,\beta}\|, \\ \beta &= -\frac{h(\delta + h\|f_{\alpha,\beta}\|)}{\|f_{\alpha,\beta}\|}, \end{aligned} \tag{15.12}$$

where $f_{\alpha,\beta}$ is the solution of the operator equation (15.11) for the fixed α, β.

An iterative procedure for approximating the pair (α, β) in (15.12) has been proposed in [25]. It should be noted that $\beta < 0$ in (15.12). On the other hand, the operator equation (15.11) with $\alpha > 0$ and $\beta > 0$ arises in the application of the so-called MPR (see, e.g., [8, 21]) to the operator equation (15.9), where the following minimization problem is considered:

$$\|A_h f - g_\delta\|^2 + \alpha\|Bf\|^2 + \beta\|f\|^2 \to \min_f, \tag{15.13}$$

with $\alpha > 0$ and $\beta > 0$.

For the application of DRTLS and MPR one needs to select the operator B, and one needs a procedure, the so-called parameter choice rule [15], to select the appropriate parameters (α, β). Parameter choice rules in the regularization methods need the noise levels in the considered ill-posed inverse problem [3]. In our case, as we mentioned in Sect. 15.2, we propose to view the operator $S_{\mathbf{x}}^* S_{\mathbf{x}}$ and the right-hand side $S_{\mathbf{x}}^* \mathbf{y}$ in (15.5) as the noisy versions of the operator $\mathscr{J}^* \mathscr{J}$ and the right-hand side $\mathscr{J}^* f_p$ in (15.3).

Thus, for the parameter choice rules, we need to estimate perturbation levels measured by

$$\|\mathscr{J}^*\mathscr{J} - S_{\mathbf{x}}^* S_{\mathbf{x}}\| \text{ and } \|\mathscr{J}^* f_p - S_{\mathbf{x}}^* \mathbf{y}\|. \tag{15.14}$$

These estimations are derived in the next section.

15.4 Estimations of the Operator and Data Noise

In the analysis of the problems in the statistical learning theory, one often assumes (see, e.g., [10]) that there are constants $\{\Sigma, M\} \subset (0, +\infty)$ such that

$$\int_Y \left(\exp\left(\frac{|y - f^\dagger(x)|}{M} \right) - \frac{|y - f^\dagger(x)|}{M} - 1 \right) p_{y|x}(y|x)\mathrm{d}y \leq \frac{\Sigma^2}{2M^2} \tag{15.15}$$

for almost all $x \in X$.

Now, we specify the structure of the subspace $\mathscr{H} \subset L^2(X, p_x)$. Since for the functions $f \in \mathscr{H}$ we are interested in their values $f(x)$ for $x \in X$, it is natural to require that the functionals $f(x)$ are continuous on $\mathscr{H}$. Reproducing Kernel Hilbert spaces (RKHS) [2, 6, 12] gives a rich variety of such spaces.

An RKHS is defined by a symmetric positive definite function $K(x, \tilde{x}) : X \times X \to \mathbb{R}$, which is called a kernel. Let us recall that a function $K(x, \tilde{x})$ is symmetric if $K(x, \tilde{x}) = K(\tilde{x}, x)$, and it is positive definite if for any $n \in \mathbb{N}$, any $\{x_1, \ldots, x_n\} \subset X$, and any $\{a_1, \ldots, a_n\} \subset \mathbb{R}$, with at least one $a_i \neq 0$,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j K(x_i, x_j) > 0. \tag{15.16}$$

This property allows to define the scalar product for the functions of the form

$$\begin{aligned} f(x) &= \sum_{i=1}^{n} a_i K(x, x_i), \\ g(x) &= \sum_{j=1}^{m} b_j K(x, \tilde{x}_j) \end{aligned} \tag{15.17}$$

as follows:

$$(f, g)_K = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j K(x_i, \tilde{x}_j). \tag{15.18}$$

The RKHS that is defined (induced) by K is built as the completion of the space of all finite linear combinations (15.17) with respect to the norm that is induced by the scalar product (15.18). This RKHS is denoted by $\mathscr{H}_K$.

Let us also note the following property of linear combinations (15.17) that easily follows from (15.16).

Proposition 1. *Two functions $f(x)$ and $g(x)$ of the form*

$$f(x) = \sum_{i=1}^{n} a_i K(x, x_i),\ \{a_i\}_{i=1}^{n} \subset \mathbb{R},$$

$$g(x) = \sum_{i=1}^{n} \tilde{a}_i K(x, x_i),\ \{\tilde{a}_i\}_{i=1}^{n} \subset \mathbb{R}$$

are equal if and only if $a_i = \tilde{a}_i$ for $i = 1, \ldots, n$.

It is common to put the following additional assumptions on the kernel K [4, 11].

Assumption 1. The kernel K is measurable. It is bounded with

$$\sup_{x \in X} \sqrt{K(x,x)} \leq \kappa < +\infty.$$

The induced RKHS $\mathscr{H}_K$ is separable.

With (15.15) and Assumption 1, we derive the estimates for the operator and data noise (15.14) in the following proposition.

Proposition 2. *Let $f^\dagger$ be the solution of (15.3) with $\mathscr{H} = \mathscr{H}_K$, and let (15.15) and Assumption 1 hold. For $\eta \in (0, 1]$, consider the following set of events:*

$$G_\eta = \{\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in (X \times Y)^n \,|\, \| \mathscr{J}^* \mathscr{J} - S_{\mathbf{x}}^* S_{\mathbf{x}} \| \leq h, \| \mathscr{J}^* f_p - S_{\mathbf{x}}^* \mathbf{y} \| \leq \delta \},$$

with

$$h = h(n, \eta) = \frac{1}{\sqrt{n}} 2\sqrt{2}\kappa^2 \log \frac{4}{\eta},$$

$$\delta = \delta(n, \eta) = 2 \left(\frac{\kappa M}{n} + \frac{\kappa \Sigma + \sqrt{2}\kappa^2 \|f^\dagger\|}{\sqrt{n}} \right) \log \frac{4}{\eta}.$$

Then, $P[G_\eta] \geq 1 - \eta$.

Proof. In [4], the following set of events was considered:

$$G'_\eta = \{\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in (X \times Y)^n \,|\, \| \mathscr{J}^* \mathscr{J} - S_{\mathbf{x}}^* S_{\mathbf{x}} \| \leq h, \| S_{\mathbf{x}}^* S_{\mathbf{x}} f^\dagger - S_{\mathbf{x}}^* \mathbf{y} \| \leq \delta' \},$$

with $\delta' = \delta'(n, \eta) = 2\left(\frac{\kappa M}{n} + \frac{\kappa \Sigma}{\sqrt{n}} \right) \log \frac{4}{\eta}$. Using the results from [9, 14, 28], it was shown in [4] that $P[G'_\eta] \geq 1 - \eta$.

Now, consider $\mathbf{z} \in G'_\eta$, and let us estimate the corresponding data noise from (15.14):

$$\|\mathscr{J}^* f_p - S^*_{\mathbf{x}}\mathbf{y}\| = \|\mathscr{J}^*\mathscr{J} f^\dagger - S^*_{\mathbf{x}}\mathbf{y}\| \le \|\mathscr{J}^*\mathscr{J} f^\dagger - S^*_{\mathbf{x}}S_{\mathbf{x}} f^\dagger\| + \|S^*_{\mathbf{x}}S_{\mathbf{x}} f^\dagger - S^*_{\mathbf{x}}\mathbf{y}\|$$
$$\le h(n,\eta)\|f^\dagger\| + \delta'(n,\eta) = 2\left(\frac{\kappa M}{n} + \frac{\kappa\Sigma + \sqrt{2}\kappa^2\|f^\dagger\|}{\sqrt{n}}\right)\log\frac{4}{\eta}.$$

Thus, $\mathbf{z} \in G_\eta$; therefore, $G_\eta \supset G'_\eta$, and

$$P[G_\eta] \ge P[G'_\eta] \ge 1 - \eta.$$

□

Remark 15.4.1. Since $\frac{1}{n} \le \frac{1}{\sqrt{n}}$ for $n \in \mathbb{N}$, the considered errors can be estimated as

$$\|\mathscr{J}^*\mathscr{J} - S^*_{\mathbf{x}}S_{\mathbf{x}}\| \le \frac{c_h}{\sqrt{n}},$$
$$\|\mathscr{J}^* f_p - S^*_{\mathbf{x}}\mathbf{y}\| \le \frac{c_\delta}{\sqrt{n}}, \tag{15.19}$$

with some constants $\{c_h, c_\delta\} \subset (0, +\infty)$. These estimations can be used in the numerical realization of the regularization methods, which are used for solving (15.5).

15.5 Numerical Realization and Tests

In order to apply DRTLS and MPR to the operator equation (15.5) with $\mathscr{H} = \mathscr{H}_K$ one has to choose the weighted operator B. The simplest choice of this operator is the identity operator $I : \mathscr{H}_K \to \mathscr{H}_K$. With this choice, both DRTLS and MPR become the Tikhonov regularization (TR). Now, let us check the extrapolating properties of the estimators, which are obtained by TR.

In the context of the extrapolating estimators, additionally to the inputs $\{x_i\}_{i=1}^n$, which are presented in the given empirical data $\mathbf{z}$, one also deals with the inputs $\{x_i\}_{i=n+1}^m$ for which the corresponding outputs $\{y_i\}_{i=n+1}^m$ are not known. Moreover, the additional inputs $\{x_i\}_{i=n+1}^m$ are usually outside the $X_e := \mathrm{conv}\{x_i,\ i = 1,\dots,n\}$. Thus, for a good extrapolating estimator, one expects additionally to a good approximation of the ideal estimator $f^\dagger$ over the set X_e also a good approximation of $f^\dagger$ over the $\mathrm{conv}\{x_i,\ i = n+1,\dots,m\}$.

In the statistical learning theory, the following function is often used as an ideal estimator for testing learning algorithms (e.g., [26]):

$$f^\dagger(x) = \frac{1}{10}\left(x + 2\left(\mathrm{e}^{-8\left(\frac{4\pi}{5}-x\right)^2} - \mathrm{e}^{-8\left(\frac{\pi}{2}-x\right)^2} - \mathrm{e}^{-8\left(\frac{3\pi}{2}-x\right)^2}\right)\right), \quad x \in [0, 2\pi].$$

This function belongs to the RKHS that is generated by the kernel $K(x,\tilde{x}) = x\tilde{x} + \exp(-8(x-\tilde{x})^2)$. We will use the RKHS that is generated by this kernel as the space $\mathscr{H}$.

The inputs $\mathbf{x}$ in the empirical data are taken as follows:

$$x_i = \frac{\pi}{10}(i-1),\ i = 1,\dots,15, \tag{15.20}$$

and the outputs $\mathbf{y}$ in the empirical data are generated as follows:

$$y_i = f^\dagger(x_i) + \hat{\delta}\xi_i,\ i = 1,\dots,15, \tag{15.21}$$

where $\{\xi_i\}$ are independent random variables with the uniform distribution over $[-1,1]$. We consider $\hat{\delta} = 0.02$.

The estimator f_β that is constructed by TR with $\mathscr{H} = \mathscr{H}_K$ has the following representation:

$$f_\beta = \sum_{i=1}^{n} c_i K(x,x_i). \tag{15.22}$$

The coefficients $\mathbf{c} = (c_1,c_2,\dots,c_n)'$ in this representation satisfy the following system of linear equations (e.g., [20, 30]):

$$(\mathbf{K} + \beta n\mathbf{I})\mathbf{c} = \mathbf{y}, \tag{15.23}$$

where $\mathbf{I}$ is the identity matrix of order n and $\mathbf{K} = (K(x_i,x_j))_{i,j=1}^n$.

Now, consider the situation when there is an additional input $x_{16} = \frac{\pi}{10}15$. Denote $\|f-g\|_{[a,b]}^\infty := \max\limits_{x\in[a,b]} |f(x)-g(x)|$. In Fig. 15.1, one sees the estimator f_β, which is constructed by TR, and has the minimal extrapolating error $\min\limits_{\beta\in(0,1]} \|f^\dagger - f_\beta\|_{[x_{15},x_{16}]}^\infty$. While it is possible to find such an estimator f_β that has a rather small interpolating error $\|f^\dagger - f_\beta\|_{[x_1,x_{15}]}^\infty$, the result in Fig. 15.1 shows that TR-estimators have rather bad extrapolating properties. Thus, other choices for the operator B in DRTLS and MPR are needed.

The sampling operator, which is scaled with the factor $\sqrt{n}$ for convenience, i.e., $\sqrt{n}S_{\mathbf{x}} : \mathscr{H}_K \to \mathbb{R}^n$, can be proposed as a next choice for the operator B. Such an operator can be viewed as a statistical approximation of the identity operator. But in the contrast to the identity operator such a choice leads to a multiparameter regularization method that is different from TR. In this case, in the application of DRTLS to (15.5), one considers for several pairs of the parameters (α,β) the following operator equation:

$$(T_{\mathbf{x}}^* T_{\mathbf{x}} + \alpha n S_{\mathbf{x}}^* S_{\mathbf{x}} + \beta I)f = T_{\mathbf{x}}^* S_{\mathbf{x}}^* \mathbf{y}, \tag{15.24}$$

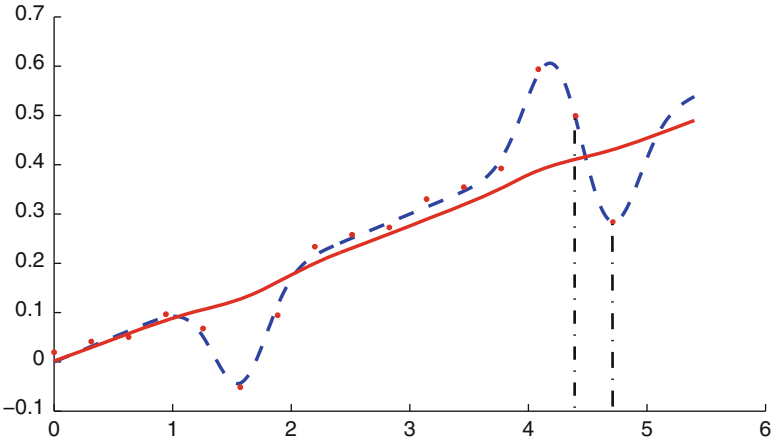

Fig. 15.1 The graph of the TR-estimator f_β (*red curve*) with the smallest extrapolating error $\|f^\dagger - f_\beta\|_{[x_{15},x_{16}]}^\infty$. *Red points* correspond to the empirical data (15.20), (15.21). *Blue dashed curve* is the graph of the ideal estimator $f^\dagger$. The extrapolating interval $[x_{15},x_{16}]$ is located between two vertical *black dashed lines*

where $T_{\mathbf{x}} := S_{\mathbf{x}}^* S_{\mathbf{x}}$. As in the case of TR, the estimator $f_{\alpha,\beta}$ that is constructed by DRTLS with $\mathscr{H} = \mathscr{H}_K$ and $B = \sqrt{n} S_{\mathbf{x}}$ has the representation (15.22). The coefficients $\mathbf{c}$ in this representation satisfy the system of linear equations that is derived in the next proposition.

Proposition 3. *The function $f \in \mathscr{H}_K$ of the form*

$$f(x) = \sum_{i=1}^{n} c_i K(x, x_i) \tag{15.25}$$

solves the operator equation (15.24) *with $\mathscr{H} = \mathscr{H}_K$ if and only if the coefficients $\mathbf{c} = (c_1, c_2, \ldots, c_n)'$ satisfy the following system of linear equations:*

$$[\mathbf{K}^2 + n^2(\alpha \mathbf{K} + \beta \mathbf{I})]\mathbf{c} = \mathbf{K}\mathbf{y}, \tag{15.26}$$

where $\mathbf{I}$ is the identity matrix of order n and $\mathbf{K} = (K(x_i, x_j))_{i,j=1}^n$.

Proof. The derivation of the system (15.26) is similar to the derivation of the system (15.23) (see, e.g., [20, 30]).

It can be shown (e.g., [11, 20]) that the operator $S_{\mathbf{x}}^* : \mathbb{R}^n \to \mathscr{H}_K$ is the following:

$$(S_{\mathbf{x}}^* \mathbf{y})(x) = \frac{1}{n} \sum_{i=1}^{n} K(x, x_i) y_i.$$

For the functions of the form (15.25) we have that

$$T_{\mathbf{x}} f = S_{\mathbf{x}}^* S_{\mathbf{x}} f = \frac{1}{n} \sum_{i=1}^{n} K(x, x_i) \sum_{j=1}^{n} K(x_i, x_j) c_j.$$

Since $T_{\mathbf{x}}^* = T_{\mathbf{x}}$, we get that

$$T_{\mathbf{x}}^* S_{\mathbf{x}}^* \mathbf{y} = T_{\mathbf{x}} S_{\mathbf{x}}^* \mathbf{y} = \frac{1}{n^2} \sum_{i=1}^{n} K(x, x_i) \sum_{j=1}^{n} K(x_i, x_j) y_j,$$

$$T_{\mathbf{x}}^* T_{\mathbf{x}} f = T_{\mathbf{x}}^2 f = \frac{1}{n^2} \sum_{i=1}^{n} K(x, x_i) \sum_{k=1}^{n} K(x_i, x_k) \sum_{j=1}^{n} K(x_k, x_j) c_j.$$

Thus, substituting the function (15.25) into the equation (15.24), we obtain in the left- and right-hand side of this equation a linear combination of functions $\{K(x, x_i)\}_{i=1}^{n}$. Since these linear combinations are equal only if their coefficients are equal (Proposition 1), we obtain the system of linear equations (15.26). □

In the case when the additional inputs $\{x_i\}_{i=n+1}^{m}$ are given, it makes sense to include them into the sampling operator for the operator B. So, let us denote all given inputs as $\tilde{\mathbf{x}} = \{x_i\}_{i=1}^{m}$. Then, instead of $B = \sqrt{n} S_{\mathbf{x}}$, one can propose to consider

$$B = \sqrt{m} S_{\tilde{\mathbf{x}}}. \tag{15.27}$$

The estimator $f_{\alpha,\beta}$, which is constructed by DRTLS with such an operator B, has the following representation:

$$f(x) = \sum_{i=1}^{m} c_i K(x, x_i). \tag{15.28}$$

The system of linear equations for the coefficients $\mathbf{c}$ can be derived similarly to the system (15.26). This system is the following:

$$\left[\mathbf{J}'\mathbf{K}\mathbf{J}\tilde{\mathbf{K}} + n^2(\alpha\tilde{\mathbf{K}} + \beta\tilde{\mathbf{I}})\right] \mathbf{c} = \mathbf{J}'\mathbf{K}\mathbf{y},$$

where $\tilde{\mathbf{I}}$ is the identity matrix of order m, $\tilde{\mathbf{K}} = (K(x_i, x_j))_{i,j=1}^{m}$, and $\mathbf{J} = (a_{ij} \mid i = 1, \ldots, n;\ j = 1, \ldots, m)$ with $a_{ii} = 1$, and $a_{ij} = 0$ when $i \neq j$.

Now, let us check the extrapolating properties of the estimators, which are constructed by DRTLS with B from (15.27). Let us take the empirical data from the test of TR, i.e., (15.20), (15.21), and let us consider two cases of additional inputs:

1. One additional input:

$$x_{16} = \frac{\pi}{10} 15; \tag{15.29}$$

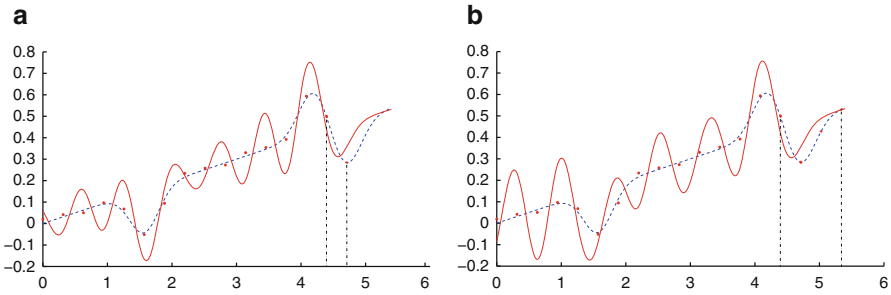

Fig. 15.2 The graphs of the DRTLS-estimators $f_{\alpha,\beta}$ (*red curves*) with the smallest extrapolating errors $\|f^\dagger - f_{\alpha,\beta}\|^\infty_{[x_{15},x_{16}]}$ (**a**) and $\|f^\dagger - f_{\alpha,\beta}\|^\infty_{[x_{15},x_{18}]}$ (**b**). *Red points* correspond to the empirical data (15.20), (15.21). *Blue dashed curve* is the graph of the ideal estimator $f^\dagger$. The extrapolating intervals $[x_{15},x_{16}]$ (a) and $[x_{15},x_{18}]$ (b) are located between two vertical *black dashed lines*

2. Three additional inputs:

$$x_i = \frac{\pi}{10} i,\ i = 16, 17, 18. \tag{15.30}$$

In Fig. 15.2a, b, one sees the estimators $f_{\alpha.\beta}$, which are constructed by DRTLS and which have the minimal extrapolating errors. These estimators have better extrapolating properties than estimators that are constructed by TR, but the approximation of the ideal estimator over the set X_e is rather poor. Can another choice of the operator B improve this situation?

Recently [5], in the context of the statistical learning theory the following penalty functional was considered:

$$\rho(f) = \sum_{i,j=1}^{m} (f(x_i) - f(x_j))^2 w_{ij},$$

where w_{ij} are weights factors, which can be interpreted as edge weights in the data adjacency graph and are usually taken as $w_{ij} = \exp(-(x_i - x_j)^2)$. This functional can be represented as

$$\rho(f) = \|Bf\|^2_K, \text{ with } B = (S^*_{\tilde{\mathbf{x}}} L S_{\tilde{\mathbf{x}}})^{1/2}, \tag{15.31}$$

where the matrix L is the so-called graph Laplacian that is given by $L = D - W$, $W = (w_{ij})_{i,j=1}^m$, $D = (d_{ij})_{i,j=1}^m$ is a diagonal matrix with $d_{ii} = \sum_{j=1}^m w_{ij}$. Thus, $\rho(f)$ can be used in DRTLS.

As in the previous choice of the operator B, it can be shown that the estimator $f_{\alpha,\beta}$, which is constructed by DRTLS with B from (15.31), has the representation (15.28). The coefficients $\mathbf{c}$ in this representation satisfy the following system of linear equations:

$$\left[\mathbf{J}'\mathbf{K}\mathbf{J}\tilde{\mathbf{K}} + n^2(\alpha L\tilde{\mathbf{K}} + \beta\tilde{\mathbf{I}})\right]\mathbf{c} = \mathbf{J}'\mathbf{K}\mathbf{y}.$$

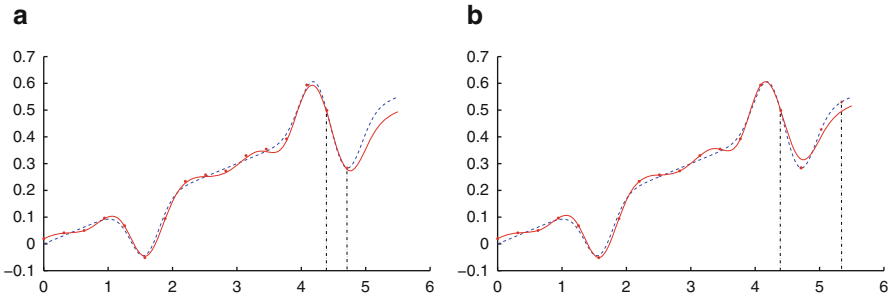

Fig. 15.3 The graphs of the MPR-estimators $f_{\alpha,\beta}$ (*red curves*) with the smallest extrapolating errors $\|f^\dagger - f_{\alpha,\beta}\|^\infty_{[x_{15},x_{16}]}$ (**a**) and $\|f^\dagger - f_{\alpha,\beta}\|^\infty_{[x_{15},x_{18}]}$ (**b**). *Red points* correspond to the empirical data (15.20), (15.21). *Blue dashed curve* is the graph of the ideal estimator $f^\dagger$. The extrapolating intervals $[x_{15}, x_{16}]$ (a) and $[x_{15}, x_{18}]$ (b) are located between two vertical *black dashed lines*

Our numerical experiments show that the obtained estimators are similar to the estimators that correspond to the choice (15.27). Thus, the choice (15.31) of the operator B does not improve the estimators that are constructed by DRTLS. However, MPR with the operator B from (15.31) gives much better estimators. Note, that in contrast to DRTLS, in MPR both regularization parameters α, β are positive.

In Fig. 15.3a, b, one sees the estimators $f_{\alpha,\beta}$, which are constructed by MPR and which have the minimal extrapolating errors. These estimators have not only the best extrapolating properties among the estimators that were considered so far, but they also approximate well the ideal estimator on the set X_e.

In practice, as any regularization method, MPR requires a rule for the choice of the involved regularization parameters. Such a rule is proposed in the next section.

15.6 The Choice of the Regularization Parameters in MPR

The so-called discrepancy principle (DP) (see, e.g., [15]) is a well-known choice rule for the parameters in the regularization methods. Let us consider the general framework of the Sect. 15.3. Denote $\{f_r\}$ the family of the regularized solutions of (15.9) that are constructed by a regularization method. Then, according to DP, one chooses f_r such that

$$\|A_h f_r - g_\delta\| = C\delta,\ C > 1. \tag{15.32}$$

There is a difficulty in using DP for the operator equation (15.5). Namely, a sharp estimate of the noise level δ is not available. Although Proposition 2 and Remark 15.4.1 give theoretical estimations of the noise level δ, in practice the choice of the involved constants there, in particular the constant c_δ in (15.19), is not clear. Moreover, the y-values in the empirical data have often the form (15.21), and

a good estimate of $\hat{\delta}$ can be assumed to be known. In this case, it seems reasonable instead of the condition (15.32), which in the case of the inverse problem (15.5) has the form

$$\|S_{\mathbf{x}}^* S_{\mathbf{x}} f_r - S_{\mathbf{x}}^* \mathbf{y}\| = C\delta, \tag{15.33}$$

to consider the following condition:

$$\|S_{\mathbf{x}} f_r - \mathbf{y}\| = \hat{C}\hat{\delta}. \tag{15.34}$$

Note, that the norms in the above conditions are connected through the following estimate:

$$\|S_{\mathbf{x}}^* S_{\mathbf{x}} f_r - S_{\mathbf{x}}^* \mathbf{y}\| \le \|S_{\mathbf{x}}^*\| \cdot \|S_{\mathbf{x}} f_r - \mathbf{y}\|.$$

Thus, the control of the modified discrepancy $\|S_{\mathbf{x}} f_r - \mathbf{y}\|$ leads to the control of the original discrepancy $\|S_{\mathbf{x}}^* S_{\mathbf{x}} f_r - S_{\mathbf{x}}^* \mathbf{y}\|$. This may be used in the theoretical justification of the condition (15.34).

In MPR $f_r = f_{\alpha,\beta}$, and the condition (15.34), as well as the original condition (15.33), does not uniquely identify the pair of the regularization parameters (α, β). The set of parameters that satisfy (15.34) can be called the discrepancy curve [22].

Among the pairs (α, β) on the discrepancy curve, one can look for the pair that defines the estimator with good extrapolating properties. For this purpose we propose to employ the so-called quasi-optimality principle [32]. The whole procedure for the choice of the appropriate pair of the regularization parameters (α, β) is presented below.

In the numerical realization of the regularization methods, the discrete sets of the regularization parameters in the form of the geometric sequence are frequently used. So, let us consider the following sequence for the parameters β:

$$\beta_k = \beta_0 q^k,\ q > 1,\ k = 0, 1, \ldots, k_{\max}.$$

For each β_k, let us determine α_k for which the condition (15.34) is satisfied, i.e.,

$$\|S_{\mathbf{x}} f_{\alpha_k,\beta_k} - \mathbf{y}\| = \hat{C}\hat{\delta}. \tag{15.35}$$

This can be done using the so-called model function approach [19, 25, 34].

Now, let us define a closeness functional $d(f_{\alpha,\beta}, f_{\alpha',\beta'})$ that describes how close in some sense is the estimator $f_{\alpha,\beta}$ to the estimator $f_{\alpha',\beta'}$. For example, if $x_b \in X$ is an input point of interest, which can be an input without the corresponding output as (15.29), then $d(\cdot,\cdot)$ can be taken as follows:

$$d(f_{\alpha,\beta}, f_{\alpha',\beta'}) = |f_{\alpha,\beta}(x_b) - f_{\alpha',\beta'}(x_b)|. \tag{15.36}$$

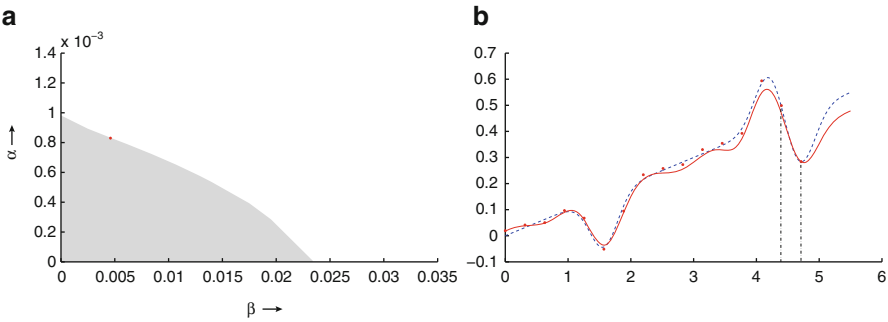

Fig. 15.4 (**a**) The discrepancy region with $\hat{C} = 1$. The *red point* corresponds to the pair (α_k, β_k) that is selected by the principle (15.37). (**b**) The graph of the corresponding MPR-estimator f_{α_k,β_k}. *Red points* correspond to the empirical data (15.20), (15.21). *Blue dashed curve* is the graph of the ideal estimator $f^\dagger$. The extrapolating interval $[x_{15}, x_{16}]$ is located between two vertical *black dashed lines*

Using the idea of the quasi-optimality principle and the chosen closeness functional d, among the pairs (α_k, β_k) that satisfy (15.35), one chooses such a pair that minimizes $d(f_{\alpha_k,\beta_k}, f_{\alpha_{k-1},\beta_{k-1}})$, i.e., one chooses the pair (α_k, β_k) with the following index k:

$$k = \operatorname{argmin}_{k=1,\dots,k_{\max}} \{d(f_{\alpha_k,\beta_k}, f_{\alpha_{k-1},\beta_{k-1}})\}. \tag{15.37}$$

Let us test the proposed procedure. Consider the empirical data (15.20) and (15.21). For these data $\hat{\delta} = 0.02$. Let us consider one additional input x_{16} from (15.29). First, let us take $\hat{C} = 1$. As the closeness functional d, we take (15.36) with $x_b = x_{16}$. In Fig. 15.4a, the discrepancy region, i.e., the region that contains (α, β) that satisfy

$$\|S_{\mathbf{x}} f_{\alpha,\beta} - \mathbf{y}\| \le \hat{C}\hat{\delta}$$

is presented. The red point depicts the pair (α_k, β_k) that is selected by the principle (15.37). In Fig. 15.4b, the corresponding estimator f_{α_k,β_k} is presented. One observes that the chosen estimator is rather close to the best extrapolating estimator in Fig. 15.3a, which demonstrates effectiveness of the proposed parameters choice rule.

By varying the value of the constant $\hat{C}$ one can obtain even better estimators. This is demonstrated in Fig. 15.5, where the results for $\hat{C} = 0.1$ can be found. This suggests that the influence of the constant $\hat{C}$ should be studied in detail.

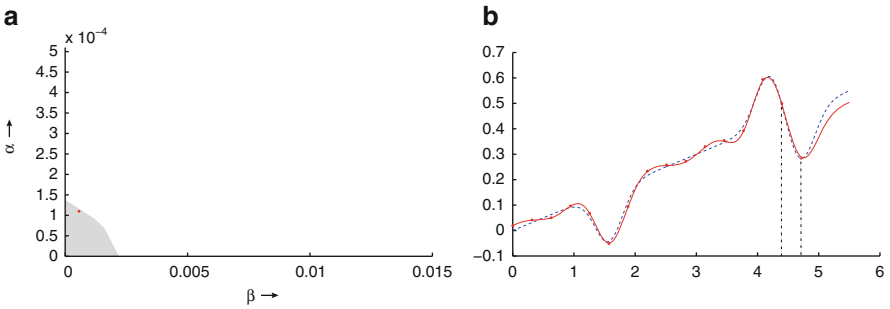

Fig. 15.5 (**a**) The discrepancy region with $\hat{C} = 0.1$. The *red point* corresponds to the pair (α_k, β_k) that is selected by the principle (15.37). (**b**) The graph of the corresponding MPR-estimator f_{α_k,β_k}. *Red points* correspond to the empirical data (15.20), (15.21). *Blue dashed curve* is the graph of the ideal estimator $f^\dagger$. The extrapolating interval $[x_{15}, x_{16}]$ is located between two vertical *black dashed lines*

15.7 Conclusions and Outlook

Construction of good extrapolating estimators requires novel approaches to the problem of constructing an estimator in the statistical learning theory, which can be formulated as an operator equation. We showed that this operator equation can be viewed as a perturbed operator equation with a perturbed operator and perturbed right-hand side. This view suggests the application of the multi-parameter regularization methods, such as DRTLS and MPR. Our numerical tests showed that among the considered realizations of DRTLS and MPR, a realization of MPR gives best extrapolating estimators, and thus, it can be proposed as a method of choice for constructing good extrapolating estimators. As any regularization method, MPR requires an automatic procedure for selecting the involved regularization parameters. We proposed such a procedure and demonstrated its successful performance.

Future research can be concentrated in the following directions.

We derived the perturbation levels in the operator equation for the estimator. This can be considered as a first step in the analysis of the application of multiparameter regularization methods, in particular MPR, for construction of extrapolating estimators. This analysis should be continued until the derivation of the estimates of the estimator general and extrapolating errors.

Other B-operators in MPR, such as (15.27) , can be tried.

It is notable that with (15.27) the system of linear equations, which appears in the numerical realization of the corresponding MPR, has simpler structure than with (15.31). Thus, it is of particular interest to compare the quality of the extrapolating estimators that are constructed by these realizations of MPR.

One can also view

$$S_{\mathbf{x}} f = \mathbf{y} \tag{15.38}$$

as a perturbed operator equation. Application of DRTLS and MPR to (15.38) is quite straightforward, and it is remarkable that the systems of linear equations, which appear in the numerical realization , have a simpler structure in comparison to the systems that arise in the application of DRTLS and MPR to the operator equation (15.5). It remains to be verified whether this application leads to better extrapolating estimators. It should be also noted that the estimation and the interpretation of the perturbation levels in (15.38) have to be addressed.

Finally, a theoretical justification of the proposed choice rule for the regularization parameters in MPR is required. A more detailed study of the influence of the constant $\hat{C}$ in (15.34) and of the connection between the conditions (15.34) and (15.33) should be also done.

Acknowledgments S. Lu is supported by the National Natural Science Foundation of China (No.11101093) and Shanghai Science and Technology Commission (No.11ZR1402800, No.11PJ1400800). S. Sampath is supported by EU-project "DIAdvisor" performed within 7th Framework Programme of EC.

References

1. Alpaydin E (2004) Introduction to machine learning (adaptive computation and machine learning). MIT Press
2. Aronszajn N (1950) Theory of reproducing kernels. Trans Am Math Soc 68:337–404
3. Bakushinskii AB (1984) Remarks on choosing regularization parameter using the quasi-optimality and ratio criterion. USSR Comp Math Math Phys 24:181–182
4. Bauer F, Pereverzev S, Rosasco L (2007) On regularization algorithms in learning theory. J Complex 23:52–72
5. Belkin M, Niyogi P, Sindhwani V (2006) Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. J Mach Learning Res 7:2399–2434
6. Berg C, Christensen JPR, Ressel P (1984) Harmonic analysis on semigroups. Theory of positive definite and related functions. Springer, New York
7. Bishop CM (2006) Pattern recognition and machine learning. Springer, Berlin
8. Brezinski C, Redivo-Zaglia M, Rodriguez G, Seatzu S (2003) Multi-parameter regularization techniques for ill-conditioned linear systems. Numer Math 94:203–228
9. Caponnetto A, De Vito E (2005) Fast rates for regularized least-squares algorithm. In: CBCL Paper 248/AI Memo 2005-013. Massachusetts Institute of Technology, Cambridge, MA
10. Caponnetto A, De Vito E (2007) Optimal rates for the regularized least-squares algorithm. Found Comput Math 7:331–368
11. Carmeli C, De Vito E, Toigo A (2006) Vector valued reproducing kernel Hilbert spaces of integrable functions and Mercer theorem. Anal Appl Singap 4:377–408
12. Cucker F, Smale S (2002) On the mathematical foundations of learning. Bull Am Math Soc New Ser 39:1–49
13. De Vito E, Rosasco L, Caponnetto A (2006) Discretization error analysis for Tikhonov regularization. Anal Appl Singap 4:81–99
14. De Vito E, Rosasco L, Caponnetto A, De Giovannini U, Odone F (2005) Learning from examples as an inverse problem. J Mach Learning Res 6:883–904
15. Engl HW, Hanke M, Neubauer A (1996) Regularization of inverse problems. Kluwer Academic Publishers, Dordrecht

16. Girosi F, Poggio T (1990) Regularization algorithms for learning that are equivalent to multilayer networks. Science 247:978–982
17. Hofmann B (1999) Mathematik inverser Probleme. Teubner, Stuttgart
18. Kirsch A (1996) An introduction to the mathematical theory of inverse problems. Springer, Berlin
19. Kunisch K, Zou J (1998) Iterative choices of regularization parameters in linear inverse problems. Inverse Probl 14:1247–1264
20. Kurkova V (2010) Learning as an inverse problem in reproducing kernel hilbert spaces. Technical report, institute of computer science, academy of sciences of the Czech Republic
21. Lu S, Pereverzev SV (2011) Multi-parameter regularization and its numerical realization. Numer Math 118:1–31
22. Lu S, Pereverzev SV, Shao Y, Tautenhahn U (2010) Discrepancy curves for multi-parameter regularization. J Inv Ill-Posed Probl 18:655–676
23. Lu S, Pereverzev SV, Tautenhahn U (2008) Dual regularized total least squares and multi-parameter regularization. Comp Meth Appl Math 8:253–262
24. Lu S, Pereverzev SV, Tautenhahn U (2009) Regularized total least squares: computational aspects and error bounds. SIAM J Matrix Anal Appl 31:918–941
25. Lu S, Pereverzev SV, Tautenhahn U (2010) A model function method in regularized total least squares. Appl Anal 89:1693–1703
26. Micchelli CA, Pontil M (2005) Learning the kernel function via regularization. J Mach Learning Res 6:1099–1125
27. Mitchell TM (1997) Machine learning. McGraw Hill, New York
28. Pinelis I, Sakhanenko A (1986) Remarks on inequalities for large deviation probabilities. Theor Probab Appl 30:143–148
29. Poggio T, Girosi F (1990) Networks for approximation and learning. Notices of AMS 78:1481–1497
30. Poggio T, Smale S (2003) The mathematics of learning: dealing with data. Notices Am Math Soc 50(5):537–544
31. Rieder A (2003) Keine Probleme mit inversen Problemen. Eine Einführung in ihre stabile Lösung. Vieweg, Wiesbaden
32. Tikhonov AN, Glasko VB (1965) Use of the regularization method in non-linear problems. USSR Comp Math Math Phys 5:93–107
33. Vapnik VN (1998) Statistical learning theory. Wiley, NY
34. Xie J, Zou J (2002) An improved model function method for choosing regularization parameters in linear inverse problems. Inverse Probl 18:631–643

Index

A
Abramovich, F., 323, 326, 329, 334, 337, 342
Achim, A., 341
Adaptive Bayesian wavelet shrinkage (ABWS), 325–326, 330
Ahsanullah, M., 298, 299
Almost-shift-invariant space
  Paley–Wiener space, 97
  sampling
    modeling signals, 93
    sinc function, 94
Ambler, G.K., 341
Amoroso distribution
  cd f, 301
  characterizations
    first order statistic, 303–304
    *j*th order statistic, pdf, 301–302
    *n*th order statistic, 302–303
  description, 299
  family, 299, 300
  pdf, 299
  special rvs, gamma, 299, 301
Angelini, C., 327
Approximation order, 55, 77
Approximation, polynomials $P_{E-E}[-1,1]$, 34–35
  intervals, 35
  Sinc points, 34, 35
Approximation, polynomials $P_{E-M}[-1,1]$
  classical polynomial, 36
  polynomial approximation, 37
  arbitrary Arc Γ, 36–37
  equi-spaced interpolation, 37
  explicit transformations, 38
Approximation, transformed polynomials, 33–34
Array processing
  applications, 108
  DOA, 106
  grating lobes, 114
ASH. *See* Average shifted histogram (ASH)
Autocorrelation
  deterministic construction, 186
  harmonic analysis, 184
  sampling and reconstruction, 106
  snapshot averaging, 109
  Wiener–Wintner theorem, 185
  WSS, 107
Autocorrelation sequences on Z
  CAZACs and Hadamard sequences, 192–195
  circulant Hadamard matrices and CAZACs, 188–191
  definition, 184–185
  Fejér function, 184
  Fourier transforms, 183
  Hadamard matrices, 186–188
  harmonic analysis, 184
  radar, 184
  triangular (*see* Triangular autocorrelation)
  unimodular sequence, 186
  waveform design, 184
  Wiener–Wintner theorem, 185
Average sampling
  acquisition device, 102
  density, exponentials, 103
  Hilbert space, 102
Average shifted histogram (ASH), 276

B
Ball Joseph, A., 180
BAMS. *See* Bayesian adaptive multiresolution shrinkage (BAMS)

X. Shen and A.I. Zayed (eds.), *Multiscale Signal Analysis and Modeling*,
DOI 10.1007/978-1-4614-4145-8, © Springer Science+Business Media New York 2013

Barber, S., 333, 335, 342
Bargmann–Segal–Foch space
  Cauchy–Schwarz inequality, 154
  chromatic derivatives, 157–158
  Hilbert space, 153
  orthonormal family, 154
  Taylor series, 153–154
Bargmann transform
  Gabor representation problem, 154–155
  Hermite polynomial, 155–156
  Parseval's relation, 156–157
Bargmann, V., 140
Bayesian adaptive multiresolution shrinkage (BAMS)
  $\varepsilon-$contamination prior uses, 322
  double exponential and empirical distribution, 321
  exponential power model, 321
  hard *vs.* soft thresholding rules, 322
  Laplace transforms, 322
  MED rule, 324, 325
  posterior distribution, 323
  posterior median, 323–324
  squared error loss and regular models, 323
Bayesian block shrinkage (BBS)
  Bernoulli random variable, 332
  chi-square random variable and correlations, 331
  diagonal method, 328
  empirical approach and nonlinear, 329
  Gibbs sample and matrix, 331
  HMT and MRF, 333
  hyperparameters and covariance matrix, 329
  mixture structure, 331–332
  MSE and power prior, 332
  Nelder–Mead simplex search method, 329
  NIG distributions and ABWS, 330
  posterior mean, 330
Bayesian hierarchical model, 335
Bayesian inference, 331, 339
Bayesian wavelet regression
  BBS, 328–333
  Box-Cox family, 342
  complex shrinkage, 333–335
  curve classification, 338–340
  FDR, 342
  GCD and BKF densities, 341
  Gibbs sampling, 335–338
  posterior mode, 341
  term-by-term shrinkage
    ABWS and SSVS, 325–326
    cumulative and posterior distribution, 327
    heavy-tailed density and hyperparameters, 326
    linear and orthogonal transform, 325
    LPM, 328
    MAP principle, 327–328
    Γ-minimax, 327
    mixture prior distributions, 326
    nonparametric, 324
    posterior mean rule, 327, 328
    quasi-Cauchy distribution, 327
  time and domain, 341
Bayesian wavelet shrinkage
  BAMS, 321–324
  block model, 320–321
  discrete transformations, 319–320
  Heisenberg's principle, 318
  MCMC, 321
  modeling building blocks, 318
  nonparametric statistical modeling, 317
  realistic statistical models, 320
  regression (*see* Bayesian wavelet regression)
  scaling laws, 318
  three-step procedure, 317–318
  thresholding rules, 320
  whitening property, 318
BBS. *See* Bayesian block shrinkage (BBS)
Berger, J., 322
Bessel K forms (BKF) densities, 341
Biorthogonal Spline Family (BSF), 237–238
BKF densities. *See* Bessel K forms (BKF) densities
Block shrinkage. *See* Bayesian block shrinkage (BBS)
Box-Cox family, 342
Breiman, L., 284
B-splines, 74, 75
Bultheel, A., 333

C

CCR. *See* Correct classification rate (CCR)
CDFs. *See* Cumulative distribution functions (CDFs)
CEB. *See* Complex empirical Bayes (CEB)
Centered Fourier matrix, 258, 259
Certain solutions of generalized Pearson differential equations, 304–305
CGSWS. *See* Complex Gibbs sampling wavelet smoother (CGSWS)
Chaos theory, 213, 223, 224
Characterizations of distributions
  Amoroso, 299–304
  Burr system, 298

Pearson system, 298
power function and uniform distributions, 298
SK, 306–307, 312–313
SKS, 305–309
SKS-type, 306, 309–312
SSK, 304–305
stochastic model, 297
Chebyshev polynomial approximation
error, 45, 46
use, 45
CHF. *See* Congestive heart failure (CHF)
CHf. *See* Continuous Hermite functions (CHf)
Chipman, H., 325, 330
Chromatic derivatives
Fourier transform, 141
integral transforms, 143–145
Legendre polynomials converge, 142
multidimensional, 149–152
polynomials, 149
Sturm–Liouville differential operator, 142–143
Taylor-type expansion, 141
Chromatic series expansions
Bargmann–Segal–Foch space (*See* Bargmann–Segal–Foch space)
Bargmann transform, 154–157
derivatives (*See* Chromatic derivatives)
description, 139, 140
Fourier transformation, 139
orthogonal polynomials
expansions, 148–149
variables, 145–147
quantum mechanical system, 140
Sturm–Liouville boundary-value problems, 140
Clary, S., 258
Classification tree (CT), 283
Clyde, M., 341
Coarrays
MRAs, 111
symmetric, 110
Coiflet, 254
Coifman, R.R., 275
Complex empirical Bayes (CEB)
MeanKill procedure, 335
phase and modulus, wavelet, 333–334
posterior mean, 334–335
Complex Gibbs sampling wavelet smoother (CGSWS), 337
Complex-valued wavelets, 333–335
Compression
datasets, 282
smooth shrinkage rules, 323
wavelet transform, 234
Conditional probability density, 350
Congestive heart failure (CHF), 225–226
Contamination prior, 322, 327
Continuous Hermite functions (CHf)
analogous properties, 273
DoG function, 261–262
finite discrete domain, 258
Gaussian derivatives, 262–264
Hermite polynomials, 258
localized analysis
digital weighting function, 265
window and weighting function, 264
monic polynomial, 258
multiscale analysis, 257
shapes, 258
Convergence, classical cardinal series
Cauchy's estimate, 11
classes, data
arithmetic means, sequence, 22
classical sampling theory, 20
Fourier transforms, 19
standard sampling theorem, 21
symmetric partial sums, 23
definition, 3
functions, analogy, 5
generic constants, 8
hypothesis, coefficients, 14
limiting function, 7
mutatis mutandis, 10
oscillatory nature, 6
sampling-type theorem, 4
special functions
elementary curiosity, 17
expansions, 17
plots, Sgn(x), 17, 18
specific bounds, 15
technical lemma, 12–13
Convolution
discrete wavelet transforms, 239
filters and filter banks, 235
Convolution sampling
Banach spaces, 83
reproducing Kernel Banach subspaces, 99–100
Convolution systems
Cauchy–Schwarz's inequality, 59–60
impulse response, 59
linear time-invariant systems, 55
Parseval's equality, 60
types, 58–59

Coprime arrays
autocorrelation, 108
coarrays, 110–111
extra elements, 111–112
spatial
DOA, 109
mutual coupling, 110
snapshot averaging, 109
uniform samplers, 107, 108
Coprime circulants, 129–130
Coprime filter banks
cascade effect, 116
DFT, 114, 115
nonoverlapping bands, 118
sensor spacings, 115, 117
simulated beams, 118, 119
transfer function, 115
Coprime matrices
construction, integer, 125
families, integer
Bezout's identity, 126
skew-circulant adjugate pair, 127
generation pairs, minors
adjugates, 130–131
arbitrary sizes, 128
circulants, 129–130
Coprime sampling and arrays
application, DOA estimation
MUSIC spectrum, 114
plane wave, 112, 113
commuting coprime matrices, 125–131
coprime arrays, one dimension, 107–112
DFT filter banks and beamforming, 114–118
DOA, 105
fractionally spaced equalizers
equivalent multirate representation, 134
oversampling receiver, 134
zero-forcing, 135
identification, sinusoids, 106
multiple dimensions, 118–125
notations, 107
pulsing, 106
system identification, 131–134
Corradi, F., 341
Correct classification rate (CCR), 340
Correlation integral, 216, 223
Courtney, D., 166
Crooks, G.E., 299
CT. *See* Classification tree (CT)
Cumulative distribution functions (CDFs), 279, 309–310
Cuntz relations
Blaschke product, 170, 173
gram matrix, 170
kernel Hilbert space, 171
orthogonality, 171
pick matrix, 170
vector spaces, 172
Curve classification, Bayesian wavelet shrinkage
binary and multicategory data, 338
functional data analysis, 338
leaf data, 339–340
logistic model and data set, 338
standard classification and unified hierarchical model, 339
Cutillo, L., 327

D
Daubechies family (DBF), 237
Daubechies, I., 234, 237, 238, 243, 245, 248, 251, 252
de Branges, L., 178
De Canditiis, D., 329
Denoising
hard thresholding, 244
signal interference, 243
Detrended fluctuation analysis (DFA), 221, 224
DFA. *See* Detrended fluctuation analysis (DFA)
DHf. *See* Discrete Hermite functions (DHf)
DHmT. *See* Discrete Hermite transform (DHmT)
Difference of two Gaussians (DoG)
darker curve, 262
definition, 261
discrete, 262
retinal ganglion cells, 261–262
scale-space representation, 261
standard deviation, 261
Dilation
DHmT, 267
digital weighting function, 265
discrete Gaussian, 263
parameter, 260
standard deviation, 261
Dimension, 217, 220, 222, 223
Direction of Arrival (DOA)
MRAs, 111
MUSIC spectrum, 114
plane wave, 112, 113
Direction of arrival (DOA), 106
Discrepancy curve, 362
Discrepancy principle (DP), 361–362
Discrete Fourier Transform (DFT)

filter banks and beamforming (*see* Coprime filter banks)
sampling array, 106
Discrete Gaussian derivatives
DHf, 273
Fourier transforms, 264
windowed discrete signal, 265
Discrete Hermite functions (DHf)
advantages, 260
centered Fourier matrix, 259
CHf (*See* Continuous Hermite functions (CHf))
definition, 259
digital function, 261
dilation parameter, plots, 260
multiscale analysis, 265–272
plots, h0,h1,h2, and h61, vectors, 259
Discrete Hermite transform (DHmT)
analysis, input signal, 270
arbitrary digital signal, 263
dilation parameter σ, 267
first-level, 268, 269
inverse, 267
scale parameter, 265
second-level, 269
subband decomposition, 272
subband method, 269
third-level, 269, 270
Discrete wavelet transforms
convolution approach, 239
matrix approach, 240
Discriminant measure
additive, 280
LDB, 293
signature time–frequency map and EMD, 281
subspace, 281
symmetric relative entropy, 283–284
DOA. *See* Direction of Arrival (DOA)
DoG. *See* Difference of two Gaussians (DoG)
Donoho, D.I., 243, 329, 332, 334
DP. *See* Discrepancy principle (DP)
DRTLS. *See* Dual regularized total least squares (DRTLS)
Dual frames, 56, 69
Dual regularized total least squares (DRTLS)
estimator graphs, 360
Hilbert spaces, 352, 353
ill-posed problem, 353
multi-parameter regularization methods, 364
noisy versions, 352
operator B, 357, 359
operator equation, 356, 365
parameter choice rule, 353–354
systems, linear equations, 365
TR, 356

E
Earth mover's distance (EMD)
advantages, 277
cdfs, 279
color and textured images, 277
empirical cdfs, 1D signatures, 279–280
Euclidean distance and signatures, 278–279
Monge-Kantorovich mass transportation problem, 278
optimal and total flow, 278
piles, soil, 277–278
signature, 277
ECDFS. *See* Empirical cumulative distribution functions (ECDFS)
EEG analysis. *See* Electroencephalography (EEG) analysis
Einthoven, W., 233
Electrocardiogram (ECG)
analysis and synthesis filter banks, 240–241
biomedical signals, 241
convolution and filter banks, 235
diagnostic tool, 233
discrete wavelet transforms, 239–240
families
BSF, 237–238
DBF, 237
properties, 238
GBL, 243
heart, 241, 242
multiresolution analysis, 235–236
PRD, 234
process 1
comparison methods, 245–247
denoising, 243–244
PRD and SNR selection, 248–249
term-by-term thresholding, 244–245
process 2
comparison method, 251
global thresholding, 250
rankings, 251
steps, 250–251
process 1 vs. process 2
Biorthogonal 2.4, 253
coiflet, 254
properties, 243
signals, patients, 241, 242
SNR, 234
systems, 236–237
time-frequency analysis techniques, 255

Electroencephalography (EEG) analysis
box-counting dimension, 223
chaotic signals, 223
characteristic scale, 225
DFA, 224
entropy measures, 224
epileptic seizures, 220
Hurst parameter, 223
KS entropy, 223
LZ complexity, 220
random fractal theory, 221
seizure and non-seizure segments, 221
small and large scales, 221, 222
EMD. *See* Earth mover's distance (EMD)
EMD based LDB
additive discriminant measures, 280
definition, 281
dimension, 282
discriminant power, subspace, 281
signature, 281–282
time–frequency map, 280, 281
tree-structure notation, 281
Empirical Bayes, 326, 327, 341
Empirical cumulative distribution functions (ECDFS), 283
Empirical data, 348, 349, 351, 356–361, 363, 364
Empirical error, 351
Entire functions of exponential type, 4, 84
Entropy, 220, 222–224
Exponential power model, 321
Extrapolating estimators. *See* Multiparameter regularization
Extrapolation. *See* Multiparameter regularization

F
False discovery rate (FDR), 342
fBm process. *See* Fractional Brownian motion (fBm) process
FDR. *See* False discovery rate (FDR)
Fejér function, 184, 199, 204–205
Figueiredo, M., 341
Filter. *See* Filter banks
Filter banks
analysis and synthesis, 240–241
cascade effect, 116
convolution (*see* Convolution)
coprime arrays, 114, 115
snapshot averaging, 118
transfer function, 115
Finite unit normed tight frames (FUNTFs), 191
Foch,V., 140
Fractional Brownian motion (fBm) process, 217
Fractionally spaced equalizers, 134, 135
Frames
Laguerre transforms, 93
Riesz bases, 56
signal analysis, 155
Fryzlewicz, P., 333
Functions of generalized order statistics, 310–311
Functions of order statistics, 310
FUNTFs. *See* Finite unit normed tight frames (FUNTFs)

G
Galambos, J., 298
Gaussian derivatives
DHmT, 263
discrete and Fourier transforms, 263, 264
filter functions, 263
vs. Hermite functions, 263
Hermite polynomials, 262
GBL. *See* Global thresholding (GBL)
GCD. *See* Generalized Gaussian distribution (GCD)
Generalized gamma distribution, 300, 301
Generalized Gaussian distribution (GCD), 341
Generalized order statistics (GOS), 310
Generalized Pearson differential equations, 304–305, 307
Generalized sampling
band-limited functions, 52
Cauchy–Schwarz's inequality, 57
convolution systems (*see* Convolution systems)
evaluation functionals, 57
Hilbert space, 58
kernel transform, 58
lattice, Z^d
invariant subspaces, 62
optimal frame bounds, 63
orthogonal relationship, 61
square matrix, 66
vector function, 64–65
regular sampling, $V^2\phi$
Bessel sequence, 69
canonical dual frame, 70
Fourier transform, 70
Hermite cubic splines, 76
isomorphism, 66, 67
L^2-approximation properties, 76–78
Laurent polynomials, 75
pseudo-inverse matrix, 68

reconstruction functions, properties, 72–74
Riesz basis, 71
shifting property, 67–68
scaling functions, 55
sequence, reconstruction functions, 56
shift-invariant spaces, 53–55
WSK sampling theorem, 51
George, E.I., 325, 341
Gibbs sampling
Bayesian hierarchical model, 335
CGSWS, 337
conjugate inverse gamma prior, 336
heavy-tailed mixture prior, 336
IPD, 337–338
MCMC, 335
multivariate exponential power distribution, 335
scale mixtures, 336–337
univariate double exponential prior, 336
Glänzel, W., 298, 307, 314, 315
Glass, L., 217
Global thresholding (GBL), 243
GOS. *See* Generalized order statistics (GOS)
Graph Laplacian, 360
Grassberger, P., 216

H
Hadamard, J., 183, 186–188, 190–195, 206
Hadamard matrices
CAZACs
circulant matrix, 188
constructions, sequence, 192–195
FUNTFs, 191
Parseval's relation, 188, 189
unimodular sequences, 190
square matrix, 186–187
Walsh functions, 187–188
Hall, P., 187, 328
Hamedani, G.G., 298, 299, 306, 309, 315
Hardy space functions
Blaschke product, 161–162
Branges–Rovnyak spaces, 166
Cuntz relations (*see* Cuntz relations)
De Branges Rovnyak spaces
Cuntz relations, 178–179
kernel, 178
multipoint interpolation problems, 180
$H_2^{p\times 2}$, 169
$H_2^{p\times q}$, 174–175
interpolation problems, 175–177
isometries, 164
kernel spaces techniques, 166
Leech's theorem, 167–169
linear dynamical systems, 164–165
matrix functions, 167
matrix-valued Schur functions, 164
multipoint interpolation problem, 162–163
tangential interpolation problems, 162, 163
transfer function, 165
Heart rate variability (HRV) analysis
cardiovascular function, 225
CHF, 225, 226
plane separation, 226, 227
small-scale behaviors, 226
wavelet-based multifractal analysis, 225
Heavy-tailed prior, 336
Hermite cubic splines, 76
Hermite polynomials
CHf, 258
functions, 263
Gaussian derivatives, 262–263
Hidden Markov tree (HMT), 333
Hierarchical model. *See* Bayesian hierarchical model
HMT. *See* Hidden Markov tree (HMT)
HRV analysis. *See* Heart rate variability (HRV) analysis
Huerta, G., 331

I
Ignjatovic, A., 139
Ill-posed inversed problem, 348, 353
Inductance plethysmography data (IPD), 337–338
Intermittent Chaos, 218, 220
Interpolating estimators, 348, 349, 352
Interpolation
cardinal series, 4
polynomial-like interpolation., 26
and Sinc notation (*see* Sinc notation and interpolation formulas)
Interpolation problems, 175–177
Inverse problem
expansion theorems, 81
regularization methods., 348
sampling sequence, 82
IPD. *See* Inductance plethysmography data (IPD)

J
Jansen, M., 333
Johnstone, I.M., 326, 329, 332, 334

K
Kamps, U., 310
Kibria, B.M., 306
Kohn, R., 342
Kolmogorov–Sinai (KS) entropy, 223
Kotz, S., 298
KS entropy. *See* Kolmogorov–Sinai (KS) entropy
Kuruoglu, E.E., 341

L
Lagrange polynomial approximation, 29
Laplace transforms, 322
Larger posterior mode (LPM), 328
Lattices sampling
  coprime, 120
  geometries, 106
  Z^d (*see* Generalized sampling)
Laurent polynomials
  reconstruction functions, 73, 76
  sampling functions, 76
Lawton, W., 333
LDA. *See* Linear discriminant analysis (LDA)
LDB. *See* Local discriminant basis (LDB)
Leech's theorem, 167–169
Left inverse matrices, 70
Lei, J.J., 77
Lempel–Ziv (LZ) complexity, 220, 224
Levy, D., 218
Lina, J.-M., 333
Linear discriminant analysis (LDA), 285, 292
Local discriminant basis (LDB)
  algorithm performance
    bell waveforms, 288–290
    chirp waveforms, 290–293
    dimension reduction, 282–283
    LDA and CT, 283
    LDBK, 283
    LDBKASH *vs.* LDBKEMD, 283, 293
    shape waveforms, 287–288
    synthetic signal datasets, 283
    training and test waveforms misclassification, 283–284
    triangular waveforms, 284–287
  ASH, 276
  classification and discrimination, 276
  depiction, WPT, 276
  EMD (*see* Earth mover's distance (EMD))
  time–frequency dictionary, 275–276
  time–frequency map, 276–277
Lorenz systems, 216
LPM. *See* Larger posterior mode (LPM)
LZ complexity. *See* Lempel–Ziv (LZ) complexity

M
Machine learning algorithms, 348
Mackey, M.C., 217
Mallat, S., 319
Mallick, B.K., 341
Marginal probability density, 350
Markov Chain Monte Carlo (MCMC), 321, 335
Markov random field (MRF), 333
Martens, J.B., 266
Maximum likelihood estimation, 326, 327, 332
Mayrand, M., 333
McCulloch, R., 325
MCMC. *See* Markov Chain Monte Carlo (MCMC)
Minimum redundancy arrays (MRAs), 111
Model function approach, 362
Monge-Kantorovich mass transport, 278
Moore-Penrose left pseudo-inverse, 68, 70
MPR. *See* Multi-penalty regularization (MPR)
MRAs. *See* Minimum redundancy arrays (MRAs)
MRF. *See* Markov random field (MRF)
Mugler, D.H., 258
Müller, P., 322
Multidimensional sampling
  coprimality, 120
  2D frequency tiling
    dense tiling, frequency plane, 124
    DFT, 122
    mutual coupling, 125
    passband regions, 123
  lattice arrays, 118, 120
  monochromatic plane waves, 125
  properties of multidimensional coarrays, 121–122
Multiparameter regularization
  DRTLS and MPR, 352–354
  empirical data, 348
  machine learning algorithms, 348
  numerical realization and tests
    additional input cases, 359–360
    graph Laplacian, 360
    MPR estimators $f_{\alpha,\beta}$, 361
    TR, 356–359
  regularization parameter choice, 361–364
  RKHS, 354–356
  statistical learning theory (*see* Statistical learning theory)

Multi-penalty regularization (MPR)
B-operators, 364
closeness functional descriptions, 362–363
description, 353
discrepancy curve, 362
DP, 361–362
DRTLS, 349, 356, 364
estimators $f_{\alpha,\beta}$ *graphs, 361*
minimal extrapolating errors, 361
model function approach, 362
quasi-optimality principle, 362, 363
Multiresolution analysis, 235–236
Multiscale analysis, DHf
broad context, polynomial transforms, 266
DHmT, 267–270
discrete signals
DHmT transform values, 266
downsampling, 266–267
Fourier transforms, 267
lower-indexed values, 267
plot, input signal, 267, 268
wavelet analysis
biorthogonal 3.3, 270
decomposition, 270–272
wavelets, 266
MUSIC algorithm, 113

N
Nashed, M.Z., 88
Nason, G.P., 333, 335
Newton-Cotes polynomials, 38–39
Nonparametric regression, 320, 324
Nonstationarity, 218–220
Nowak, R., 341

O
Olhede, S., 342
One-parameter regularization methods, 349
Order statistics
first, 298, 299, 303, 309
generalized, 310
*j*th, 301
*n*th, 298, 299, 302, 304, 309
progressive type II, 310
*r*th, 310

P
Paley, R.E.A.C, 187
Paley-Wiener spaces ($PW_{1/2}$)
band-limited functions, 52
Fourier duality technique, 55
properties, Sinc function, 84–86
and properties, sinc function
orthogonality property, 85
shift-invariant space, 94, 97
WSK sampling theorem, 84
shift-invariant spaces, 97
Parameters choice rule, 353
Pearson, K., 298
Percentage root mean square difference (PRD)
and SNR selection, 248–249
wavelet denoising, 234
Perfect autocorrelation. *See* Autocorrelation sequences on Z
Pericchi, L.R., 327
p-frame, 101
Pižurica, A., 333
Polynomial approximation
error, transformed Chebyshev, 45, 46
Sinc and Lagrange, 29
uniform Chebyshev, 45
Polynomial-like interpolation., 26
Portilla, J., 333
Posterior mean
Bayesian estimators, 321
CEB, 335
linear combination, 330
marginal distribution, 332
rule, 327–328
standard estimator, 337
Posterior median, 323, 324, 327, 329, 332
PRD. *See* Percentage root mean square difference (PRD)
Procaccia, I., 216

Q
Quadrature mirror filters, 319
Quasi-Cauchy distribution, 327
Quasi-optimality principle, 362, 363

R
Ray, S., 341
Regularization methods. *See* Multiparameter regularization
Regularization parameter, 349, 351, 352, 361–364
Reproducing kernel Banach spaces, 101–102
Reproducing kernel Hilbert spaces (RKHS)
corresponding data noise from, 356
definition, 354
numerical realization, regularization methods, 356

Reproducing kernel Hilbert spaces (*Cont.*)
operator and data noise, 355
properties, multiplication and composition operators, 172–173
property, linear combinations, 355
Riesz representation theorem, 90
scalar product, 354
square-integrable functions, 91
Riesz bases, 56
RKHS. *See* Reproducing kernel Hilbert spaces (RKHS)
Rochberg, R., 167
Rosenblum, M., 168
Rovnyak, J., 178
Rubner, Y., 277–279
Runge, C., 38–44
Runge's function
derivative, Sinc approximation error, 39, 41
exponential polynomial error, 41, 42
f minus trans circ poly approx error, 44
f-prime minus trans algebraic polynomial prime, 41, 43
Newton–Cotes polynomial error, 39, 40
Sinc approximation error, 39, 40
trans algebraic polynomial error, 41, 43
trans circ poly prime error, 44

S

Saito, N., 275, 276, 283, 284
Sampling
and arrays (*see* Coprime sampling and arrays)
classical sampling theory, 20
function spaces (*see* Sampling expansions)
generalized (*see* Generalized sampling)
Sampling expansions
average, 102–103
Banach space, 82
convolution, reproducing Kernel Banach subspaces, 99–100
engineering approach, 91–92
expansion theorems, 81
Fourier series/integral approach, 83–84
function spaces, 88–91
Kernel Hilbert space, 100–101
mathematical deficiencies
definition, 88
replication images, 86
Shannon sampling theorem, 87
non-bandlimited signals, 82
properties, Sinc Function and the Paley–Wiener space, 84–86
reproducing kernel banach spaces, 101–102
reproducing kernel banach subspaces, Lp, 97–99
RKHS, 92–93
series expansions and integral representations, 81
shift-invariant spaces, 93–94
signals, finite rate of innovation, 95–97
unitarily translation-invariant Hilbert spaces, 94–95
Sampling operator, 351, 357, 359
Sampling theory
band-limited functions, 20
functional analysis and harmonic analysis, 83
lattice geometries, 106
modeling signals, 95
shift-invariant spaces., 94
signal and image processing, 82
Sapatinas, T., 327
Sarason, D., 169
Scale-dependent Lyapunov exponent (SDLE)
biological data analysis
EEG (*see* Electroencephalography (EEG) analysis)
HRV (*see* Heart rate variability (HRV) analysis)
characterization, 227–228
complex systems, 212
definitions and fundamental properties
chaos theory, 213
computational procedure, 215
fBm process, 217
inherent scales and characteristic scale, 215
logarithm and averaging, 215
Lorenz systems, 216, 217
phase space, 214
reconstructed vectors, 214
scaling law, 217, 218
stochastic processes, 217, 218
time series, 214
detection, intermittent Chaos, 218, 219
HRV, 212
multiscale complexity measure, 227
multiscale nature, signals, 212, 213
nonstationarity, 218–220
Scaling law, 217, 218
Schur functions, 167–169
SDLE. *See* Scale-dependent Lyapunov exponent (SDLE)
Selesnick, I.W., 333
Sendur, L., 333

Shakil–Kibria (SK) distributions
characterizations, 312–313
five-parameter solution, 306–307
Shakil–Kibria–Singh (SKS)-type distributions
characterizations
arbitrary positive constant, 311–312
CDF, 309–310
conditional pdf, 311
first and *n*th order statistic, 309
GOS, 310
real-valued function, 309
description, 306
log-normal and inverse Gaussian, 305
Pearson system defined, 306
truncated moments, 305–306
Shakil, M., 304–306
Shakil–Singh–Kibria (SSK) distributions
characterizations
continuous random variable, 307
differential equation, 308–309
real functions, 307–308
confluent hypergeometric functions, 305
parameter, 304–305
Shanbhag, D.N., 298
Shannon, D.C., 224
Shen, X., 25
Shift-invariant space. *See* Almost-shift-invariant space
Shift-invariant spaces
Fourier duality technique, 55
function spaces, 83
Riesz sequence, 53
sampling, 93–94
Zak transform, 54
Signal to-noise ratio (SNR), 234
Signal with finite of innovation
representative sampling expansions, 83
sampling
regularity and decay properties, 97
shift-invariant space, 96
theory, 95
Signature
EMD, 277–280
time–frequency map, 281
Silverman, B.W., 326, 341
Sinc interpolation, 29
Sinc methods, 26, 27, 34
Sinc notation and interpolation formulas
basis functions, 28, 29
derivative, p(x), 30
error estimates, data, 31–33
and Lagrange polynomial approximation, 29
Lagrange polynomial interpolation, 30
mathematical notation, 27
Sinc points
definition, function, 29
Rational function methods, 29
Smith, A.F.M., 327
SNR. *See* Signal to-noise ratio (SNR)
Sobolev space, 89
Sparse sampling, 120
Sparsity, 107, 322, 336, 339
Spectral problem, 103
SSVS. *See* Stochastic search variable selection (SSVS)
Stable generators. *See* Generalized sampling
Statistical learning theory
definition, 348
estimator construction
approximation quality measurement, 350
empirical data, 349
empirical error, 351
expected error, 350
extrapolating estimators, 352
marginal probability and conditional probability density, 350
realizations, random variables, 349
regularization parameter, 351–352
sampling operator, 351
TR, 351
estimator types, 348
ideal estimator, testing learning algorithms, 356
ill-posed, 348
TR, 349
Stein, E.M., 244
Stochastic search variable selection (SSVS), 325–326
Strang-fix conditions, 76–78, 96
Subband decomposition
DHmT values, 272
digital function, 261
multiscale analysis, 266
Support vector machine (SVM) classifier, 340
SVM classifier. *See* Support vector machine (SVM) classifier
Sylvester, J.J., 187, 188
Symmetric partial sums
cardinal series, 20–23
convergence, 4
Symmetric relative entropy, 283
System identification
blocked version, 133
continuous-time channel, 131
multirate building blocks, 132

T
Ter Braak, C.J.F., 328
Thresholding
  analytic wavelet, 342
  BAMS-MED rule, 324
  characterization, classical, 333
  ECG signals, 243
  empirical Bayes, 340
  GBL, 243
  global, 250
  hard and soft, 244, 322
  individual and block, 329
  term-by-term, 244–245
  wavelet shrinkage, 319–320
  wavelet transform, 234
Tikhonov regularization (TR)
  additional input cases, 359–360
  DRTLS-estimator graphs, 360
  empirical data, 357
  estimator graph, 357, 358
  extrapolating properties, 359
  ideal estimator, testing learning algorithms, 356
  linear equations, 358–359
  RKHS, 357
Time-delay embedding, 214
Tomasi, C., 278, 279
TR. *See* Tikhonov regularization (TR)
Triangular autocorrelation
  Fejér function, 195–199
  function, sequence, 195
  Lebesgue measure, 199
  unimodular functions, 199–205
Truncated moments
  Amoroso distribution, 301
  first order statistic, 298, 303–304
  *n*th order statistic, 298, 302–303
  ratio, 314
Two dimensional sampling, 106, 115

U
Unimodular sequences, 183, 190, 192
Unitarily translation-invariant Hilbert spaces, 94–95
Unser, M., 52

V
Vannucci, M., 341
Vetterli, M., 95
Vidakovic, B., 322, 327, 329
von Goethe, J.W., 180

W
Walden, A., 342
Walsh, J.L., 187, 188
Walter, G.G., 25, 53–55, 88
Wang, X., 331, 338, 339, 341
Waveforms
  bell, 288–290
  chirp, 290–293
  shape, 287–288
  triangular, 284–287
Wavelet approximation
  error, 47
  odd period function, 45
  transformed, error, 47
Wavelets, 25
Wavelet shrinkage
  Bayesian, 338–340
  complex
    bivariate real-valued random variables, 334
    CEB, 333–335
    Daubechies and zero-flipping, 333
    Gibbs sampling (*see* Gibbs sampling)
    valued and matrix, 333
  discrete transformations
    coefficients and decomposition, 319
    linear and orthogonal, 319
    quadrature mirror filters, 319
    thresholding, 319–320
Wavelet transform
  discrete, 239–240
  shrinkage, 319–320
Whittaker cardinal series, 142
Whittaker–Shannon–Kotel'nikov (WSK) sampling theorem, 51
Wide-sense stationary (WSS), 107
Wiener, N., 185, 196
Wiener–Wintner theorem, 185
Windowed input signal, 265, 266
Wintner, A., 185
Wolf, A., 215, 218, 223, 224
Wood, A.T.A., 331
WSS. *See* Wide-sense stationary (WSS)

Z
Zak transform, 54, 58