RECURSIVE ESTIMATION AND SPECTRAL ESTIMATION FOR 2-D ISOTROPIC RANDOM FIELDS

by

Ahmed Hossam Tewfik

B.Sc., Cairo University (1982)

S.M., Massachusetts Institute of Technology (1984)

E.E., Massachusetts Institute of Technology (1985)

SUBMITTED TO THE DEPARTMENT OF
ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1987

Signature of Author _____
Department of Electrical Engineering and Computer Science
January 5, 1987

Certified by _____
Bernard C. Levy
Thesis Supervisor

Certified by _____
Alan S. Willsky
Thesis Supervisor

Accepted by _____
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Recursive Estimation and Spectral Estimation for 2-D Isotropic Random Fields

by

Ahmed Hossam Tewfik

Submitted to the Department of Electrical Engineering and Computer Science on January 13, 1987 in partial fulfillment of the requirements for the degree of Doctor of Science

Abstract

In this thesis, we develop recursive estimation algorithms and spectral estimation algorithms for 2-D isotropic random fields. Isotropic random fields are characterized by a mean value that is a constant independent of position and a covariance function that is invariant under all rigid body motions. In some sense, isotropy is the natural extension to multidimensions of the 1-D notion of stationarity. Furthermore, isotropic random fields arise in a number of practical problems. An important property of 2-D isotropic random fields is that when they are expanded in a Fourier series in terms of the polar coordinate angle θ, the Fourier coefficient processes of different orders are uncorrelated.

By expanding 2-D isotropic random fields into Fourier series in terms of the coordinate angle θ, we show that a a wave-number-limited isotropic random field $z(\vec{r})$ whose spectral density function is zero outside a disk of radius B can be reconstructed in the mean-square sense from its observation on the *circles* of radii $r_i = i\pi/B$, $i \in \mathbf{N}$, or of radii $r_i = a_{i,n}/B$, $i \in \mathbf{N}$, where $a_{i,n}$ denotes the ith zero of the nth order Bessel function $J_n(x)$, and where n is arbitrary.

Furthermore, by using Fourier series expansions of such fields, we develop efficient *recursive* smoothing algorithms for isotropic random fields described by *non-causal* internal differential models involving a Laplacian operator. The recursions here are with respect to the radius r in a polar coordinate representation of the fields. We show that the 2-D estimation problem is equivalent to a countably infinite set of 1-D separable two-point boundary value smoothing problems which we solve by using either a Markovianization approach followed by a standard 1-D smoothing algorithm, or by using a smoothing technique recently developed for two-point boundary value problems. We also show that the class of fields that we consider includes all isotropic random fields that can be represented as the output of rational 2-D filters driven by white noise.

In the second part of the thesis, we develop two spectral estimation techniques for isotropic random fields. The first technique is designed to estimate 2-D isotropic covariance functions which are equal to weighted sums of cylindrical harmonics. Such fields are often used to model background noises in geophysics and in ocean acoustics. Our procedure is based on an *eigenanalysis* of the matrix of samples of the *non-stationary* 1-D covariance function corresponding to the zeroth-order Fourier coefficient process of the underlying observations. The second technique is a more general *maximum entropy* spectral estimation algorithm. We show that the maximum entropy spectral estimation problem for isotropic random fields has a *linear* solution which can be computed efficiently by using fast recursions that were developed for the problem of estimating an isotropic random field on the boundary of a disk of observation given noisy measurements of the field inside the disk. The maximum entropy spectral estimation algorithm is based on a key spectral factorization result that we derive. Specifically, we show that any isotropic power spectrum corresponding to a covariance function that is zero outside of a finite disk $2R$ in the spatial domain has a non-symmetric half plane factorization where the spectral factors have a *finite* spatial support which is concentrated on a disk of radius R. Our spectral estimation algorithms are based on the knowledge of the covariance functions of the Fourier coefficient processes corresponding to the underlying field and we present an unbiased and consistent method for estimating these covariance functions from observations of the underlying isotropic field.

Finally, we develop a new uncorrelated series expansion and a spectral representation for isotropic random vector fields. This representation is based on the decomposition of isotropic random vector fields into their solenoidal and potential parts which was studied earlier by a number of researchers. We believe that this series representation may help to extend all of the results of this thesis to the vector case.

The Fourier series expansion approach used in this thesis is easily generalized to higher dimensions where isotropic random fields have to be expanded in terms of spherical harmonics rather than in terms of exponential functions. In fact, this approach can be used to solve problems involving any general random field $u(\cdot)$ defined over a homogeneous space provided that its covariance function is invariant with respect to some group G of motions. In this case, the methods introduced in this thesis can be used if the field $u(\cdot)$ is expanded in terms of the characters of the group G.

Thesis Supervisor: Bernard C. Levy
Title: Associate Professor of Electrical Engineering
Thesis Supervisor: Alan S. Willsky
Title: Professor of Electrical Engineering

Acknowledgement

I wish to thank my thesis supervisors Professors Bernard Levy and Alan Willsky for contributing to the success of this work. Special thanks are due to Professor Levy for his accessibility, his insights and ideas and for the time and effort that he has put into both this thesis and my graduate education. I am most grateful to Professor Willsky for his enthusism, his insights and for the time that he invested in this thesis. It has been a pleasure to work with both Professors Levy and Willsky.

I am also grateful to other members of my doctoral committee Professors Arthur Baggeroer and John Tsitsiklis and Dr. Steve Lang for their comments, help and discussions.

Throughout my graduate years at MIT, I was fortunate to interact with other members of the Faculty. Of those, I would like to thank in particular Professor Jeffrey Shapiro who served as my Master's thesis supervisor and my graduate advisor, and Professors George Verghese and Sanjoy Mitter who gave me an opportunity to teach and to participate in the development of a new course.

I would also like to thank my fellow graduate students, past and present, Carey Bunks, Ken Chou, Brian Coleman, Peter Doershuk, Richard Lamb, Ramine Nikoukhah, Cuneyt Ozveren, Jerry Prince and Robin Rohlicek for their friendship and assistance during my research. Special thanks are due to Richard Lamb for his help with the computer systems and for the nice discussions that we had.

My parents deserve more than what I can express for being there when I needed them, for helping out in any way they could, both financially and spiritually, and for all their love, guidance and encouragement throughout my life.

Finally, I would like to express my deep gratitude to my wife Heba. I will always remember her love, kindness, patience, and encouragement during my last year as a graduate student and our first year of marriage. May this thesis be a partial compensation to her and to my parents for all these years away from home.

To Heba and our expected baby,

To Nadia and Hossam,

To Kamla, Abdallah, and their
expected twins.

Contents

List of Figures

Chapter 1

INTRODUCTION

Problems involving spatially-distributed data and phenomena arise in various fields including image processing, meteorology, geophysical signal processing, oceanography and optical processing. However, the presence of more than one independent variable in such problems leads to an increased computational complexity and raises some interesting questions. How can one deal effectively with the computational complexity of multidimensional problems given that the dimension has a strong effect on the size of a problem? How can recursions be efficiently implemented in several dimensions? What are the new types of estimation and signal processing problems of interest that arise in several dimensions and which do not exist in one-dimension?

A variety of researchers whose work is reviewed below, have made contributions which have helped to answer parts of these questions, and we feel that this thesis is of value in this larger effort. However, much of the existing literature deals with relatively direct extensions of formulations and approaches which have worked well for known classes of problems, namely causal estimation problems. Furthermore, efficient estimation and signal processing algorithms which are general enough to be applicable to a large class of problems of practical interest, are lacking in several dimensions. To gain a better understanding of the problems that we have listed, we focus our attention in this thesis on a class of random fields which is of substantial practical interest and which has enough structure to allow the development of a detailed theory. Specifically, we study *isotropic* random fields (see Chapter 2 for

references), i.e. fields whose mean value is a constant independent of position and whose autocovariance function is invariant under all rigid body motions. Note that isotropy is the natural extension of the 1-D notion of stationarity. Furthermore, isotropic random fields arise in a number of practical problems such as the black body radiation problem [49], the study of the background noise in seismology [13] and ocean acoustics [5], [11], [36] and the investigation of temperature and pressure distributions at a constant altitude in the atmosphere [31]. Our work deals with various aspects of state-space estimation and spectral estimation for isotropic random fields in *two dimensions*, and it emphasizes the use of geometrical symmetries in the development of fast algorithms for isotropic random fields. There is no loss of generality in focusing our interest on two-dimensional estimation problems, since all the essential features of multidimensional estimation and signal processing problems appear already in two dimensions. In fact, all the results that we obtained are directly extendible to higher dimensions.

1.1 LITERATURE REVIEW

In this section we review some previous results on recursive estimation and spectral estimation for 2-D fields. Most of these results are very general in nature and were not specifically developed for isotropic random fields.

A Recursive Estimation for 2-D Fields

The main objective of the recursive estimation part of this thesis is to exploit the geometrical structure of models for spatial processes, and to exploit the geometrical symmetries associated with such processes, in order to develop efficient estimation algorithms for such fields. Central difficulties here are the issues of recursions and the organization of the required computations. In one dimension the ways in which data can be organized for efficient processing are extremely limited, and causality typically provides a natural choice. Furthermore, in one dimension internal differential realizations of random processes have been exploited to develop an efficient

estimation algorithm, namely the Kalman filtering technique. This has led researchers in estimation theory to investigate the extension of the methods for causal 1-D processes to non-causal 2-D fields. The work of Woods and Radewan [91], Habibi [24], Attasi [4], Jain and Angel [28], Wong [87], Wong and Ogier [53] to name a few, has shown that such extensions do exist. However, the methods developed by these researchers are either approximate or can be applied only to a limited class of 2-D fields, namely to fields that can be described by hyperbolic partial differential equations and which therefore are causal in some sense.

We believe that 2-D fields are more naturally and efficiently described by different types of models which are fundamentally non-causal in nature and whose structure may offer significant possibilities for obtaining efficient solutions to estimation problems. Some of these non-causal models have been investigated by Rodriguez [66] and by Adams [2]. In particular, Adams studied estimation techniques for fields $x(\cdot)$ governed by an internal model of the form

$$Lx = Bu \tag{1.1.1}$$

with the boundary conditions

$$v = Vx_b \tag{1.1.2}$$

where L is a linear differential or difference operator, B and V are linear operators, and u and v are multidimensional noise fields. His results show that even though the process $x(\cdot)$ is fundamentally non-causal, the observations associated to $x(\cdot)$ can be processed recursively to generate smoothed estimates of $x(\cdot)$. In this thesis, we exploit the new techniques proposed by Adams together with the geometrical structure of isotropic random vector fields to develop fast recursive estimation algorithms in two dimensions.

It should be mentioned at this point that our work is also motivated by the results of Levy and Tsitsiklis [41] who exploited the geometry of isotropic random vector fields to develop efficient Levinson-like recursions for the input-output estimation of such fields. Note that, unlike the one dimensional case, the most natural problem in higher dimensions is the smoothing problem rather than the causal filtering problem. This is what the works of Adams and Levy and Tsitsiklis have

in common: both focus from the beginning on smoothing problems, even though Adams uses an internal approach, while Levy and Tsitsiklis consider input-output estimation.

Finally, note that an important problem of great practical interest which is special to applications in several dimensions, is the problem of estimating random *vector* fields governed, for example, by Maxwell's equations, the heat equation or the gravitational field equations. This problem has not yet been considered in the literature. An interesting subclass of random vector fields is the subclass of *isotropic* random vector fields. Such fields are characterized by a mean vector and a covariance matrix that are unaffected by any translation, rotation and/or reflection accompanied by simultaneous rotation and/or reflection of the coordinate system with respect to which the components of the vector are determined. Isotropic random vector fields arise in a number of physical problems such as the study of black body radiation [49], the analysis of turbulence [50] and the examination of the gravitational fields of planets [62]. In Chapter 6 of this thesis, we develop a mathematical theory for isotropic random vector fields which we believe will help in extending our results to such fields. However, more work is still required in this area to answer the open problems pertaining to some modeling issues and solution formulation questions. We expect that the study of estimation problems for such fields will help us to deepen our understanding of two and higher dimensional estimation theory in general and will point out the differences existing between one dimensional and multidimensional problems.

B 2-D Spectral Estimation Techniques

As mentioned earlier, a goal of this thesis is the development and understanding of 2-D spectral estimation techniques for isotropic random fields. Multidimensional spectral analysis arises in many applications such as geophysics, sonar, radar, radio astronomy, and for that reason, has received considerable attention. The fundamental problem considered in the literature is that of estimating a power spectrum given samples of a stationary and homogeneous underlying random field. Most of the existing 1-D estimation techniques have been extended to several dimensions.

For example, 2-D maximum entropy procedures have been proposed by Woods [90], Jain and Raganath [29], Lim and Malik [42], Roucos and Childers [69], Lang and McClellan [39], Schott and McClellan [74] to name a few.

It is noteworthy that none of the desirable properties of the 1-D MEM spectral estimation algorithms has yet been established in the 2-D case. In particular, the 1-D MEM method has a linear solution which can be implemented efficiently via the Levinson recursions, whereas the 2-D MEM algorithms involve solving a non-linear constrained optimization problem. Furthermore, some of these 2-D procedures are slow, particularly when they are used to estimate highly peaked spectra, and some are not even guaranteed to converge. By exploiting the special structure of isotropic random fields, we show in Chapter 5 that a linear solution to the isotropic MEM spectral estimation problem can be computed via fast Levinson-like recursions which were developed in the filtering context for isotropic fields by Levy and Tsitsiklis [41]. Another interesting feature in one dimension is that the Levinson recursions give rise to lattice realizations of the underlying stationary process [56]. These lattice realizations have the feature that their stability can be checked by inspection, and have also the property of having a very low sensitivity to roundoff errors. In addition, because of their modular, cascaded structure, the lattice filters give rise to simple signal processing architectures which are now used widely in speech processing and communications applications. Unfortunately, no such structures are available in two or higher dimensions for homogeneous random fields. Our work seems to indicate that such lattice realizations might exist for isotropic fields. However, more work is still required in this area.

Note also that 1-D harmonic retrieval techniques which rely on an eigenanalysis of a covariance matrix, such as Pisarenko's method [58] or the MUSIC method [73], [9], have been extended to the 2-D case by Lang and McClellan [38] and by Wax and Kailath [84] respectively. These researchers extended the 1-D techniques to solve the problem of estimating 2-D power spectra which are equal to a weighted sum of multidimensional *point* impulses. In contrast, we present in Chapter 4 an eigenstructure method for estimating power spectra that are equal to a weighted sum of 2-D *cylindrical* impulses, i.e. which are equal to a sum of impulsive cylindrical

sheets in the wave-number plane.

1.2 THESIS CONTRIBUTIONS

Throughout this thesis, Fourier series expansions of isotropic random fields are used in solving a variety of sampling, recursive estimation and spectral estimation problems for 2-d isotropic random fields. The Fourier coefficient processes in a Fourier series expansion of an isotropic random field in terms of the coordinate angle θ in a polar coordinate representation of the plane, are *uncorrelated.* This is just a consequence of the fact that isotropic covariance functions are invariant under all rigid body motions. Hence, by expanding an isotropic random field in a Fourier series, we are able to reduce any 2-D problem involving such a field into a countably infinite number of *uncorrelated* 1-D problems for the Fourier coefficients. The Fourier series expansion approach is easily generalized to higher dimensions where isotropic random fields have to be expanded in terms of spherical harmonics rather than in terms of exponential functions. In fact, this approach can be used to solve problems involving any general random field $u(\cdot)$ defined over a homogeneous space provided that its covariance function is invariant with respect to some group G of motions. In this case, the methods introduced in this thesis can be used if the field $u(\cdot)$ is expanded in terms of the characters of the group G. Finally, note that Fourier series expansions of 2-D isotropic random fields were used previously by Yadrenko [92] and by Levy and Tsitsiklis [41] to solve filtering and estimation problems for such fields.

By using Fourier series expansions of isotropic random fields we develop in this thesis the following new results:

1. **Sampling theorems for isotropic random fields:**
   In Chapter 2, we develop new sampling theorems for the Fourier coefficient processes corresponding to a wave-number-limited isotropic random field $z(\vec{r})$ whose spectral density function is zero outside a disk of radius B centered at the origin of the wave-number plane. By using these theorems, we show that $z(\vec{r})$ can be reconstructed in the mean-square sense from its observation

on the circles of radii $r_i = i\pi/B$, $i \in \mathbf{N}$, or of radii $r_i = a_{i,n}/B$, $i \in \mathbf{N}$, where $a_{i,n}$ denotes the ith zero of the nth order Bessel function $J_n(x)$, and where n is arbitrary. This result is partly new and partly a generalization of an earlier result of Yadrenko [92] who showed that any wave-number-limited isotropic random field $z(\vec{r})$ whose spectral density function is zero outside a disk of radius B centered at the origin of the wave-number plane can be reconstructed in the mean-square sense from its observation on the circles of radii $r_i = a_{i,0}/B$, $i \in \mathbf{N}$, where $a_{i,0}$ denotes the ith zero of the zeroth order Bessel function $J_0(x)$.

2. **Recursive estimation for isotropic random fields:**
We develop in Chapter 3 efficient *recursive* smoothing algorithms for isotropic random fields described by *non-causal* internal differential models involving a Laplacian operator. The 2-D estimation problem is shown to be equivalent to a countably infinite set of 1-D separable two-point boundary value smoothing problems which we solve by using either a Markovianization approach followed by a standard 1-D smoothing algorithm, or by using the smoothing technique recently developed by Adams [2] for two-point boundary value problems. The desired field estimate is then obtained as a properly weighted sum of the 1-D smoothed estimates. This result is important because, as mentioned in Section 1.1, most of the earlier 2-D recursive smoothing algorithms are either approximate or apply only to the limited class of 2-D fields that are described by hyperbolic partial differential equations and which therefore are causal in some sense.

3. **Eigenstructure spectral estimation method for isotropic random fields:**
In Chapter 4, we present a high resolution spectral estimation method for 2-D isotropic random fields with covariance functions equal to weighted sums of cylindrical harmonics. Such fields are often used to model some types of background noises in geophysics and in ocean acoustics. The approach that we present differs from previous 2-D spectral estimation techniques by the

fact that we take maximal advantage of the symmetries implied by both the isotropy and the special covariance structure of these fields. Our approach is similar in spirit to 1-D harmonic retrieval techniques, such as the MUSIC method, which rely on an eigenanalysis of the covariance matrix. In the 2-D isotropic context, we obtain a spectral estimate by performing an eigenanalysis of the covariance matrix of samples of the zeroth-order Fourier coefficient process in order to extract the cylindrical harmonics. However, unlike the 1-D MUSIC and Pisarenko's methods which use sample values of a stationary 1-D covariance matrix, our procedure uses samples of the *non-stationary* 1-D covariance function corresponding to the zeroth-order Fourier coefficient process of the underlying observations. We also present in Chapter 4 a method for estimating the covariance functions of the Fourier coefficient processes from observations of the underlying isotropic field. Furthermore, we show that our estimates are both unbiased and consistent.

4. **Maximum entropy spectral estimation for isotropic random fields:** We present in Chapter 6 a new *linear* maximum entropy method (MEM) algorithm for 2-D isotropic random fields. Unlike general 2-D covariances, isotropic covariance functions which are positive definite on a disk are known to be extendible. Here, we develop a computationally efficient procedure for computing the MEM isotropic spectral estimate corresponding to an isotropic covariance function which is given over a finite disk of radius $2R$. We show that the isotropic MEM problem has a linear solution and that it is equivalent to the problem of constructing the optimal linear filter for estimating the underlying isotropic field at a point on the boundary of a disk of radius R given noisy measurements of the field inside the disk. Our procedure is based on Fourier series expansions in both the space and wave-number domains of the inverse of the MEM spectral estimate. Furthermore, our method is guaranteed to yield a valid isotropic spectral estimate and it is computationally efficient since it requires only $O(BRL^2)$ operations where L is the number of points used to discretize the interval $[0, R]$, and where B is the bandwidth in the wave-number plane of the spectrum that we want to estimate. We believe

that this result is the most important contribution of this thesis. As mentioned earlier, all of the previous 2-D MEM algorithms involve solving a nonlinear constrained optimization problem because of the lack of useful spectral factorization theorems in the discrete 2-D case. Our method is based on a key spectral factorization result that we derive in an appendix to Chapter 6. Specifically, we show that any isotropic power spectrum corresponding to a covariance function that is zero outside of a finite disk $2R$ in the spatial domain has a non-symmetric half plane factorization where the spectral factors have a *finite* spatial support which is concentrated on a disk of radius R.

5. **Series expansions for isotropic random vector fields:**
We develop a new uncorrelated series expansion and a spectral representation for isotropic random vector fields. This representation is based on the decomposition of isotropic random vector fields into their solenoidal and potential parts which was studied earlier by Yaglom [93], Ito [27] and Ogura [54]. We believe that this series representation may help to extend all of the results of this thesis to the vector case.

1.3 THESIS OVERVIEW

This thesis is organized as follows:

In Chapter 2 we present an overview of the mathematical properties of 2-D isotropic scalar random fields. In particular, we focus our attention on Fourier series expansions of such fields in terms of the coordinate angle θ in a polar coordinate representation of the underlying plane. Fourier series expansions of 2-D isotropic random fields have the important property that the Fourier coefficients processes of different orders are uncorrelated. Such expansions will be used throughout this thesis to solve and understand a number of problems for 2-D isotropic random fields. In particular, they are used in Chapter 2 to develop new sampling theorems for isotropic random fields where the fields are sampled along circles rather than at a discrete set of points.

In Chapter 3, we introduce a non-causal internal differential model for a class of 2-D isotropic fields. The model that we consider is expressed in terms of the Laplacian operator, and it defines a class of isotropic fields which includes fields that can be represented as the output of rational 2-D filters driven by white noise. The smoothing problem for isotropic random fields in this class given noisy measurements over a disk of radius R is defined and is reduced to a countably infinite set of decoupled 1-D estimation problem. Two-point boundary value models are then developed to describe the 1-D estimation problems. Furthermore, we outline two solutions to the 1-D two-point boundary value smoothing problems and we study the asymptotic behavior of the 1-D filters that we introduce to solve the 1-D problems.

Following some motivation for considering the class of isotropic covariance functions that are equal to a weighted sum of cylindrical harmonics, we formulate in Chapter 4 the problem of retrieving cylindrical harmonics from the covariance of samples of the zeroth-order Fourier coefficient process. We then present a new algorithm for recovering the cylindrical harmonics by performing an eigenanalysis of the covariance matrix of samples of the zeroth order Fourier coefficient process corresponding to the measurements. Both the theory behind this algorithm and its numerical implementation are discussed. The proposed algorithm is very similar in spirit to the eigenstructure approach developed by Schmidt [73], and by Bienvenu and Kopp [9] for solving the 1-D harmonic retrieval problem. Furthermore, we develop a method for estimating the covariance matrix that is used as an input to our procedure. Both the statistical properties of the covariance estimate and its practical implementation are presented. In particular, our estimate is shown to be both unbiased and consistent, and an example is provided to demonstrate this fact. Finally, two examples are presented to illustrate the high resolution and robustness properties of our method.

In Chapter 5, we consider the maximum entropy method (MEM) spectral estimation problem for isotropic random fields. We begin by deriving an expression for the isotropic MEM estimate. The MEM spectral estimation problem is then related to the problem of finding the best linear filter for estimating an isotropic

field on the boundary of a disk given noisy observations of the field inside the disk. By using Fourier expansions of the optimal linear estimation filter and the efficient recursions of [41], a fast and robust method for computing the MEM estimate is developed. The numerical implementation of our procedure is also described in this chapter. Particular attention is given to the issues of numerical stability and convergence of our implementation. Finally, several examples are presented to illustrate the behavior of our algorithm and particularly to demonstrate its high resolution property.

In Chapter 6, we review some of the mathematical properties of isotropic random vector fields. Furthermore, we develop a new series expansion for such fields. This series expansion is seen to be the generalization to the vector case of the Fourier series expansions for isotropic scalar fields. We describe some recursive estimation and spectral estimation problems for isotropic random vector fields which we believe can be solved by using the series representation that we developed and by using some of the techniques that we introduced earlier in the thesis. We also discuss some open modeling problems for random vector fields.

Finally, in Chapter 7, we summarize our results, and suggest areas where future research may be fruitful.

Chapter 2

MATHEMATICAL THEORY OF ISOTROPIC RANDOM FIELDS

The objective of this chapter is to introduce the reader to some of the mathematical properties of 2-D isotropic random fields, and in particular to Fourier series expansions of such fields. Fourier series expansions of isotropic random fields will be used throughout this thesis to solve a variety of smoothing and spectral estimation problems for 2-D isotropic random fields in Chapters 3, 4 and 5. Specifically, these expansions will allow us to reduce the general 2-D problems that we address to a possibly countably infinite number of corresponding 1-D problems. In some cases, we will even be able to solve our problems by looking at a single 1-D problem by properly exploiting the properties of the Fourier coefficients corresponding to the underlying 2-D isotropic random field. The usefulness of Fourier series expansions is also illustrated in this chapter by the sampling theorems that we have developed for 2-D isotropic random fields and which we present in Section 2.2.

2.1 FOURIER SERIES FOR ISOTROPIC FIELDS

The covariance function

$$K(\vec{r}) = E[z(\vec{v})z(\vec{v}+\vec{r})] \tag{2.1.1}$$

of any zero-mean isotropic random field $z(\vec{r})$ [1], is a function of r only, so that, by abuse of notation we can write

$$K(\vec{r}) = K(r). \tag{2.1.2}$$

Since $K(r)$ is a function of r only, it is straightforward to show that the power spectrum $S(\vec{\lambda})$ of the field $z(\vec{r})$, i.e. the 2-D Fourier transform of $K(\vec{r})$, is actually a function of $\lambda = |\vec{\lambda}|$ only, and with a slight abuse of notation we will write this as $S(\lambda)$. This follows from the fact that

$$\begin{aligned} S(\vec{\lambda}) &= \int_{\mathbf{R}^2} K(r)\, e^{-j\vec{\lambda}\cdot\vec{r}}\, d\vec{r} \\ &= \int_0^{\infty}\int_0^{2\pi} K(r)\, e^{-j\lambda r\cos(\theta-\gamma)}\, r\, dr\, d\theta \\ &= 2\pi\int_0^{\infty} K(r) J_0(\lambda r)\, r\, dr \\ &= S(\lambda), \end{aligned} \tag{2.1.3}$$

where we have used the fact that

$$J_0(z) = \frac{1}{2\pi}\int_0^{2\pi} e^{jz\cos\theta}\, d\theta, \tag{2.1.4}$$

and where $\vec{\lambda} = (\lambda, \gamma)$. In (2.1.3) and (2.1.4) $J_0(\cdot)$ is the Bessel function of order 0. Observe that $S(\lambda)$ is nothing more than 2π times the Hankel transform [55] of $K(r)$ viewed as a function of the scalar $r = |\vec{r}|$. Equation (2.1.3) can also be inverted to yield

$$K(r) = \frac{1}{2\pi}\int_0^{\infty} S(\lambda) J_0(\lambda r)\, \lambda\, d\lambda. \tag{2.1.5}$$

Note that the covariance function

$$K(|\vec{r} - \vec{s}|) = K((r^2 + s^2 - 2rs\cos(\theta - \phi))^{1/2}) \tag{2.1.6}$$

is a positive-definite symmetric function of the scalar variables θ and ϕ where $\vec{r} = (r, \theta)$ and $\vec{s} = (s, \phi)$. Hence, if we assume that the random field $z(\cdot)$ has a finite energy, then the covariance function $K(\cdot)$ viewed as a function of the variables

[1] Throughout this thesis we use $\vec{r}$ to denote a point in 2-D Cartesian space. The polar coordinates of this point are denoted by r and θ. The Cartesian coordinates of $\vec{r}$ are denoted by r_1 and r_2.

θ and ϕ must have a symmetric eigenfunction expansion according to Mercer's theorem [63]. By using (2.1.5) and the addition theorem for Bessel functions [7],

$$J_0(|\vec{r}-\vec{s}|) = \sum_{n=-\infty}^{\infty} J_n(\lambda r) J_n(\lambda s)\, e^{jn(\theta-\phi)} \tag{2.1.7}$$

where $\vec{r} = (r,\theta)$ and $\vec{s} = (s,\phi)$, we can write

$$K(|\vec{r}-\vec{s}|) = \sum_{n=-\infty}^{\infty} k_n(r,s) e^{jn(\theta-\phi)}, \tag{2.1.8}$$

where

$$k_n(r,s) = \frac{1}{2\pi}\int_0^{\infty} J_n(\lambda r) J_n(\lambda s) S(\lambda)\lambda\, d\lambda, \tag{2.1.9}$$

In (2.1.9) $J_n(\cdot)$ is the Bessel function of order n. Alternatively, $k_n(r,s)$ can be computed from $K(\cdot)$ as

$$k_n(r,s) = \frac{1}{2\pi}\int_0^{2\pi} K((r^2+s^2-2rs\cos\theta)^{1/2})\, e^{-jn\theta}\, d\theta. \tag{2.1.10}$$

Note that since $K(\cdot)$ is a real and even function of θ, then

$$k_n(r,s) = k_{-n}(r,s). \tag{2.1.11}$$

Alternatively, (2.1.11) can be derived by using (2.1.9) and the fact that

$$J_n(x) = (-1)^n J_{-n}(x). \tag{2.1.12}$$

Note also that both the covariance function $K(r)$ of the full process $z(\vec{r})$ and its power spectrum $S(\lambda)$ can be recovered exactly from $k_0(r,s)$ as follows. By specializing (2.1.10) to the case $n=0$ we obtain for $s=0$

$$K(r) = k_0(r,0). \tag{2.1.13}$$

Furthermore, (2.1.9) implies that $S(\lambda)$ can be recovered from $k_n(r,s)$ in general by taking the nth order Hankel transform [55] of $k_n(r,s)$ with respect to the variable r and dividing by $J_n(\lambda s)/2\pi$, i.e.

$$S(\lambda) = 2\pi \int_0^{\infty} J_n(\lambda r) k_n(r,s)\, r\, dr \,/\, J_n(\lambda s). \tag{2.1.14}$$

Thus, if we are interested in extracting information from $K(r)$, or from its associated power spectrum $S(\lambda)$, we can as well focus our attention on the covariance $k_0(r,s)$ without any loss of information. This important observation is the key to the high resolution spectral estimation procedure that we present in Chapter 4.

By using (2.1.3) and the Karhunen-Loeve theorem [81] we can expand $z(\vec{r})$ as [93], [92], [88]

$$z(\vec{r}) = \sum_{n=-\infty}^{\infty} z_n(r)\, e^{jn\theta}, \tag{2.1.15}$$

$$z_n(r) = \frac{1}{2\pi}\int_0^{2\pi} z(\vec{r}) e^{-jn\theta}\, d\theta, \tag{2.1.16}$$

where

$$E[z_n(r) z_m(s)] = k_n(r,s)\delta_{n,m}, \tag{2.1.17}$$

and where $\delta_{n,m}$ is a Kronecker delta function. Note that equations (2.1.9) and (2.1.17) imply that although $z(\vec{r})$ is isotropic, $z_n(r)$ is *not* a stationary process, i.e. $k_n(r,s)$ is *not* a function of $r-s$. Equation (2.1.15) is very interesting since it can also be interpreted as a Fourier series expansion of the field $z(\vec{r})$ in terms of the coordinate angle θ. In particular, (2.1.15)-(2.1.17) state that the Fourier coefficients $z_n(r)$ in a Fourier series expansion of $z(\vec{r})$ in terms of the angle θ are *independent*. This observation plays a key role in a number of works dealing with isotropic random fields (e.g. [92],[41]) and we shall use it throughout this thesis to solve a variety of smoothing and spectral estimation problems for isotropic random fields.

Finally, note that a spectral representation for the Fourier coefficient processes $z_n(r)$ has been developed in [93] and [92]. Specifically, it is shown in [93] and [92] that $z_n(r)$ has the representation

$$z_n(r) = \int_0^{\infty} J_n(\lambda r) Z_n(d\lambda) \tag{2.1.18}$$

where the $\{Z_n(d\lambda)\}$ are uncorrelated orthogonal random measures satisfying

$$E[Z_n(d\lambda)] = 0 \tag{2.1.19}$$

$$E[Z_n(d\lambda_1) Z_m^*(d\lambda_2)] = \delta(\lambda_1-\lambda_2) S(\lambda_1)\, \lambda_1\, d\lambda_1\, d\lambda_2\, \delta_{n,m}, \tag{2.1.20}$$

and where $Z_m^*(\cdot)$ denotes the complex conjugate of $Z_m(\cdot)$. Hence, by combining (2.1.15) and (2.1.18)-(2.1.20) we find that the isotropic random field $z(\vec{r})$ has the spectral representation

$$z(\vec{r}) = \sum_{n=-\infty}^{\infty} e^{jn\theta} \int_0^{\infty} J_n(\lambda r) Z_n(d\lambda). \tag{2.1.21}$$

2.2 RESTRICTION OF 2-D ISOTROPIC FIELDS TO A LINE

The restriction of a 2-D isotropic random field to a line is clearly a 1-D stationary random process. Similarly, the restriction of a 3-D isotropic random field to a line is also a 1-D stationary random process. In this short section, we derive a relationship between the 1-D power spectrum $S_1(\lambda_1)$ of the restriction of a 2-D isotropic random field to a line L and the 2-D power spectrum $S(\lambda)$ of the isotropic random field. Without loss of generality, let us assume that the line L coincides with the r_1 axis. Accordingly, we have

$$\begin{aligned} S_1(\lambda_1) &= \int_{-\infty}^{\infty} K(r_1) e^{-j\lambda_1 r_1}\, dr_1 \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\lambda)\, d\lambda_2 \\ &= \frac{1}{\pi} \int_{\lambda_1}^{\infty} \frac{\lambda S(\lambda)}{\sqrt{\lambda^2 - \lambda_1^2}}\, d\lambda, \end{aligned} \tag{2.2.1}$$

and where $\lambda = (\lambda_1, \lambda_2)$ in a Cartesian representation of the wave-number plane. To obtain $S(\lambda)$ from $S_1(\lambda_1)$, we note that (2.2.1) is an Abel integral equation [80]. Hence, by using known results concerning these equations we find that [10]

$$S(\lambda) = -2 \int_{\lambda}^{\infty} \frac{d}{d\lambda_1} [S_1(\lambda_1)] \frac{d\lambda_1}{\sqrt{\lambda_1^2 - \lambda^2}}. \tag{2.2.2}$$

The fact that a the restriction of a 2-D isotropic random field to a line is a stationary random process suggests a different approach than the one we have taken

in this thesis for solving problems involving 2-D isotropic random fields. Specifically, one might think of using a generalization of the 1-D techniques that were developed for stationary random processes to solve problems involving 2-D isotropic random fields. For example, one may try to use 1-D spectral estimation techniques to estimate $S_1(\lambda_1)$ and then use (2.2.2) to compute an estimate of $S(\lambda)$. In theory, such an approach is possible provided that one makes sure that all the constraints implied by (2.2.1)-(2.2.2) are satisfied. The constraints implied by (2.2.1)-(2.2.2) are easier to visualize in the 3-D case where (2.2.1)-(2.2.2) take the form [10]

$$S_1(\lambda_1) = \int_{\lambda_1}^{\infty} \lambda S(\lambda)\, d\lambda \tag{2.2.3}$$

$$S(\lambda) = -\frac{1}{\lambda}\frac{d}{d\lambda} S_1(\lambda). \tag{2.2.4}$$

In (2.2.3)-(2.2.4) $S(\lambda)$ and $S_1(\lambda_1)$ are respectively the 3-D power spectrum of a 3-D isotropic random field and the 1-D power spectrum of the restriction of the same 3-D isotropic random field to a line. In particular, note that since $S(\lambda)$ is nonnegative then $S_1(\lambda_1)$ is a nonincreasing function of $|\lambda_1|$. Hence, if one uses 1-D spectral estimation techniques to compute an estimate of $S_1(\lambda_1)$, one has to make sure that the computed 1-D spectral estimate is nonincreasing for all $|\lambda_1|$.

2.3 SAMPLING THEOREMS

In this section, we develop two different procedures for sampling and reconstructing the 1-D *Fourier coefficient processes* associated with a given isotropic random field. Using the sampling theorems for the Fourier processes, we show that a wave-number limited isotropic random field can be reconstructed from its observations on a countably infinite number of concentric *circles* with a vanishing mean-square error. In [57], Petersen and Middleton extended the 1-D Shannon sampling theorem to m-dimensional Euclidean spaces. In particular, they developed efficient *point* sampling and reconstruction techniques for wave-number limited homogeneous random fields which minimize the number of sample points required per unit area to reconstruct a given field in the sense of a vanishing mean-square

error. Here, by contrast, we present a new reconstruction procedure for 2-D wave-number-limited isotropic random fields sampled along circles in the Euclidean plane, where the reconstruction is to be understood in a mean-square sense.

A Sampling the Fourier processes covariance functions

Let us begin by presenting two different sampling procedures for the covariance function of the nth order Fourier coefficient process.

Theorem 2.1 *The nth order Fourier coefficient process covariance function $k_n(r,s)$ of an isotropic random field $z(\vec{r})$ whose spectral density function $S(\lambda)$ is wave-number limited to the region $\lambda < B$, can be reconstructed exactly from the sample values of the mth order Fourier coefficient process covariance function $k_m(r,s)$ taken over a lattice of points $\{(a_{i,m}/B, a_{j,m}/B);\ i,j \in \mathbf{N}\}$, where $a_{i,m}$ is the ith zero of the mth order Bessel function $J_m(x)$.*

Proof

Over the interval $0 < \lambda < B$, $J_n(\lambda r)$ can be expanded into a Fourier-Bessel series of the form [30]

$$J_n(\lambda r) = \sum_{i=1}^{\infty} b_i(r) J_m(\frac{a_{i,m}\lambda}{B}),\ 0 < \lambda < B \tag{2.3.1}$$

where

$$b_i(r) = \frac{2\int_0^B J_n(\lambda r) J_m(\frac{a_{i,m}\lambda}{B})\lambda\, d\lambda}{B^2 J_{m+1}(a_{i,m}) J_{m-1}(a_{i,m})}, \tag{2.3.2}$$

Substituting (2.3.1) for $J_n(\lambda r)$ and $J_n(\lambda s)$ into (2.1.9) yields the desired result

$$k_n(r,s) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} b_i(r) b_j(s) k_m(\frac{a_{i,m}}{B}, \frac{a_{j,m}}{B}). \tag{2.3.3}$$

Note that, according to Theorem 2.1, it is possible to reconstruct the covariance functions of the Fourier coefficient processes of *all* orders given sampled values of the covariance function of a *single* Fourier coefficient process of *any* order. This should not come as a surprise: if one can reconstruct $k_m(r,s)$ exactly from its sample values on the grid $\{(\frac{a_{i,m}}{B}, \frac{a_{j,m}}{B});\ \ i,j \in \mathbf{N}\}$ then one can easily compute $S(\lambda)$ (e.g. by

taking the mth order Hankel transform of $k_m(r,s)$ with respect to r and dividing by $J_m(\lambda s)$). Given $S(\lambda)$, one can then evaluate $k_n(r,s)$ for all n through (2.1.9). We now state and prove a second sampling theorem for the covariance function $k_n(r,s)$ of the nth order Fourier coefficient process. In this case we use samples of $k_n(r,s)$, rather than of $k_m(r,s)$.

Theorem 2.2 *The nth order Fourier coefficient process covariance function $k_n(r,s)$ of an isotropic random field $z(\vec{r})$ whose spectral density function $S(\lambda)$ is wave-number limited to the region $\lambda < B$, can be reconstructed exactly from its own sample values taken over a lattice of points $\{i\pi/B, j\pi/B);\ i,j \in \mathbf{N}\}$, or over a lattice of points $\{(a_{i,m}/B, a_{j,m}/B);\ i,j \in \mathbf{N}\}$, where $a_{i,m}$ is the ith zero of the mth order Bessel function $J_m(x)$.*

Proof

Consider the identity (see Appendix 2.A)

$$J_n(\lambda r) = \sum_{i=0}^{\infty} d_i^n(r) J_n(\frac{i\pi}{B}\lambda),\ 0 < \lambda < B \tag{2.3.4}$$

where

$$d_i^n(r) = \frac{1}{(1+\delta_{0,n})}((-1)^n \frac{sin(B(r+\frac{i\pi}{B}))}{B(r+\frac{i\pi}{B})} + \frac{sin(B(r-\frac{i\pi}{B}))}{B(r-\frac{i\pi}{B})}). \tag{2.3.5}$$

Substituting (2.3.4) into (2.1.9) yields

$$k_n(r,s) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} d_i^n(r) d_j^n(s) k_n(\frac{i\pi}{B}, \frac{j\pi}{B}). \tag{2.3.6}$$

Similarly, by substituting the identity (see Appendix 2.A)

$$J_n(\lambda r) = \sum_{i=1}^{\infty} c_{i,m}^n(r) J_n(\frac{a_{i,m}}{B}\lambda),\ 0 < \lambda < B \tag{2.3.7}$$

where

$$c_{i,m}^n(r) = \frac{2a_{i,m}J_m(Br)}{(a_{i,m}^2 - B^2r^2)J_{m+1}(a_{i,m})}(\frac{a_{i,m}}{r})^{|m-n|}, \tag{2.3.8}$$

into (2.1.9) we obtain

$$k_n(r,s) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} c_{i,m}^n(r)c_{j,m}^n(s)k_n(\frac{a_{i,m}}{B},\frac{a_{j,m}}{B}). \tag{2.3.9}$$

□ □

Observe that Theorem 2.2 asserts that the same sampling grid can be used for *all* the Fourier coefficient process covariance functions. The sampling grid can be selected to be $\{(i\pi/B, j\pi/B)\}$ or $\{(a_{i,m}/B, a_{j,m}/B)\}$ where m is fixed but arbitrary. This fact will prove useful in deriving sampling theorems for isotropic random fields.

B Sampling Isotropic Fields

Theorem 2.2 can now be used to prove the following important result.

Theorem 2.3 *The nth order Fourier coefficient process $z_n(r)$ corresponding to an isotropic random field $z(\vec{r})$ with a wave-number limited spectrum $S(\lambda)$, where $S(\lambda) = 0$ for $\lambda > B$, can be reconstructed with zero mean-square error from its samples $\{z_n(\frac{a_{i,m}}{B});\quad i \in \mathbf{N}\}$, where $a_{i,m}$ is the ith zero of the mth order Bessel function $J_m(x)$ as,*

$$z_n(r) = \sum_{i=1}^{\infty} c_{i,m}^n(r) z_n(\frac{a_{i,m}}{B}), \tag{2.3.10}$$

and from its samples $\{z_n(\frac{i\pi}{B});\ i \in \mathbf{N}\}$ as,

$$z_n(r) = \sum_{i=0}^{\infty} d_i^n(r) z_n(\frac{i\pi}{B}), \tag{2.3.11}$$

where (2.3.10) and (2.3.11) hold in a mean-square sense, and where $d_i^n(r)$ and $c_{i,m}^n(r)$ are defined respectively by (2.3.5) and (2.3.8).

Proof

Let

$$\hat{z}_n^B(r) = \sum_{i=1}^{\infty} c_{i,m}^n(r) z_n(\frac{a_{i,m}}{B}). \tag{2.3.12}$$

From (2.3.1) and (2.1.9), it can be shown that

$$k_n(r,s) = \sum_{i=1}^{\infty} c_{i,m}^n(s) k_n(r,\frac{a_{i,m}}{B}), \tag{2.3.13}$$

and that

$$k_n(\frac{a_{i,m}}{B}, r) = \sum_{j=1}^{\infty} c_{j,m}^n(r) k_n(\frac{a_{i,m}}{B}, \frac{a_{j,m}}{B}). \tag{2.3.14}$$

Using the above two equations we obtain

$$\begin{aligned} E[z_n(r)(z_n(r) - \hat{z}_n^B(r))] &= k_n(r,r) - \sum_{i=1}^{\infty} c_{i,m}^n(r) k_n(r, \frac{a_{i,m}}{B}) \\ &= 0, \end{aligned} \tag{2.3.15}$$

and

$$\begin{aligned} E[z_n(\frac{a_{i,m}}{B})(z_n(r) - \hat{z}_n^B(r))] &= k_n(\frac{a_{i,m}}{B}, r) - \sum_{i=1}^{\infty} c_{i,m}^n(r) k_n(\frac{a_{i,m}}{B}, \frac{a_{j,m}}{B}) \\ &= 0. \end{aligned} \tag{2.3.16}$$

Combining (2.3.15) with (2.3.16), it follows that

$$E[|z_n(r) - \hat{z}_n^B(r)|^2] = 0. \tag{2.3.17}$$

By using a similar approach it can be shown that

$$E[|z_n(r) - \hat{z}_n^F(r)|^2] = 0, \tag{2.3.18}$$

where

$$\hat{z}_n^F(r) = \sum_{i=0}^{\infty} d_i^n(r) z_n(\frac{i\pi}{B}). \tag{2.3.19}$$

□ □

Now recall that the knowledge of $z(\vec{r})$ on a circle of radius r' is sufficient to compute *all* of the Fourier coefficient processes $z_n(r)$ at the location $r = r'$. Hence, we have the following important result.

Theorem 2.4 *Any isotropic random field $z(\vec{r})$ with a wave-number limited spectrum $S(\lambda)$, where $S(\lambda) = 0$ for $\lambda > B$, can be reconstructed with zero mean-square error from its samples on the countable set of circles of radii $r_i = \frac{a_{i,m}}{B}$, $i \in \mathbf{N}$, where $a_{i,m}$ is the ith zero of the mth order Bessel function $J_m(x)$, as*

$$z(\vec{r}) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \sum_{i=1}^{\infty} c_{i,m}^n(r) \int_0^{2\pi} z(\frac{a_{i,m}}{B}, \phi) e^{-jn\phi} \, d\phi \, e^{jn\theta}, \tag{2.3.20}$$

and from its samples on the countable number of circles of radii $r_i = \frac{i\pi}{B}$, $i \in \mathbf{N}$, *as*

$$z(\vec{r}) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \sum_{i=0}^{\infty} d_i^n(r) \int_0^{2\pi} z(\frac{i\pi}{B}, \phi) e^{-jn\phi}\, d\phi\; e^{jn\theta}, \qquad (2.3.21)$$

where equations (2.3.20) and (2.3.21) are to be understood in a mean-square sense.

Theorem 2.4 follows directly from Theorem 2.3 and equation (2.1.15), and is a generalization of a result in [92]. A natural question to be asked here, is which of the above two sampling schemes —i.e. the *Bessel sampling scheme* involving sampling on circles of radii $r_i = a_{i,m}/B$, or the *uniform sampling scheme* using circles of radii $r_i = i\pi/B$— is more efficient in terms of minimizing the number of sampling circles per unit radial length? This leads us to examine the distribution of the zeros $a_{i,m}$ of the mth order Bessel function $J_m(x)$, along the positive real axis. For large i, and a fixed value m, the zeros of the mth order Bessel function are approximately given by the McMahon expansion [1]

$$a_{i,m} \approx (i + \frac{1}{2}m - \frac{1}{4})\pi - \frac{(4m^2 - 1)}{8\pi(i + \frac{1}{2}m - \frac{1}{4})} \cdots, \qquad (2.3.22)$$

which shows that the separation $\Delta_{i,m}$ between two successive large zeros $a_{i+1,m}$ and $a_{i,m}$ of $J_m(x)$, with $i >> m$, is approximately equal to

$$\Delta_{i,m} = a_{i+1,m} - a_{i,m} \approx \pi + \frac{(4m^2 - 1)}{8i^2} + O(i^{-3}). \qquad (2.3.23)$$

In particular, two successive large zeros of $J_0(x)$ are separated by a distance slightly less than π, whereas two successive large zeros of $J_m(x)$ for $m \neq 0$, are separated by a distance slightly larger than π. As the order i of the zeros $a_{i,m}$ of $J_m(x)$ tends to infinity the separation $\Delta_{i,m}$ between successive zeros is asymptotically equal to π, for all m. Furthermore, examination of the small zeros of the mth order Bessel function reveals that even for $i = 2$, $\Delta_{i,m}$ is approximately equal to π. Hence, the Bessel sampling scheme is slightly more efficient than the uniform sampling scheme if the zeros of a large order Bessel function are used to generate the nonuniform circular sampling grid. However, the Bessel sampling scheme is primarily of theoretical interest, while the uniform sampling scheme is of more practical value since it does not require the knowledge of a large number of zeros of one of the Bessel functions.

Finally, observe that in practice one does not need to sample the field $z(\vec{r})$ continuously as a function of θ along any of the circles r_i. Note that along any of these circles $z(r_i, \theta)$ is a stationary process with the covariance function

$$\begin{aligned} K(r_i; \theta, \phi) &= E[z(r_i, \theta)(z(r_i, \phi)] \\ &= \sum_{n=-\infty}^{\infty} k_n(r_i, r_i)\, e^{jn(\theta-\phi)}. \end{aligned} \tag{2.3.24}$$

Examination of a plot of $J_n(x)$ [1] reveals that

$$J_n(x) \approx 0 \quad \text{for} \quad x \gg 1 \quad \text{and} \quad n > x. \tag{2.3.25}$$

Hence, by using (2.1.9) and the Lebesgue dominated convergence theorem [72] to interchange the operations of limit and integration, we obtain

$$k_n(r_i, r_i) \approx 0 \quad \text{for} \quad Br_i \gg 1 \quad \text{and} \quad n > Br_i. \tag{2.3.26}$$

Equation (2.3.26) implies that along any circle of radius r_i, $z(r_i, \theta)$ can be approximated with a small mean square error by the finite sum

$$z(r_i, \theta) = \sum_{n=-N}^{N} z_n(r_i)\, e^{jn\theta} \tag{2.3.27}$$

where $N \geq Br_i$. In particular, (2.3.20)and (2.3.21) can be approximated in the mean square sense as

$$z(\vec{r}) = \sum_{i=1}^{\infty} \sum_{n=-N_i}^{N_i} z_n(r_i)\, e^{jn\theta} \tag{2.3.28}$$

where

$$r_i = \frac{i\pi}{B} \qquad \text{and} \qquad N_i \gg i\pi \tag{2.3.29}$$

or

$$r_i = \frac{a_{i,m}}{B} \qquad \text{and} \qquad N_i \gg a_{i,m}. \tag{2.3.30}$$

The coefficients $z_n(r_i)$, $-N_i \leq n \leq N_i$, can be determined by sampling $z(r_i, \theta)$ at $2N_i + 1$ points.

2.4 CONCLUSION

In this chapter we have described the Fourier expansions for 2-D isotropic random fields. We have illustrated the usefulness of these expansions with some new sampling theorems for 2-D isotropic random fields, where the fields are sampled along concentric circles rather than at discrete points. The true power of these expansions will become clear in the next three chapters in which we use them to reduce a number of 2-D smoothing and spectral estimation problems to a possibly countably infinite number of corresponding 1-D problems. The results of this chapter, and indeed the whole analysis developed in this thesis, can easily be extended to isotropic random fields in higher dimensions. In the general m-D case, isotropic random fields can be expanded in terms of spherical harmonics [93], [92], [88] rather than in Fourier series. In fact, expansions similar to (2.1.15) can be developed for homogeneous fields $z(\vec{r})$ defined over a homogeneous space $\mathcal{R}$ provided that the covariance function of such fields satisfies a condition of the form

$$K(\vec{r},\vec{r'}) = K(g\vec{r},g\vec{r'}) \tag{2.4.31}$$

where g belongs to an arbitrary group of motions G (see [95] and [94] for more details). Such expansions provide a means for generalizing most of our results to such fields.

APPENDIX 2.A

Proof of (2.3.4)

Let $p_B(\lambda)$ be the function

$$p_B(\lambda) = \begin{cases} 1 & \lambda < B \\ 0 & \text{otherwise.} \end{cases} \tag{2.A.1}$$

By using the identity [20], p. 43 and p. 99

$$\int_{-\infty}^{\infty} J_n(\lambda r) e^{-jur}\, dr = \begin{cases} 2(-j)^n \frac{T_n(\frac{u}{\lambda})}{(\lambda^2-u^2)^{1/2}} & 0 < |u| < \lambda \\ 0 & \text{otherwise,} \end{cases} \tag{2.A.2}$$

where $T_n(x)$ is a Chebyshev polynomial of order n, it can be shown that the Fourier transform of $J_n(\lambda r)p_B(\lambda)$ with respect to r is bandlimited to B radians per unit distance. Hence, $J_n(\lambda r)p_B(\lambda)$ can be written as [30]

$$J_n(\lambda r)p_B(\lambda) = \sum_{i=0}^{\infty} d_i^n(r) J_n(\frac{i\pi}{B}\lambda),\ 0 < \lambda < B \tag{2.A.3}$$

where

$$d_i^n(r) = \frac{1}{(1+\delta_{0,n})}((-1)^n \frac{sin(B(r+\frac{i\pi}{B}))}{B(r+\frac{i\pi}{B})} + \frac{sin(B(r-\frac{i\pi}{B}))}{B(r-\frac{i\pi}{B})}), \tag{2.A.4}$$

and where $\delta_{0,n}$ denotes the Kronecker delta function.

Proof of (2.3.7)

Consider the identity [7], p. 72

$$J_l(\lambda r) = \sum_{i=1}^{\infty} c_{i,l}(r) J_l(\frac{a_{i,l}}{B}\lambda),\ 0 < \lambda < B \tag{2.A.5}$$

where

$$c_{i,l}(r) = \frac{2a_{i,l}J_l(Br)}{(a_{i,l}^2 - B^2r^2)J_{l+1}(a_{i,l})}. \tag{2.A.6}$$

By repeatedly differentiating both sides of (2.A.5) with respect to λ, and by using the identity

$$\frac{d}{d\lambda}J_m(\lambda r) = rJ_{m-1}(\lambda r) - \frac{m}{\lambda}J_m(\lambda r) \tag{2.A.7}$$

for the case where $n < l$, and the identity

$$\frac{d}{d\lambda}J_m(\lambda r) = -rJ_{m+1}(\lambda r) + \frac{m}{\lambda}J_m(\lambda r) \tag{2.A.8}$$

for the case where $n > l$, we obtain

$$J_n(\lambda r) = \sum_{i=1}^{\infty} c_{i,l}^n(r)J_n(\frac{a_{i,l}}{B}\lambda),\ 0 < \lambda < B \tag{2.A.9}$$

where

$$c_{i,l}^n(r) = c_{i,l}(r)(\frac{a_{i,l}}{r})^{|l-n|}. \tag{2.A.10}$$

Chapter 3

RECURSIVE ESTIMATION FOR A CLASS OF ISOTROPIC RANDOM FIELDS

As mentioned in Chapter 1, problems involving spatially-distributed data and phenomena arise in various fields including image processing, meteorology, geophysical signal processing, oceanography and optical processing. A major challenge in any such problem is to develop algorithms capable of dealing effectively with the increased computational complexity of multidimensional problems and which can be implemented in a recursive fashion. In one dimension the ways in which data can be organized for efficient processing are extremely limited and causality typically provides a natural choice. Furthermore, in one dimension, internal differential realizations of random processes were exploited to develop an efficient estimation algorithm,namely the Kalman filtering technique. This has led researchers in estimation theory to investigate the extension of 1-D Kalman filtering and smoothing methods to non-causal 2-D random fields. However, all of the estimation algorithms developed in higher dimensions are either approximate or can be applied only to a limited class of 2-D fields, namely to fields that can be described by hyperbolic partial differential equations, and which therefore are causal in some sense.

The objective of this chapter is to study the smoothing problem for a class of random fields which have non-causal internal differential realizations but which

also have enough structure to allow the development of efficient recursive smoothing algorithms. Note that, unlike one dimension, the most natural estimation problem in higher dimensions is the smoothing problem, rather than the causal filtering problem. This is due to the fact that in higher dimensions, the filtering problem requires an artificial partition of the data between past and future, whereas the smoothing problem does not assume any causal ordering of the data. Specifically, in this chapter we investigate efficient recursive smoothing techniques for isotropic random fields which can be represented as the output of rational 2-D filters driven by white noise, and which admit therefore simple internal differential models. The smoothing problem for 2-D isotropic random fields has been studied from an input-output point of view by Levy and Tsitsiklis [41] who used Fourier expansions of these fields to develop efficient Levinson-like recursions for the smoothing filter. Here, by contrast, we consider the smoothing problem for isotropic random fields $z(\vec{r})$ having an internal differential realization involving the Laplacian operator. The motivation for studying a model of this form is that any isotropic process that can be obtained by passing 2-D white noise through a rational linear filter has an internal realization of this type (see Section 3.1). Another motivation for considering such a model is that it can be used to describe a large class of physical phenomena such as the variation of the electric potential created by a random charge distribution. Given noisy observations of the isotropic random field $z(\vec{r})$ over a finite disk of radius R , our approach is to reduce the 2-D smoothing problem to a countable set of decoupled 1-D smoothing problems for the uncorrelated Fourier coefficient processes $z_k(r)$ corresponding to the process $z(\vec{r})$. Using the internal model of the process $z(\vec{r})$, we construct 1-D state space two-point boundary value (TPBV) models for the Fourier coefficient processes . We then solve the resulting 1-D TPBV smoothing problems using either a Markovianization technique which transforms the non-causal model to a causal one to which standard 1-D smoothing techniques can be applied, or directly by using the method of Adams et al. [2]. Finally, we obtain the best linear least squares estimate of $z(\vec{r})$ given the observations as a properly weighted combination of the 1-D smoothed estimates of all the Fourier coefficient processes $z_k(r)$. Observe that by properly exploiting the structure of

isotropic random fields, a *recursive* solution to the smoothing problem for a non-causal isotropic process has thus been constructed . The recursions here are with respect to the radius r in a polar coordinate representation of the fields.

3.1 INTERNAL MODEL

A Differential Model

The class of random fields considered in this chapter is described over the plane $\mathbf{R}^2$ by the differential model

$$(I_n\nabla^2 - A^2)x(\vec{r}) = Bu(\vec{r}) \tag{3.1.1}$$

$$z(\vec{r}) = Cx(\vec{r}) \tag{3.1.2}$$

with the asymptotic condition

$$E[x(\vec{r})x^T(\vec{s})] \to 0 \quad as \quad |\vec{r}-\vec{s}| \to \infty. \tag{3.1.3}$$

Here, $x(\vec{r}) \in \mathbf{R}^n$, $u(\vec{r}) \in \mathbf{R}^m$, $z(\vec{r}) \in \mathbf{R}^p$, and A , B, and C are real matrices of appropriate dimensions. The eigenvalues of the matrix A are assumed to have strictly positive real parts. This assumption insures that there exists a solution $x(\vec{r})$ to (3.1.1) that obeys the asymptotic condition (3.1.3). In equation (3.1.1) $u(\vec{r})$ is a random zero-mean two-dimensional white Gaussian noise process with

$$E[u(\vec{r})u^T(\vec{s})] = I_m\delta(\vec{r}-\vec{s}), \tag{3.1.4}$$

where I_m is the $m \times m$ identity matrix. The class of random fields that can be modeled by (3.1.1)-(3.1.2) includes a large number of physical phenomena. Furthermore, as we shall show latter, this class includes also all isotropic random fields which can be represented as the output of rational 2-D filters driven by white noise.

From an input-output point of view equations (3.1.1)-(3.1.2) together with the asymptotic condition (3.1.3) are equivalent to the representation

$$x(\vec{r}) = -\frac{1}{2\pi}\int_{\mathbf{R}^2} K_0(A|\vec{r}-\vec{r'}|)Bu(\vec{r'})\,d\vec{r'}, \quad \vec{r} \in \mathbf{R}^2 \tag{3.1.5}$$

$$z(\vec{r}) = Cx(\vec{r}) \tag{3.1.6}$$

where $d\vec{r}' = dx'\,dy'$ denotes an element of area. Here, $K_0(Ar)$ denotes a matrix modified Bessel function of the second kind and of order zero [7]. In fact, $G(\vec{r},\vec{s}) = \frac{1}{2\pi}K_0(A|\vec{r}-\vec{s}|)$ is the Green's function associated to the differential equation (3.1.1), i. e. $G(\vec{r},\vec{s})$ satisfies the equation

$$(I_n\nabla^2 - A^2)G(\vec{r},\vec{s}) = -I_n\delta(\vec{r}-\vec{s}) \tag{3.1.7}$$

for $\vec{r},\vec{s} \in \mathbf{R}^2$, with the asymptotic condition

$$G(\vec{r},\vec{s}) \to 0 \quad as \quad |\vec{r}-\vec{s}| \to \infty. \tag{3.1.8}$$

Matrix modified Bessel functions of the first and second kinds arise naturally in the study of rational isotropic random fields. A brief discussion of some of their properties appears in Appendix 3.A. (For more details see [7] and the references therein).

The main property of the process $x(\vec{r})$ defined by (3.1.1) and (3.1.3) is that it is a 2-D rational isotropic random field as is shown below.

Theorem 3.1 *The process $x(\vec{r})$ defined by equation (3.1.1) together with the asymptotic condition (3.1.3)is an isotropic random field, i.e. its autocorrelation function $R_x(\vec{r},\vec{s}) = E[x(\vec{r})x^T(\vec{s})]$ is invariant under translations and rotations.*

Proof

We will first show that $R_x(\vec{r},\vec{s})$ is invariant under translation. From (3.1.5) we have

$$R_x(\vec{r},\vec{s}) = E[x(\vec{r})x^T(\vec{s})] \tag{3.1.9}$$

$$= \frac{1}{4\pi^2}\int_{\mathbf{R}^2} K_0(A|\vec{r}-\vec{u}|)BB^TK_0(A|\vec{s}-\vec{u}|)\,d\vec{u} \tag{3.1.10}$$

Now perform the transformation

$$\vec{v} = \vec{u} + \vec{h} \tag{3.1.11}$$

to obtain

$$R_x(\vec{r},\vec{s}) = \frac{1}{4\pi^2}\int_{\mathbf{R}^2} K_0(A|\vec{r}+\vec{h}-\vec{v}|)BB^TK_0(A|\vec{s}+\vec{h}-\vec{v}|)\,d\vec{v}. \tag{3.1.12}$$

This shows that $R_x(\vec{r},\vec{s})$ is invariant under translation. Using this fact, we can write

$$R_x(\vec{r},\vec{s}) = R_x(\vec{v},0) \tag{3.1.13}$$

where $\vec{v} = \vec{r} - \vec{s}$. Hence,

$$\begin{aligned} R_x(\vec{r},\vec{s}) &= \frac{1}{4\pi^2}\int_{\mathbf{R}^2} K_0(A|\vec{v}-\vec{u}|)BB^T K_0(A|\vec{u}|)\,d\vec{u} && (3.1.14)\\ &= \frac{1}{4\pi^2}\int_{\mathbf{R}^2} K_0(A(v^2+u^2-2uv\cos(\theta-\phi))^{\frac{1}{2}})BB^T K_0(Au)\,d\vec{u} && (3.1.15)\end{aligned}$$

where $\vec{v} = (v,\phi)$ and $\vec{u} = (u,\theta)$. Letting $\alpha = \phi - \theta$, we conclude from the above equation and the periodicity of $\cos\alpha$ that

$$R_x(\vec{r},\vec{s}) = \frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{\infty} K_0(A(v^2+u^2-2uv\cos\alpha)^{\frac{1}{2}})BB^T K_0(Au)u\,du\,d\alpha \tag{3.1.16}$$

□□

Theorem (3.1) implies also that the output process $z(\vec{r})$ is isotropic with auto-correlation function

$$R_z(\vec{r},\vec{s}) = CR_x(\vec{r},\vec{s})C^T. \tag{3.1.17}$$

Since $R_x(\cdot)$ is translation-invariant we can define its spectral density matrix $S_x(\vec{\lambda})$, which is the 2-D Fourier transform of $R_x(\vec{r})$:

$$\begin{aligned} S_x(\vec{\lambda}) &= \int_{\mathbf{R}^2} R_x(\vec{r})e^{-j\vec{\lambda}\cdot\vec{r}}\,d\vec{r} && (3.1.18)\\ &= 2\pi\int_0^{\infty} R_x(r)J_0(\lambda r)r\,dr && (3.1.19)\\ &= (\lambda^2 I_n + M)^{-1}BB^T(\lambda^2 I_n + M^T)^{-1} && (3.1.20)\\ &= S_x(\lambda), && (3.1.21)\end{aligned}$$

where we have taken advantage of the circular symmetry of $R_x(\vec{r})$,and where $M = A^2$. Here M^T denotes the transpose of M. Observe that $S_x(\lambda)$ is *rational* in λ, the magnitude of $\vec{\lambda}$. Furthermore, the poles of the spectrum $S_x(\lambda)$, obtained by setting $p = j\lambda$ in (3.1.20), have a quadrantal symmetry property when plotted in the complex p-plane. Another important property of the process $x(\vec{r})$ that follows

from equation (3.1.20) is that $x(\vec{r})$ is pseudo-Markovian of order 1 [59], i.e. the value of $x(\vec{r})$ inside a closed curve Γ is independent of the value of $x(\vec{r})$ outside Γ given the value of $x(\vec{r})$ and of its normal derivative along Γ. In the sequel, we shall make extensive use of the isotropic and pseudo-Markovian nature of the process $x(\vec{r})$.

B Motivation

The motivation for considering model (3.1.1)-(3.1.3) is that it can be used to describe a large class of physical phenomena such as the variation of the electric potential created by a uniformly distributed random sources in a lossy medium, where the loss is described here by A^2. Another important motivation for considering such a model is given in the following theorem.

Theorem 3.2 *Any isotropic process that is obtained by passing 2-D white noise through a rational and proper 2-D circularly symmetric linear filter has an internal realization of the form (3.1.1)-(3.1.2).*

Proof Consider the scalar 2-D random field $z(\vec{r})$ described by the partial differential equation

$$P(\frac{\partial}{\partial r_1}, \frac{\partial}{\partial r_2})z(\vec{r}) = Q(\frac{\partial}{\partial r_1}, \frac{\partial}{\partial r_2})u(\vec{r}) \tag{3.1.22}$$

where $u(\vec{r})$ is a 2-D white noise process of intensity I_m. Here, $P(s_1, s_2)$ and $Q(s_1, s_2)$ are 2-D polynomials in the variables s_1 and s_2. Equation (3.1.22) implies that $z(\vec{r})$ is the output of a rational 2-D filter $H(\vec{\lambda})$ driven by the noise process $u(\vec{r})$, where

$$H(\vec{\lambda}) = \frac{Q(j\lambda_1, j\lambda_2)}{P(j\lambda_1, j\lambda_2)}. \tag{3.1.23}$$

The spectrum of $z(\vec{r})$ is given by

$$S_z(\vec{\lambda}) = |H(\vec{\lambda})|^2. \tag{3.1.24}$$

In [92], Yadrenko shows that the process $z(\vec{r})$ is isotropic if and only if the 2-D polynomials $P(\cdot,\cdot)$ and $Q(\cdot,\cdot)$ are functions of $\lambda = (\lambda_1^2 + \lambda_2^2)^{1/2}$ only, i.e. if $P(\cdot,\cdot)$

and $Q(\cdot,\cdot)$ are of the form

$$P(j\lambda_1, j\lambda_2) = \sum_{k=0}^{n} p_k(-\lambda^2)^k \tag{3.1.25}$$

$$= P(-\lambda^2) \tag{3.1.26}$$

$$Q(j\lambda_1, j\lambda_2) = \sum_{k=0}^{q} q_k(-\lambda^2)^k \tag{3.1.27}$$

$$= Q(-\lambda^2). \tag{3.1.28}$$

In this case, the model (3.1.22) reduces to

$$P(\nabla^2)z(\vec{r}) = Q(\nabla^2)u(\vec{r}). \tag{3.1.29}$$

Furthermore, if $n > q$, by writing

$$H(-\lambda^2) = \frac{Q(-\lambda^2)}{P(-\lambda^2)} \tag{3.1.30}$$

$$= C(-\lambda^2 I_n - M)^{-1}B \tag{3.1.31}$$

and using any of the standard 1-D state-space realization techniques with the variable s replaced by λ^2 and the operator $\frac{d}{dt}$ by the operator ∇^2, we can obtain a state space realization of $z(\vec{r})$ in the form (3.1.1)-(3.1.2).

□□

We see therefore that the class of random fields with an internal differential realization of the form (3.1.1)-(3.1.2) is quite large. It is, in fact, the analog of the class of 1- D stationary processes which are obtained by passing white-noise through a finite dimensional, linear time-invariant filter.

C Model over a Finite Disk

Over the finite disk $D_R = \{\vec{r} : r \leq R\}$, the field $z(\vec{r})$ defined by (3.1.1)-(3.1.3) or alternatively by the integral representation (3.1.5) can be modeled by (3.1.1)-(3.1.2) together with a suitable boundary condition on the edge Γ of D_R.

In addition to well-posedness, we would like to specify this boundary condition so that it is independent of the noise $u(\vec{r})$ inside the disk D_R. This will allow us later to directly apply the results of Adams et al. [2] for the estimation of boundary value processes. We shall call a boundary condition that satisfies the above conditions an *admissible* boundary condition. An admissible boundary condition can be specified as follows.

Theorem 3.3 *An admissible boundary condition for the process $x(\vec{r})$ over the disk D_R is given by*

$$\int_\Gamma [G(\vec{R},\vec{s})\frac{\partial x}{\partial n}(\vec{s}) - (\frac{\partial G}{\partial n}(\vec{R},\vec{s}))x(\vec{s})]\, dl = \beta(R,\theta),\ \ 0 \le \theta < 2\pi \tag{3.1.32}$$

where Γ is the circle of radius R, $G(\vec{r},\vec{s}) = \frac{1}{2\pi}K_0(A|\vec{r}-\vec{s}|)$, and

$$E[\beta(R,\theta)] = 0 \tag{3.1.33}$$

$$E[\beta(R,\theta)\beta^T(R,\phi)] = \Pi_\beta(R;\theta-\phi) = \sum_{k=-\infty}^{\infty} I_k(AR)\Pi_{\eta_k}(R)I_k^T(AR)e^{jk(\theta-\phi)} \tag{3.1.34}$$

with

$$\Pi_{\eta_k}(R) = \frac{1}{2\pi}\int_R^\infty K_k(As)BB^TK_k(As)s\, ds. \tag{3.1.35}$$

Here, $\frac{\partial}{\partial n}$ and dl denote respectively the normal derivative with respect to Γ and an infinitesimal element of arc length along Γ. The functions $I_k(Ar)$ and $K_k(Ar)$ are matrix modified Bessel functions of the first and second kind respectively, and of order k (see Appendix 3.A and [7]).

Theorem (3.3) is proved in Appendix 3.B, where by repeatedly applying Green's identity it is shown that the boundary condition (3.1.32) leads to a well-posed problem and that the process $x(\vec{r})$ given by (3.1.5) is the unique solution to eq. (3.1.1) with the boundary condition (3.1.32). It is further shown that the boundary process $\beta(R,\theta)$ is independent of the noise $u(\vec{r})$ for $r < R$. Since we are primarily interested in the smoothing problem for the field $z(\vec{r})$ over the finite disk D_R, we shall assume throughout the remainder of this chapter that $z(\vec{r})$ is described by the model (3.1.1)-(3.1.2) together with the boundary condition (3.1.32) or equivalently by equations (3.1.5)-(3.1.6).

3.2 THE SMOOTHING PROBLEM

A Problem Statement

Let

$$y(\vec{r}) = z(\vec{r}) + v(\vec{r}), \quad \vec{r} \in D_R \tag{3.2.1}$$

with $D_R = \{\vec{r} : r \leq R\}$, be noisy observations of the isotropic field $z(\vec{r})$ defined by the internal model (3.1.1)-(3.1.2) together with the boundary condition (3.1.32). Here, $v(\vec{r})$ is a two-dimensional white Gaussian noise field of dimension p uncorrelated with $u(\vec{r})$ and $\beta(R,\theta)$, and with intensity V, where V is a positive definite matrix. Thus,

$$E[v(\vec{r})u^T(\vec{s})] = 0 \tag{3.2.2}$$

$$E[v(\vec{r})\beta^T(R,\theta)] = 0 \tag{3.2.3}$$

$$E[v(\vec{r})v^T(\vec{s})] = V\delta(\vec{r}-\vec{s}) \tag{3.2.4}$$

where $\delta(\vec{r})$ denotes a two-dimensional delta function. The estimation problem that we consider here consists in computing the conditional mean

$$\hat{z}(\vec{r}|R) = E[z(\vec{r})|\, y(\vec{s}) \,:\, 0 \leq s \leq R] \tag{3.2.5}$$

for all $\vec{r} \in D_R$.

B Solution via Fourier Series Expansions

Following [41], our estimation procedure relies on the Fourier series expansions of the observation, signal and observation and process noise fields, e. g.

$$f(r,\theta) = \sum_{k=-\infty}^{\infty} f_k(r)\, e^{jk\theta}, \tag{3.2.6}$$

$$f_k(r) = \frac{1}{2\pi}\int_0^{2\pi} f(r,\theta)e^{-jk\theta}\, d\theta \tag{3.2.7}$$

where $f(\cdot)$ stands for $y(\cdot)$, $z(\cdot)$, $x(\cdot)$, $u(\cdot)$ or $v(\cdot)$. (See Section 2.1.) Note that the Fourier coefficient processes $y_k(r)$, $z_k(r)$, $u_k(r)$ and $v_k(r)$ are one dimensional

processes. Substituting the Fourier series expansions of $y(\cdot)$, $z(\cdot)$, and $v(\cdot)$ into (3.2.1) yields

$$y_k(r) = z_k(r) + v_k(r), \quad 0 \leq r \leq R. \tag{3.2.8}$$

According to the results of Section 2.1, the Fourier coefficient process of different orders are *uncorrelated*, i.e.

$$E[\alpha_k(r)\gamma_l^H(s)] = 0, \quad \text{for } k \neq l \tag{3.2.9}$$

where $\alpha(\cdot)$ and $\gamma(\cdot)$ stand for $y(\cdot)$, $z(\cdot)$, $x(\cdot)$, $u(\cdot)$ or $v(\cdot)$. Consequently, our original two-dimensional estimation problem requires only the solution of a countable set of decoupled 1-D smoothing problems for the Fourier coefficient process $z_k(r)$ given the observations $y_k(s)$ over the interval $0 \leq s \leq R$. Once the smoothed estimates $\hat{z}_k(r|R) = E[z_k(r)|y_k(s) \ : \ 0 \leq s \leq R]$ are found, $\hat{z}(\vec{r}|R)$ may be computed as

$$\hat{z}(\vec{r}|R) = \sum_{k=-\infty}^{\infty} \hat{z}_k(r|R)\, e^{jk\theta} \tag{3.2.10}$$

where the equality in (3.2.10) is to be understood in the mean-square sense. In practice, of course, one would consider only a finite number N of the above one dimensional estimation problems. We shall have more to say about this point in Section 3.4.

C State-Space Models For The Fourier Processes

Using the internal model (3.1.1)-(3.1.2) and (3.2.1) for the process $z(\vec{r})$ and the observations $y(\vec{r})$, 1-D state-space two-point boundary value models can be constructed for the Fourier coefficient processes $z_k(r)$ and $y_k(r)$ as follows.

Theorem 3.4 *A two-point boundary value (TPBV) model describing $z_k(r)$ and $y_k(r)$ over the interval [0,R] is given by*

$$\frac{d}{dr}\begin{bmatrix} \xi_k(r) \\ \eta_k(r) \end{bmatrix} = \begin{bmatrix} -rI_k(Ar)B \\ rK_k(Ar)B \end{bmatrix} u_k(r) \tag{3.2.11}$$

$$z_k(r) = [CK_k(Ar) \ \ CI_k(Ar)] \begin{bmatrix} \xi_k(r) \\ \eta_k(r) \end{bmatrix} \tag{3.2.12}$$

$$y_k(r) = z_k(r) + v_k(r), \tag{3.2.13}$$

with the boundary conditions

$$\xi_k(0) = 0 \qquad \textit{with probability 1} \tag{3.2.14}$$

and

$$\eta_k(R) \sim N(0, \Pi_{\eta_k}(R)) \tag{3.2.15}$$

where $\Pi_{\eta_k}(R)$ *is given by equation (3.1.35). Here,* $u_k(r)$ *and* $v_k(r)$ *are two one dimensional zero-mean white Gaussian noise processes with covariance*

$$E\left[\begin{bmatrix} u_k(r) \\ v_k(r) \end{bmatrix} [u_k^T(s) \;\; v_k^T(s)]\right] = \begin{bmatrix} I & 0 \\ 0 & V \end{bmatrix} \frac{\delta(r-s)}{2\pi r}. \tag{3.2.16}$$

Note that the TPBV model dynamics (3.2.11) are extremely simple, consisting of a gain matrix multiplying the input noise process $u_k(r)$. This is to be contrasted with the more complicated dynamics of an equivalent Markovian model for $z_k(r)$ that we shall develop in the next section.

Proof

To derive equations (3.2.11)-(3.2.13), we shall use the following identity [7]

$$K_0(A|\vec{r} - \vec{s}|) = \sum_k I_k(Ar_<)K_k(Ar_>)\cos(k(\theta - \phi)) \tag{3.2.17}$$

where $\vec{r} = (r, \theta)$, $\vec{s} = (s, \phi)$, $r_< = \min(r, s)$ and $r_> = \max(r, s)$. Upon multiplying both sides of (3.1.1)-(3.1.2) and (3.2.1) by $e^{-jk\theta}/2\pi$ and integrating from 0 to 2π, we obtain

$$\begin{aligned} x_k(r) &= -K_k(Ar)\int_0^r I_k(As)Bu_k(s)s\,ds \\ &\quad - I_k(Ar)\int_r^\infty K_k(As)Bu_k(s)s\,ds \qquad (3.2.18) \\ z_k(r) &= Cx_k(r) \qquad (3.2.19) \end{aligned}$$

and

$$y_k(r) = z_k(r) + v_k(r). \tag{3.2.20}$$

Define the state variables $\xi_k(r)$ and $\eta_k(r)$ by

$$\xi_k(r) = -\int_0^r I_k(As)Bu_k(s)s\,ds \tag{3.2.21}$$

and

$$\eta_k(r) = -\int_r^\infty K_k(As)Bu_k(s)s\,ds. \tag{3.2.22}$$

Then, it follows from (3.2.18)-(3.2.22) that a TPBV model describing $y_k(r)$ over the interval [0,R] is given by the system (3.2.11)-(3.2.13).

□□

Note that the boundary condition for the process $\eta_k(r)$ follows directly from the boundary condition (3.1.32) for the process $x(\vec{r})$ upon recognizing from identity (B.8) that

$$\beta(R,\theta) = \sum_{k=-\infty}^{\infty} I_k(Ar)\eta_k(R)\,e^{jk\theta}. \tag{3.2.23}$$

Note also that the TPBV model (3.2.11)-(3.2.15) is well-posed, since $z_k(r)$ can be expressed uniquely in terms of $u_k(r)$ and $\eta_k(R)$ as

$$\begin{aligned} z_k(r) \;=\; & -C(K_k(Ar)\int_0^r I_k(As)Bu_k(s)s\,ds \\ & + I_k(Ar)\int_r^R K_k(As)Bu_k(s)s\,ds + I_k(Ar)\eta_k(R)) \end{aligned} \tag{3.2.24}$$

Furthermore, observe that $\eta_k(R)$ is independent of $u_k(r)$ for $r \leq R$.

3.3 1-D SMOOTHERS

In this section we discuss two solutions to the 1-D TPBV smoothing problems for the Fourier coefficient processes. The first solution is based on a Markovianization procedure followed by standard 1-D smoothing techniques, while the second solution is a direct application of the method proposed by Adams et al. [2]. Conceptually, the difference between the two approaches lies in the way they deal with the boundary conditions for the smoother. In the method of Adams et al. the

boundary conditions are replaced initially by zero boundary conditions and a two-filter smoothing formula with simple dynamics is used. Once all the measurements $y_k(r)$ have been processed, a second step is required to take the true boundary conditions into account. On the other hand, the Markovianization approach deals with the boundary conditions directly as the measurements are processed. It does so by properly incorporating the boundary conditions into the dynamics of the estimator, a step that results in a more complicated smoother implementation.

A The Markovianization Approach

As mentioned earlier, the main feature of the TPBV model (3.2.11)-(3.2.15) describing the k^{th} order Fourier coefficient is that it is separable, i.e. the boundary conditions $\xi_k(0)$ and $\eta_k(R)$ are decoupled (cf. [35]). Hence, a Markovian model of the same order as the model (3.2.11)-(3.2.15) can be constructed for $x_k(r)$ by reversing the direction of propagation of $\eta_k(r)$ using a technique introduced by Verghese and Kailath [82] for constructing backwards Markovian models. Let

$$\mathcal{F}_r^k = \sigma\{\eta_k(s)\,,\ 0 \le s \le r\} \tag{3.3.1}$$

be the sigma field generated by the process $\eta_k(s)$ over the interval $[0, r]$. Then

$$\begin{aligned} \hat{u}_k(r) &= E[u_k(r)|\mathcal{F}_r^k] \\ &= E[u_k(r)\eta_k^H(r)]E[\eta_k(r)\eta_k^H(r)]^{-1}\eta_k(r) \\ &= -\frac{1}{2\pi}B^T K_k(Ar)\Pi_{\eta_k}^{-1}(r)\eta_k(r), \end{aligned} \tag{3.3.2}$$

where

$$\Pi_{\eta_k}(r) = \frac{1}{2\pi}\int_r^\infty K_k(As)BB^T K_k^T(As)s\,ds, \tag{3.3.3}$$

and where we have assumed that $\Pi_{\eta_k}(r)$ is non-singular. The process $\tilde{u}_k(r)$ defined by

$$\tilde{u}_k(r) = u_k(r) - \hat{u}_k(r) \tag{3.3.4}$$

is then an $\mathcal{F}_r^k$ - martingale with the same intensity $I_m/2\pi r$ as $u_k(r)$. Substituting (3.3.2) and (3.3.4) into (3.2.11)-(3.2.15) yields the forwards propagating model

$$\frac{d}{dr}\begin{bmatrix}\xi_k(r)\\ \eta_k(r)\end{bmatrix} = \begin{bmatrix}0 & G_k(r)\\ 0 & F_k(r)\end{bmatrix}\begin{bmatrix}\xi_k(r)\\ \eta_k(r)\end{bmatrix} + \begin{bmatrix}-rI_k(Ar)B\\ rK_k(Ar)B\end{bmatrix}\tilde{u}_k(r) \tag{3.3.5}$$

$$y_k = [CK_k(Ar)\ \ CI_k(Ar)]\begin{bmatrix}\xi_k(r)\\ \eta_k(r)\end{bmatrix} + v_k(r), \tag{3.3.6}$$

with

$$E[\tilde{u}_k(r)] = 0 \tag{3.3.7}$$

$$E[v_k(r)] = 0 \tag{3.3.8}$$

$$E\left[\begin{bmatrix}\tilde{u}_k(r)\\ v_k(r)\end{bmatrix}[\tilde{u}_k^H(s)\ \ v_k^H(s)]\right] = \begin{bmatrix}I & 0\\ 0 & V\end{bmatrix}\frac{\delta(r-s)}{2\pi r}, \tag{3.3.9}$$

and where

$$G_k(r) = \frac{r}{2\pi}I_k(Ar)BB^TK_k^T(Ar)\Pi_{\eta_k}^{-1}(r) \tag{3.3.10}$$

and

$$F_k(r) = -\frac{r}{2\pi}K_k(Ar)BB^TK_k^T(Ar)\Pi_{\eta_k}^{-1}(r). \tag{3.3.11}$$

The initial conditions for the state-space model (3.3.5) at r=0 are given by

$$\begin{bmatrix}\xi_k(r)\\ \eta_k(r)\end{bmatrix} \sim N(0, \Pi_k(0)) \tag{3.3.12}$$

with

$$\Pi_k(0) = \begin{bmatrix}0 & 0\\ 0 & \Pi_{\eta_k}(0)\end{bmatrix}, \tag{3.3.13}$$

where we have used the fact that

$$E[\xi_k(0)\eta_k^H(0)] = 0. \tag{3.3.14}$$

Here, $\eta_k^H(r)$ denotes the complex conjugate transpose of $\eta_k(r)$. The smoothing problem associated with the system (3.3.5)-(3.3.6) over [0,R] is a standard causal smoothing problem and can be solved using any of the 1-D smoothing techniques such as the Mayne-Fraser two-filter formula [45], [21], or the Rauch-Tung-Striebel formula [61], among others.

B The TPBV Smoother Formulation

Directly applying the results of [2] to the TPBV model (3.2.11)-(3.2.15), we find that the smoothed estimates of $\xi_k(r)$ and $\eta_k(r)$, $\hat{\xi}_k(r)$ and $\hat{\eta}_k(r)$ respectively, satisfy the following Hamiltonian TPBV system

$$\frac{d}{dr}\begin{bmatrix} \hat{\xi}_k(r) \\ \hat{\eta}_k(r) \\ \hat{\gamma}_k(r) \\ \hat{\delta}_k(r) \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2\pi r}BB^T \\ 2\pi r C^T V^{-1} C & 0 \end{bmatrix} \begin{bmatrix} \hat{\xi}_k(r) \\ \hat{\eta}_k(r) \\ \hat{\gamma}_k(r) \\ \hat{\delta}_k(r) \end{bmatrix} + \begin{bmatrix} 0 \\ -2\pi r C^T V^{-1} y_k(r) \end{bmatrix}, \tag{3.3.15}$$

where

$$B^T = [-rB^T I_k^T(Ar) \;\; rB^T K_k(Ar)] \tag{3.3.16}$$

$$C = [CK_k(Ar) \;\; CI_k(Ar)], \tag{3.3.17}$$

and with the boundary conditions,

$$\hat{\xi}_k(0) = 0, \tag{3.3.18}$$

$$\hat{\delta}_k(0) = 0, \tag{3.3.19}$$

$$\hat{\gamma}_k(R) = 0, \tag{3.3.20}$$

$$\hat{\delta}_k(R) = -\Pi_{\eta_k}^{-1}(R)\hat{\eta}_k(R). \tag{3.3.21}$$

An alternative way of deriving the Hamiltonian system (3.3.15) is to note that, since $x(\vec{r})$ is described by the model (3.1.1)-(3.1.2) with the boundary condition (3.1.32), then according to Adams et al. [2], the 2-D smoothed estimate of $x(\vec{r})$, $\hat{x}(\vec{r}|R)$, satisfies the Hamiltonian system

$$(I_n\nabla^2 - M)\hat{x}(\vec{r}|R) = BB^T\hat{\Theta}(\vec{r}|R) \tag{3.3.22}$$

$$(I_n\nabla^2 - M^T)\hat{\Theta}(\vec{r}|R) = C^T V^{-1}(y(\vec{r}) - C\hat{x}(\vec{r}|R)) \tag{3.3.23}$$

with the boundary condition [2]

$$\begin{bmatrix} -\frac{\partial\hat{\Theta}}{\partial n}(\vec{R}|R) \\ \hat{\Theta}(\vec{R}|R) \end{bmatrix} = V^* \Pi_\beta^{-1} V \begin{bmatrix} \hat{x}(\vec{R}'|R) \\ \frac{\partial}{\partial n}\hat{x}(\vec{R}'|R) \end{bmatrix} \tag{3.3.24}$$

where $\vec{R} = (R, \theta)$ and $\vec{R'} = (R, \phi)$. In the above identity, if $L_2^k(\Gamma)$ denotes the space of k vector functions which are square-integrable over Γ, the operator $V : L_2^{2n}(\Gamma) \rightarrow L_2^n(\Gamma)$ is such that for

$$f(\vec{R}) = [f_1^T(\vec{R}) \;\; f_2^T(\vec{R})]^T \tag{3.3.25}$$

we have

$$(Vf)(\vec{R}) = \int_\Gamma [-(\frac{\partial G}{\partial n}(\vec{R}, \vec{R'}))f_1(\vec{R'}) + G(\vec{R}, \vec{R'})f_2(\vec{R'})]\, dl \tag{3.3.26}$$

where $G(\vec{R}, \vec{R'}) = \frac{1}{2\pi} K_0(A|\vec{R} - \vec{R'}|)$ and where dl denotes an element of arc length along Γ. In (3.3.24) the operators V^* and Π_β^{-1} denote respectively the Hilbert adjoint of the operator V and the inverse of the correlation operator associated to the kernel $\Pi_\beta(R; \theta - \phi)$ defined in eq. (3.1.35). If we introduce the variable

$$\Psi(\vec{r}|R) = 2\pi r \hat{\Theta}(\vec{r}|R) \tag{3.3.27}$$

and substitute the Fourier expansions of $\hat{x}(\vec{r}|R)$ and $\hat{\Psi}(\vec{r}|R)$ into (3.3.22)- (3.3.23) we obtain the following Hamiltonian system for the k^{th} order Fourier coefficient process

$$(I_n(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{k^2}{r^2}) - M)\hat{x}_k(r|R) = \frac{1}{2\pi r} BB^T \hat{\psi}_k(r|R) \tag{3.3.28}$$

$$(I_n(\frac{d^2}{dr^2} - \frac{d}{dr}[\frac{1}{r}\cdot] - \frac{k^2}{r^2}) - M^T)\hat{\psi}_k(r|R) = 2\pi r C^T V^{-1} (y_k(r) - C\hat{x}_k(r|R)). \tag{3.3.29}$$

By properly selecting the state variables

$$\hat{\xi}_k(r) = r\frac{d}{dr}(I_k(Ar))\hat{x}_k(r|R) - rI_k(Ar)\frac{d}{dr}\hat{x}_k(r|R), \tag{3.3.30}$$

$$\hat{\xi}_k(r) = -r\frac{d}{dr}(K_k(Ar))\hat{x}_k(r|R) + rK_k(Ar)\frac{d}{dr}\hat{x}_k(r|R), \tag{3.3.31}$$

$$\hat{\gamma}_k(r) = (\frac{(k+1)}{r}A^{-T}K_k^T(Ar) - K_{k+1}^T(Ar))\hat{\psi}_k(r|R) - A^{-T}K_k^T(Ar)\frac{d}{dr}\hat{\psi}_k(r|R) \tag{3.3.32}$$

$$\hat{\delta}_k(r) = (\frac{(k+1)}{r}A^{-T}I_k^T(Ar) + I_{k+1}^T(Ar))\hat{\psi}_k(r|R) - A^{-T}I_k^T(Ar)\frac{d}{dr}\hat{\psi}_k(r|R) \tag{3.3.33}$$

it can be shown that the second order Hamiltonian system (3.3.28)-(3.3.29) has the realization (3.3.15).

The Hamiltonian TPBV system (3.3.15) for the smoother with the boundary conditions (3.3.18)-(3.3.21), can be solved efficiently by using the procedure proposed by Adams et al. (see [2] for details).

Once the smoothed estimates $\hat{\xi}_k(r)$ and $\hat{\eta}_k(r)$ have been computed for all k, the smoothed estimate $\hat{z}(\vec{r}|R)$ of $z(\vec{r})$ can be found as

$$\hat{z}(\vec{r}|R) = \sum_{k=-\infty}^{\infty} C(K_k(Ar)\hat{\xi}_k(r) + I_k(Ar)\hat{\eta}_k(r))e^{jk\theta}. \tag{3.3.34}$$

Finally, as noted earlier, the two efficient processing schemes that we have developed for estimating isotropic random fields with an internal differential realization of the form (3.1.1)-(3.1.2) are based on a concept of causality where the data is processed outwards or inwards with respect to a disk of observation as shown in Figs. 3.1 a and 3.1 b. Observe that this concept of causality follows naturally from the special geometrical structure of isotropic random fields.

3.4 IMPLEMENTATION ISSUES

In this section we briefly discuss some implementation issues. Specifically, we examine the problem of truncating the series (3.2.10) and the problem of implementing the 1-D smoothers of Section 3.3.

A Truncation of the Series Representation of the Smoothed Estimate

The smoothed estimate $\hat{z}(\vec{r}|R)$ is given by equation (3.2.10) as an infinite sum of the 1-D Fourier coefficient processes smoothed estimates $\hat{z}_k(r|R)$. In practice, of course, one would consider a finite set of 1-D smoothing problems and one would approximate the series (3.2.10) by the finite series

$$\hat{z}_N(\vec{r}) = \sum_{|k|\leq N} \hat{z}_k(r|R)\, e^{jk\theta}. \tag{3.4.1}$$

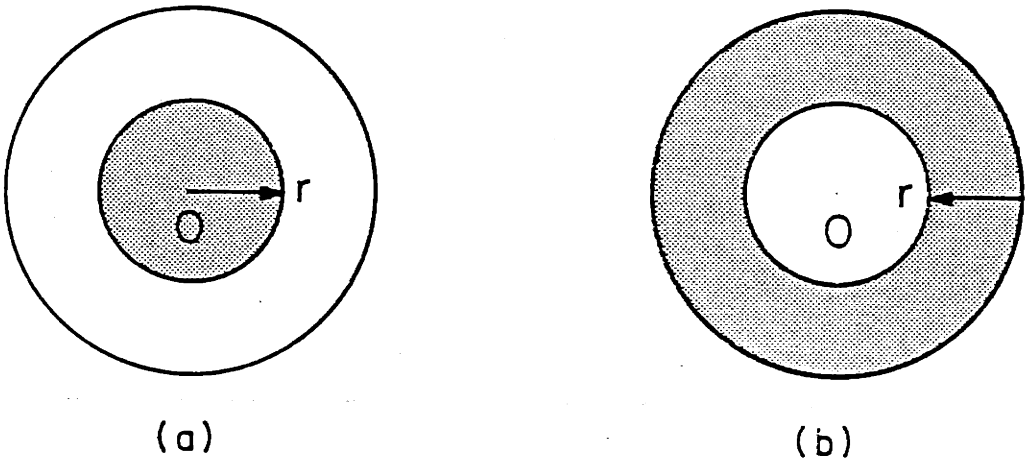

Figure 3.1: Outgoing and incoming radial recursions.

Note that, with $\tilde{z}_N(\vec{r}) = z(\vec{r}) - \hat{z}_N(\vec{r})$ and with $\tilde{z}_k(r) = z_k(r) - \hat{z}_k(r|R)$, we have

$$E[\tilde{z}_N(\vec{r})\tilde{z}_N^T(\vec{r})] = \sum_{|k|\leq N} E[\tilde{z}_k(r)\tilde{z}_k^H(r)] + \sum_{|k|>N} E[z_k(r)z_k^H(r)]. \tag{3.4.2}$$

As r tends to zero, the matrix $E[z_k(r)z_k^H(r)]$ tends to zero as r^2 for $k \neq 0$. Furthermore, as r tends to infinity, the matrix $E[z_k(r)z_k^H(r)]$ tends to zero as r^{-1} for all values of k. Hence, in order to keep the variance of the estimator error small, the number 2 N + 1 of terms to be used in (3.4.1) should increase with the distance between the origin and the point where $z(\vec{r})$ is to be estimated. If r is small, one can use very few terms in (3.4.1) and still obtain a good estimate of $z(\vec{r})$. In fact, for $r = 0$ one needs only the zeroth order Fourier coefficient process smoothed estimate, $\hat{z}_0(0)$, in order to compute $\hat{z}(0|R)$ exactly.

Finally, note that the concept of a white-noise process is really an idealization of the physical reality. In some cases of interest, we may assume that the noise sources $u(\cdot)$ and $v(\cdot)$ have covariance functions that can be well approximated by expressions of the form

$$E[u(\vec{r})u^T(\vec{s})] = \frac{\delta(r-s)}{r} \sum_{|n|\leq N} a_n e^{jn(\theta-\phi)} \tag{3.4.3}$$

$$E[v(\vec{r})v^T(\vec{s})] = V\frac{\delta(r-s)}{r} \sum_{|n|\leq N} b_n e^{jn(\theta-\phi)} \tag{3.4.4}$$

where $\vec{r} = (r,\theta)$ and $\vec{s} = (s,\phi)$ and where we have assumed that $u(\cdot)$ and $v(\cdot)$ are white in the radial direction only. In such a case where the noise processes are not exactly white in the angular variable, we can obtain an *exact* solution to the smoothing problem that we have considered in this chapter by using $2N + 1$ terms in (3.2.10).

B 1-D Smoother Implementation

At first glance the implementation of the 1-D smoothers of Section 3.3 poses some problems since the models (3.2.11)-(3.2.13) and (3.3.5)-(3.3.6) are not well behaved in the vicinity of $r = 0$ for $k \neq 0$. This can be seen from the singularity of $K_k(Ar)$ at $r = 0$, and is not surprising since the Fourier series decomposition

degenerates at the origin. We now show that this is of no practical consequence. In practice, to compute $\hat{z}(\vec{r}|R)$ we divide the intervals $[0, R]$ and $[0, 2\pi]$ into M and N subintervals of length $\Delta_1 = R/M$ and $\Delta_2 = 2\pi/N$ respectively. As a result, the Fourier coefficient processes $y_k(r)$ are available at the positions $r = m\Delta_1$, $0 \leq m \leq M$. The 1-D smoothed estimates $\hat{z}_k(r|R)$ are then found by discretizing the smoother equations corresponding to models (3.2.11)-(3.2.13) and (3.3.5)-(3.3.6). In particular, for $k \neq 0$ we consider the 1-D discretized smoother implementations for $1 \leq m \leq M$. Note that, since $z_k(0) = 0$ and $y_k(0) = 0$ with probability one for $k \neq 0$, then

$$\hat{z}_k(\Delta_1|0) = 0, \quad k \neq 0. \tag{3.4.5}$$

Thus,

$$E[z_k(m\Delta_1)|y_k(l\Delta_1) \ : \ 0 \leq l \leq M] = E[z_k(m\Delta_1)|y_k(l\Delta_1) \ : \ 1 \leq l \leq M] \tag{3.4.6}$$

for $k \neq 0$. For $k = 0$, the models (3.2.11)-(3.2.13) and (3.3.5)-(3.3.6) are well behaved at $r = 0$. Hence, we solve the 1-D discretized smoothing problem for the zeroth order Fourier coefficient for $0 \leq m \leq M$. Observe that the zeroth order Fourier coefficient process is the only process needed to compute $\hat{z}(0|R)$. Consequently, in practice, the smoothed estimate of $\hat{z}(m\Delta_1, n\Delta_2)$ is computed as

$$\hat{z}(m\Delta_1, n\Delta_2) = \begin{cases} \hat{z}_0(0) & \text{if } m = 0 \\ \sum\limits_{|k| \leq K} \hat{z}_k(m\Delta_1)\, e^{jkn\Delta_2} & \text{otherwise,} \end{cases} \tag{3.4.7}$$

where

$$\hat{z}_k(m\Delta_1) = \begin{cases} E[z_0(m\Delta_1)|y_0(l\Delta_1) \ : \ 0 \leq l \leq M] & k = 0 \\ E[z_k(m\Delta_1)|y_k(l\Delta_1) \ : \ 1 \leq l \leq M] & k \neq 0, \end{cases} \tag{3.4.8}$$

and where K is some number suitably chosen (see the previous section).

3.5 ASYMPTOTIC BEHAVIOR OF THE DIFFERENTIAL MODELS AT INFINITY

The Fourier coefficient processes $x_n(r)$ have a finite variance for all $r \in \mathbf{R}$ since by definition $x(\vec{r})$ has finite variance over the whole plane (see Section 3.1.) Hence,

the optimal estimator for the Fourier coefficient process $x_n(r)$ written in integral form, must have a well-behaved kernel for all r. However, the matrices appearing in (3.2.11)-(3.2.13) and (3.3.5)-(3.3.6) are *not* well-behaved as r tends to zero or infinity. This ill-behavior is due to the singularity of $K_k(Ar)$ and $I_k(Ar)$ as r tends to zero and infinity respectively [1], [7]. Furthermore, models (3.2.11)-(3.2.13) and (3.3.5)-(3.3.6) define a singular estimation problem as r tends to infinity. This follows from the fact that the intensity of the noise processes $u_k(r)$, $\tilde{u}_k(\cdot)$ and $v_k(\cdot)$ varies as r^{-1}. In Section 3.4, we discussed a strategy for dealing with the singularity of the models (3.2.11)-(3.2.13) and (3.3.5)-(3.3.6) as r tends to zero. Here, we introduce differential models for the Fourier coefficient processes that are well-behaved as r tends to infinity.

A Models

The models that we develop are obtained by applying the state transformation

$$\chi_k(r) = T_k(r) \begin{bmatrix} \xi_k(r) \\ \eta_k(r) \end{bmatrix} \tag{3.5.1}$$

$$T_k(r) = \begin{bmatrix} K_k(Ar) & I_k(Ar) \\ -K_{k+1}(Ar) & I_{k+1}(Ar) \end{bmatrix} \tag{3.5.2}$$

to models (3.2.11)-(3.2.13) and (3.3.5)-(3.3.6), followed by a normalization of all the processes. The normalization consists in multiplying all processes by $r^{1/2}$, which forces the intensity of the noise processes to be a constant.

Note that by using (3.2.12) we can identify

$$\chi_k(r) = \begin{bmatrix} x_k(r) \\ A^{-1}(\frac{d}{dr}x_k(r) - \frac{k}{r}x_k(r)) \end{bmatrix}. \tag{3.5.3}$$

Note also that the transformation $T_k(r)$ has the properties that

$$\frac{d}{dr}T_k(r) = \begin{bmatrix} \frac{k}{r}I & A \\ A & -\frac{(k+1)}{r}I \end{bmatrix} T_k(r) \tag{3.5.4}$$

$$T_k^{-1}(r) = r \begin{bmatrix} AI_{k+1}(Ar) & -AI_k(Ar) \\ AK_{k+1}(Ar) & AK_k(Ar) \end{bmatrix}. \tag{3.5.5}$$

Identities (3.5.4)-(3.5.5) can be derived by using the recurrence relations for modified Bessel functions [1], [7] and the Wronskian identity [1], [7]

$$I_{k+1}(Ar)K_k(Ar) + I_k(Ar)K_{k+1}(Ar) = A^{-1}r^{-1}. \tag{3.5.6}$$

If we apply the state transformation $T_k(r)$ to the model (3.2.11)-(3.2.13) and if we introduce the *normalized processes*

$$\bar{\alpha}_k(r) = \sqrt{r}\alpha_k(r) \tag{3.5.7}$$

where $\alpha_k(r)$ stands for $\chi_k(r)$, $u_k(r)$, $y_k(r)$ or $v_k(r)$, we obtain

$$\frac{d}{dr}\bar{\chi}_k(r) = (A_k(r) + \frac{I}{2r})\bar{\chi}_k(r) + \bar{B}\bar{u}_k(r) \tag{3.5.8}$$

$$\bar{y}_k(r) = \bar{C}\bar{\chi}_k(r) + \bar{v}_k(r) \tag{3.5.9}$$

where

$$A_k(r) = \begin{bmatrix} \frac{k}{r}I & A \\ A & -\frac{(k+1)}{r}I \end{bmatrix} \tag{3.5.10}$$

$$\bar{B} = \begin{bmatrix} 0 \\ A^{-1}B \end{bmatrix} \tag{3.5.11}$$

$$\bar{C} = [C \quad 0], \tag{3.5.12}$$

and where we have used (3.5.4)-(3.5.5). In (3.5.8)-(3.5.9) $\bar{u}_k(r)$ and $\bar{v}_k(r)$ are two uncorrelated zero-mean Gaussian noise processes with intensities $I/2\pi$ and $V/2\pi$ respectively. Hence, (3.5.8)-(3.5.9) does not lead to a singular estimation problem.

Similarly, by using the state transformation $T_k(r)$ and normalizing all processes we find that model (3.3.5)-(3.3.6) is transformed to

$$\frac{d}{dr}\bar{\chi}_k(r) = (A'_k(r) + \frac{I}{2r})\bar{\chi}_k(r) + \bar{B}\bar{\bar{u}}_k(r) \tag{3.5.13}$$

$$\bar{y}_k(r) = \bar{C}\bar{\chi}_k(r) + \bar{v}_k(r) \tag{3.5.14}$$

where

$$A'_k(r) = \begin{bmatrix} \frac{k}{r}I & A \\ A + D_k(r) & -\frac{(k+1)}{r}I + E_k(r) \end{bmatrix} \tag{3.5.15}$$

$$D_k(r) = -\frac{r}{2\pi}A^{-1}BB^T K_k^T(Ar)\Pi_{\eta_k}^{-1}(r)K_{k+1}(Ar)A \tag{3.5.16}$$

$$E_k(r) = -\frac{r}{2\pi}A^{-1}BB^T K_k^T(Ar)\Pi_{\eta_k}^{-1}(r)K_k(Ar)A, \tag{3.5.17}$$

and where $\bar{\tilde{u}}_k(r)$ and $\bar{v}_k(r)$ are two uncorrelated zero-mean Gaussian noise processes with intensities $I/2\pi$ and $V/2\pi$ respectively. Once more, note that this implies that (3.5.13)-(3.5.14) defines a nonsingular estimation problem.

To study the asymptotic behavior of models (3.5.8)-(3.5.9) and (3.5.13)-(3.5.14) as r tends to infinity, we note that as r tends to infinity the modified Bessel functions $K_k(Ar)$ and $I_k(Ar)$ have the asymptotic forms [1]

$$I_k(Ar) \sim (2\pi Ar)^{-\frac{1}{2}} e^{Ar} \tag{3.5.18}$$

$$K_k(Ar) \sim (\frac{2Ar}{\pi})^{-\frac{1}{2}} e^{-Ar}. \tag{3.5.19}$$

Hence, if we assume that the pair (A, B) is controllable we obtain

$$\begin{aligned} \lim_{r\to\infty} D_k(r) &= \lim_{r\to\infty} -A^{-1}BB^T e^{-A^T r}(\int_r^\infty e^{-As}BB^T e^{-A^T s}\,ds)^{-1} e^{-Ar}A \\ &= -A^{-1}BB^T Q^{-1}A \\ &= D, \end{aligned} \tag{3.5.20}$$

where Q is the matrix

$$Q = \int_0^\infty e^{-As}BB^T e^{-A^T s}\,ds. \tag{3.5.21}$$

Note that since $-A$ is a stable matrix and since the pair (A, B) is controllable then Q is the unique positive definite solution of the matrix equation [8]

$$-AQ - QA^T + BB^T = 0. \tag{3.5.22}$$

Similarly, we have

$$\lim_{r\to\infty} E_k(r) = D. \tag{3.5.23}$$

Thus, as r tends to infinity the TPBV model (3.5.8)-(3.5.9) takes the form

$$\frac{d}{dr}\bar{\chi}_k(r) = \bar{A}\bar{\chi}_k(r) + \bar{B}\bar{u}_k(r) \tag{3.5.24}$$

$$\bar{y}_k(r) = \bar{C}\bar{\chi}_k(r) + \bar{v}_k(r) \tag{3.5.25}$$

where

$$\bar{A} = \begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}, \tag{3.5.26}$$

whereas the Markovian model (3.5.13)-(3.5.14) takes the form

$$\frac{d}{dr}\bar{\chi}_k(r) = \bar{A}'\bar{\chi}_k(r) + \bar{B}\bar{\tilde{u}}_k(r) \tag{3.5.27}$$

$$\bar{y}_k(r) = \bar{C}\bar{\chi}_k(r) + \bar{v}_k(r) \tag{3.5.28}$$

where

$$\bar{A}' = \begin{bmatrix} 0 & A \\ A+D & D \end{bmatrix}. \tag{3.5.29}$$

Note that the asymptotic models (3.5.24)-(3.5.25) and (3.5.27)-(3.5.28) imply that the models (3.5.8)-(3.5.9) and (3.5.13)-(3.5.14) are well-behaved as r tends to infinity. Note also that the asymptotic models (3.5.24)-(3.5.25) and (3.5.27)-(3.5.28) are space invariant models that do not depend on the order k of the Fourier coefficient process under consideration. This reflects the fact that as r tends to infinity all the Fourier coefficient processes have an equal importance in the sense that we would have to retain a very large number of terms in (3.2.10) to obtain meaningful results, as was already observed in Section 3.4. Furthermore, model (3.5.27)-(3.5.28) provides a *stable spectral factorization* of $S_x(\lambda)$. In particular, observe that the transfer function associated with equation (3.5.27) is

$$\begin{aligned} W_f(s) &= A(sI+A)^{-1}(sI-A+A^{-1}BB^TQ^{-1}A)^{-1}A^{-1}B \\ &= (sI+A)^{-1}(sI-A+BB^TQ^{-1})^{-1}B. \end{aligned} \tag{3.5.30}$$

The formula

$$-A+BB^TQ^{-1} = QA^TQ^{-1} \tag{3.5.31}$$

(which is easily derived from (3.5.22)) now shows that $-A+BB^TQ^{-1}$ and A have the same eigenvalues. Therefore, $W_f(s)$ will have its poles in the left half-plane since all the eigenvalues of A have a positive real part by assumption. Note that this also implies that the matrix $\bar{A}'$ is a stable matrix. Furthermore, observe that

$$W_f(s)U(s) = W_b(s) \tag{3.5.32}$$

where

$$\begin{aligned} W_b(s) &= (sI + A)^{-1}(sI - A)^{-1}B \\ &= (s^2 I - A^2)^{-1}B \\ &= (s^2 I - M)^{-1}B \end{aligned} \tag{3.5.33}$$

and where

$$U(s) = I + B^T Q^{-1}(sI - A)^{-1}B. \tag{3.5.34}$$

It is easy to verify that $U(s)$ is a paraunitary or allpass transfer function in the sense that

$$U(s)U^T(-s) = U^T(-s)U(s) = I. \tag{3.5.35}$$

Hence, we have

$$\begin{aligned} W_f(s)W_f^T(-s) &= W_b(s)W_b^T(-s) \\ &= (sI + A)^{-1}(sI - A)^{-1}BB^T(-sI - A^T)^{-1}(-sI + A^T)^{-1} \\ &= S_x(\lambda)|_{\lambda=-js}, \end{aligned} \tag{3.5.36}$$

which proves that the asymptotic model (3.5.27)-(3.5.28) does lead to a stable spectral factorization of $S_x(\lambda)$. Finally, observe that the results of [82] imply that $(sI - A + BB^T Q^{-1})^{-1}B$ is the transfer function of a stable *forward* Markovian model corresponding to the stable *backwards* Markovian model with transfer function $(sI - A)^{-1}B$.

We now show that the stability of the matrix $\bar{A}'$ implies that the Kalman filter associated to (3.5.13)-(3.5.14) is stable. To do this we will need the following lemma which is an adaptation of a result of Coddington and Levinson ([15], p. 314).

Lemma 3.1 *Let*

$$\dot{x} = Ax + f(t, x) \tag{3.5.37}$$

where A is a real constant matrix with eigenvalues all having negative real parts. Furthermore, let f be real, continuous for small $|x|$ and $t \geq 0$, and such that

$$f(t, x) = o(|x|) \quad \text{as} \quad |x| \to 0 \tag{3.5.38}$$

uniformly in t, $t \geq 0$. Then, the system (3.5.37) is exponentially stable in a neighborhood of $x = 0$.

Proof

Let $\phi(t)$ be a solution of (3.5.37). So long as $\phi(t)$ exists, it follows from (3.5.37) that

$$\phi(t) = e^{At}\phi(0) + \int_0^t e^{A(t-s)} f(s, \phi(s))\, ds. \tag{3.5.39}$$

Because the real parts of the eigenvalues of A are negative, there exists positive constants K and σ such that

$$|e^{At}| \leq Ke^{-\sigma t} \quad \text{for} \quad t \geq 0. \tag{3.5.40}$$

Hence, we have

$$|\phi(t)| \leq K|\phi(0)|e^{-\sigma t} + K\int_0^t e^{-\sigma(t-s)}|f(s, \phi(s))|\, ds. \tag{3.5.41}$$

Given $\epsilon > 0$, there exists by assumption a $\delta > 0$ such that $|f(t, x)| \leq \epsilon|x|/K$ for $|x| \leq \delta$. Thus, as long as $|\phi(t)| \leq \delta$, it follows that

$$e^{\sigma t}|\phi(t)| \leq K|\phi(0)| + \epsilon\int_0^t e^{\sigma s}|\phi(s)|\, ds. \tag{3.5.42}$$

This inequality yields

$$e^{\sigma t}|\phi(t)| \leq K|\phi(0)|e^{\epsilon t}, \tag{3.5.43}$$

or

$$|\phi(t)| \leq K|\phi(0)|e^{-(\sigma-\epsilon)t} \quad \text{for} \quad t \geq 0. \tag{3.5.44}$$

The above discussion now implies that if initially $|\phi(0)| \leq \delta/K$, then $|\phi(t)|$ will decay exponentially to zero.

□□

Lemma 3.1 can now be used to prove the following result.

Theorem 3.5 *The system defined by equations (3.5.13)-(3.5.14) is exponentially stable.*

Proof

The proof follows by writing

$$A'_k(r) = \bar{A}' + \bar{A}'_k(r) \tag{3.5.45}$$

where $\bar{A}'$ is defined in (3.5.29). By taking $f(r,x)$ in Lemma 3.1 as

$$f(r,\bar{\chi}_k) = \bar{A}'_k(r)\bar{\chi}_k(r), \tag{3.5.46}$$

and noting that

$$\lim_{r\to\infty} \bar{A}'_k(r) = 0, \tag{3.5.47}$$

we obtain the desired result by invoking Lemma 3.1.

□□

By using Theorem 3.5 we can state and prove the main result of this section.

Theorem 3.6 *The Kalman filter associated with the model (3.5.13)-(3.5.14) is asymptotically stable. Furthermore, the error covariance associated with the normalized process $\bar{\chi}_k(r)$ converges to a non-negative definite matrix $\bar{P}$ as r tends to infinity, where $\bar{P}$ is the solution of the algebraic Riccati equation*

$$0 = \bar{A}'\bar{P} + \bar{P}\bar{A}'^T + \bar{B}\bar{B}^T - \bar{P}\bar{C}^T V^{-1}\bar{C}\bar{P}, \tag{3.5.48}$$

where the matrix $\bar{A}'$ is defined in (3.5.29).

Proof

The result follows by direct application of Theorem 4.11 of [37].

□□

B Discussion

The normalized models (3.5.8)-(3.5.9) and (3.5.13)-(3.5.14) that we have obtained in this section involve the state transformation $T_k(r)$ of (3.5.2) and a normalization procedure which consists in multiplying all processes by the gain $r^{1/2}$. Note that the transformation $T_k(r)$, its inverse $T_k^{-1}(r)$ and the normalization gain $r^{1/2}$ blow up as r tends to infinity. (The transformation $T_k(r)$ and its inverse $T_k^{-1}(r)$ blow up as r tends to infinity because of the singularity of the matrix functions $I_k(Ar)$ as r tends to infinity.) However, the normalized processes that appear in (3.5.8)-(3.5.9) and (3.5.13)-(3.5.14) are well-behaved and have a finite variance as r tends to infinity. In fact, by using the asymptotic forms of $K_k(Ar)$ and $I_k(Ar)$ as r tends to infinity (cf. (3.5.18)-(3.5.19)) and using equation (3.5.1), it can be shown that the process $\chi_k(r)$ has a variance that tends to zero as r^{-1} as r tends to infinity. Furthermore, recall that the intensity of the noise processes $u_k(r)$ and $v_k(r)$ is also proportional to r^{-1}. Hence, the variance of all the Fourier coefficient processes tends to zero as r^{-1} as r tends to infinity. This is precisely the reason why we have to keep a very large number of terms in (3.2.10) to obtain meaningful results as r tends to infinity. Note that this also implies that all the normalized processes are well-behaved with variances and noise intensities that tend to a finite constant as r tends to infinity.

Observe also that the models (3.5.8)-(3.5.9) and (3.5.13)-(3.5.14) show that we can interpret the Fourier coefficient process $y_k(r)$ as being the output of a cascaded system which is driven by the non-singular noise processes $\bar{u}_k(r)$ and $\bar{v}_k(r)$. The cascaded system consists of a system which is well-behaved as r tends to infinity followed by a gain stage with a gain of $r^{-1/2}$, as shown in Fig. 3.2.

Finally, observe that the stability analysis for the Kalman filters that we have presented in this section really concerns the Kalman filters associated with the normalized Markovian model (3.5.13)-(3.5.14) which corresponds to the center block in Fig. 3.3. The filtered estimates $\hat{\chi}_k(r)$ are actually obtained by feeding the observations $y_k(r)$ into an input gain stage with a gain of $r^{1/2}$ followed by the asymptotically stable Kalman filters associated with the Markovian model (3.5.13)-(3.5.14) and an output gain stage with a gain of $r^{-1/2}$, as is shown in

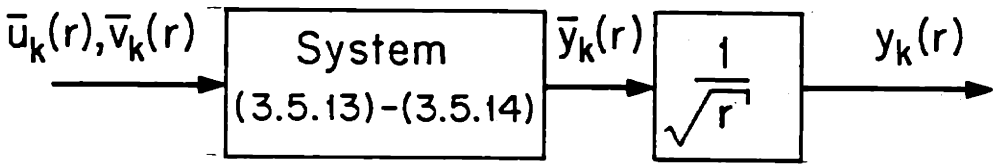

Figure 3.2: A model for $y_k(r)$ for large values of r.

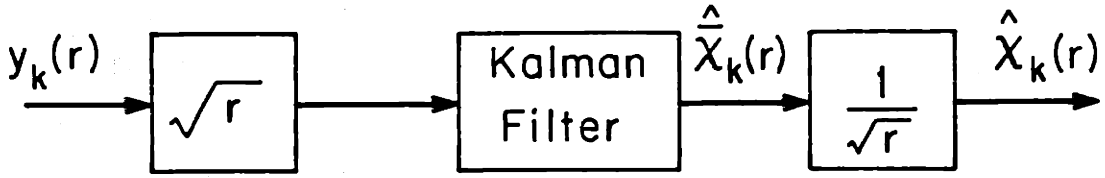

Figure 3.3: Filtering procedure for large values of r.

Fig. 3.3. Note that the overall Kalman filter based estimation procedure is well-behaved according to the above discussion.

3.6 CONCLUSION AND EXTENSIONS

In this chapter we have obtained efficient recursive estimation techniques for isotropic random fields described by non-causal internal differential realizations. By exploiting the properties of isotropic random fields, we showed that the problem of estimating an isotropic random field given noisy observations over a finite disk of radius R is equivalent to a countably infinite set of decoupled one-dimensional two-point boundary value system (TPBV) estimation problems for the Fourier coefficient processes of the random field. We then solved the 1-D TPBV estimation problems using either the method of Adams et al. [2] or by using a Markovianization approach followed by standard 1-D smoothing techniques. We have also studied the asymptotic behavior of the Markovian models that we developed as the radius R of the disk of observation tends to infinity, and we have shown that the 1-D Kalman filters associated to these models are asymptotically stable The smoothing schemes that we have developed result in a processing structure that is recursive with respect to the radius r in a polar coordinate representation of the field.

Note that the approach that we have used in this chapter carries over to the case where the source term $u(\cdot)$ appearing on the right hand side of (3.1.1) is not spatially white but has a covariance function which is invariant under rotations only. In particular, it applies to the case where the field $u(\cdot)$ has a covariance function of the form

$$\begin{aligned} E[u(\vec{r})u^T(\vec{s})] &= K_1(r,s)K_2(\theta-\phi) \\ &= K_1(r,s)\sum_k a_k e^{jk(\theta-\phi)} \end{aligned} \tag{3.6.49}$$

where $K_1(r,s)$ is a positive definite function of the variables r and s which is assumed to have a *finite* dimensional state-space realization. In such a case the noise process $u(\cdot)$ has a Fourier series expansion with uncorrelated Fourier coefficients. By substituting the Fourier series expansion of $u(\cdot)$ into (3.1.1) we find that the field $z(\cdot)$

has also a Fourier series expansion with uncorrelated coefficients. Hence, the 2-D estimation problem for the field $z(\cdot)$ can be reduced to a countably infinite number of 1-D estimation problems for its Fourier coefficient processes. Note however that in such a case the TPBV model (3.2.11)-(3.2.13) describing the 1-D Fourier coefficient processes has to be properly augmented to account for the fact that the processes $u_k(r)$ have covariance functions $a_k K_1(r,s)$ and must be realized as the output of a 1-D dynamical system driven by white noise.

The approach presented in this chapter can also be used in the case where the matrices A, B and C of (3.1.1)-(3.1.2) are functions of the polar coordinate variable r only. Again in such a case, the TPBV model (3.2.11)-(3.2.13) has to be properly modified to account for the r dependence of the matrices A, B and C. However, our method fails when the matrices A, B or C are allowed to be a function of the angular coordinate variable θ, and when the source term $u(\cdot)$ has a covariance function which is not invariant under rotations. Since this latter case is of importance in a number of applications (e.g. in ocean acoustics) where the source term $u(\cdot)$ is often homogeneous with a 2-D power spectrum that has an angular dependence only in the wave-number plane, alternative estimation approaches have to be developed. Chapter 7 discusses this case in more detail.

APPENDIX 3.A

In this chapter, we make frequent use of the matrix modified Bessel functions of the first and second kinds, $I_k(Ar)$ and $K_k(Ar)$. These functions are a generalization of the corresponding scalar modified Bessel functions, and they satisfy the matrix differential equation

$$(I_n(\frac{d^2}{dr^2}+\frac{1}{r}\frac{d}{dr}-\frac{k^2}{r^2})-A^2)F(r)=0 \tag{3.A.1}$$

with the limiting forms

$$I_k(Ar) \sim (k!)^{-1}(\frac{Ar}{2})^k \tag{3.A.2}$$

$$K_0(Ar) \sim \ln(Ar) \tag{3.A.3}$$

$$K_k(Ar) \sim \frac{(k-1)!}{2}(\frac{Ar}{2})^{-k}, \quad k \geq 1 \tag{3.A.4}$$

as r tends to zero, and with the asymptotic forms

$$I_k(Ar) \sim (2\pi Ar)^{-\frac{1}{2}}e^{Ar} \tag{3.A.5}$$

$$K_k(Ar) \sim (\frac{2Ar}{\pi})^{-\frac{1}{2}}e^{-Ar} \tag{3.A.6}$$

as r tends to infinity. Thus $I_k(Ar)$ and $K_k(Ar)$ are regular at $r=0$, and as r tends to infinity, respectively.

Bessel functions have a number of useful properties which are listed in [7].

APPENDIX 3.B

In this appendix we show that the boundary condition for the process $x(\vec{r})$ given in Theorem (3.3) is independent of the noise inside the disk D_R, where $D_R = \{\vec{r} : \quad r \leq R\}$, and that this boundary condition leads to a well-posed problem. From Green's identity, we have

$$\int_{D_R} G(\vec{r},\vec{s})((I_n\nabla^2 - A^2)x(\vec{s})) - ((I_n\nabla^2 - A^2)G(\vec{r},\vec{s}))x(\vec{s})\, d\vec{s} \quad = \int_\Gamma [G(\vec{r},\vec{s})\frac{\partial}{\partial n}x(\vec{s}) - (\frac{\partial}{\partial n}G(\vec{r},\vec{s}))x(\vec{s})]\, dl \tag{3.B.1}$$

where $\vec{r} \in D_R$, $G(\vec{r},\vec{s}) = \frac{1}{2\pi}K_0(A|\vec{r}-\vec{s}|)$, and where $\frac{\partial}{\partial n}$ denotes the normal derivative with respect to the curve $\Gamma = \{\vec{r} : \quad r = R\}$. Here, dl is an element of arc length along Γ. Equation (3.B.1) implies that

$$x(\vec{r}) = -\int_{D_R} G(\vec{r},\vec{s})Bu(\vec{s})\, d\vec{s} + \int_\Gamma [G(\vec{r},\vec{s})\frac{\partial}{\partial n}x(\vec{s}) - (\frac{\partial}{\partial n}G(\vec{r},\vec{s}))x(\vec{s})]\, dl. \tag{3.B.2}$$

However, from (3.1.5) $x(\vec{r})$ can be expressed as

$$x(\vec{r}) = -\int_{D_R} G(\vec{r},\vec{s})Bu(\vec{s})\, d\vec{s} - \int_{D_R^c} G(\vec{r},\vec{s})Bu(\vec{s})\, d\vec{s} \tag{3.B.3}$$

where D_R^c denotes the complement of D_R in $\mathbf{R}^2$. Hence, we conclude from (3.B.2) and (3.B.3) that

$$\Phi_R(\vec{r}) = \int_\Gamma [G(\vec{r},\vec{s})\frac{\partial}{\partial n}x(\vec{s}) - (\frac{\partial}{\partial n}G(\vec{r},\vec{s}))x(\vec{s})]\, dl \tag{3.B.4}$$

$$= \int_{D_R^c} G(\vec{r},\vec{s})Bu(\vec{s})\, d\vec{s} \tag{3.B.5}$$

and the above identity can be used to specify a boundary condition for $x(\vec{r})$ which is independent of the noise inside the disk D_R. Specifically, $\Phi_R(R,\theta)$ depends only on the noise $u(\vec{r})$ outside the disk D_R, and is therefore independent of the noise inside D_R. Let

$$\beta(R,\theta) = \Phi_R(R,\theta). \tag{3.B.6}$$

By taking the expression (3.B.5) into account, and using the expansion [7]

$$K_0(A|\vec{r}-\vec{s}|) = \sum_k I_k(Ar_<)K_k(Ar_>)\cos(k(\theta-\phi)) \tag{3.B.7}$$

where $\vec{r}=(r,\theta)$, $\vec{s}=(s,\phi)$, $r_< = \min(r,s)$ and $r_> = \max(r,s)$, we obtain

$$\beta(R,\theta) = -\sum_{k=-\infty}^{\infty} I_k(Ar)\int_R^{\infty} K_k(As)Bu_k(s)s\,ds\,e^{jk\theta}, \tag{3.B.8}$$

where

$$u_k(r) = \frac{1}{2\pi}\int_0^{2\pi} u(r,\theta)e^{-jk\theta}\,d\theta. \tag{3.B.9}$$

Since the random variables $u_k(s)$ and $u_l(s)$ with $k \neq l$, are independent zero-mean white Gaussian noise processes of intensity $I/2\pi s$, it follows that

$$E[\beta(R,\theta)] = 0, \tag{3.B.10}$$

$$\begin{aligned} E[\beta(R,\theta)\beta^T(R,\phi)] &= \Pi_\beta(R;\theta-\phi) \\ &= \sum_{k=-\infty}^{\infty} I_k(AR)\Pi_{\eta_k}(R)I_k^T(AR)e^{jk(\theta-\phi)}, \end{aligned} \tag{3.B.11}$$

where $\Pi_{\eta_k}(R)$ is given by (3.1.35). Then, as indicated in Theorem (3.3), equation (3.B.5) together with (3.B.10) and (3.B.11) can be used to specify a boundary condition for the 2-D field $x(\vec{r})$ in terms of the boundary process $\beta(R,\theta)$.

To show that the boundary condition (3.1.32) leads to a well-posed problem, note that (3.B.4) implies that $\Phi_R(\vec{r})$ satisfies

$$(I_n\nabla^2 - A^2)\Phi_R(\vec{r}) = 0, \quad \text{for } \vec{r}\in D_R \tag{3.B.12}$$

$$\Phi_R(R,\theta) = \beta(R,\theta). \tag{3.B.13}$$

Let $G_R(\vec{r},\vec{s})$ denote that Green's function corresponding to the system (3.B.12)-(3.B.13). Then $G_R(\vec{r},\vec{s})$ obeys the equation

$$(I_n\nabla^2 - A^2)G_R(\vec{r},\vec{s}) = -I_n\delta(\vec{r}-\vec{s}), \tag{3.B.14}$$

for $\vec{r},\vec{s}\in D_R$, with the boundary condition

$$G_R(\vec{R},\vec{s}) = 0 \text{ for } \vec{R}\in\Gamma. \tag{3.B.15}$$

Now using Green's identity, we obtain

$$\Phi_R(\vec{r}) = -\int_\Gamma (\frac{\partial}{\partial n} G_R(\vec{r},\vec{s}))\beta(\vec{R})\, dl. \tag{3.B.16}$$

Then, combining relations (3.B.3), (3.B.4) and (3.B.16), $x(\vec{r})$ can be expressed as

$$x(\vec{r}) = -\int_{D_R} G(\vec{r},\vec{s})Bu(\vec{s})\, d\vec{s} - \int_\Gamma (\frac{\partial}{\partial n} G_R(\vec{r},\vec{s}))\beta(\vec{R})\, dl, \tag{3.B.17}$$

which shows that the boundary value problem for $x(\vec{r})$ is well-posed.

Chapter 4

THE CYLINDRICAL HARMONICS RETRIEVAL PROBLEM

In Chapter 3 we have considered recursive estimation techniques for isotropic random fields. In various applications, e.g. in radar [46], sonar [6], image processing [85] etc., one is interested instead in estimating the power spectrum of an underlying random field. For example, in geophysical applications and in ocean acoustics knowledge of the frequency-wave-number power spectrum of the background noise field is important for studying the performance of optimal array processing schemes [13]. Spectral estimation techniques have also been applied to the problems of processing electroencephalographic data monitored via an array of scalp electrodes [25], and of reconstructing unknown radio brightness distributions in radio astronomy [85]. Several multidimensional spectral estimation methods (see Chapter 1 and [47] for references and discussion) have been developed in the past and have been applied to the above mentioned problems. These techniques can certainly be used to estimate the frequency-wave-number power spectrum of isotropic random fields. However, these methods are very general and do not attempt to exploit any special structure of the power spectrum to be estimated (i.e. they do not make explicit use of isotropy).

In this chapter and in the next one, we develop two spectral estimation methods which are specifically adapted to isotropic random fields. In particular, we present in this chapter a new high resolution spectral estimation method for a class of 2-D

isotropic random fields that is often used to model some types of background noises in both geophysics and ocean acoustics. Our algorithm uses the special structure of the covariance functions of isotropic random fields in this class. The specific class of 2-D isotropic fields that we consider corresponds to fields whose covariance function can be modeled as a weighted sum of *cylindrical harmonics*. The term cylindrical harmonics is used in this context to denote a radially symmetric function $f(\vec{r})$ of the form $J_0(\lambda_k r)$ where λ_k is a fixed two-dimensional "cylindrical frequency" measured in radians per unit distance. Such covariance functions arise in geophysics whenever the background noise field consists of either fundamental-mode, or higher mode, Rayleigh waves propagating uniformly from all azimuths simultaneously [13]. The covariance function of the "circle noise" [5] in ocean acoustics is also of this particular form. Note that the two-dimensional Fourier transform of a cylindrical harmonic is radially symmetric in the wave-number plane and consists of a *cylindrical* impulse at a radial frequency of λ_k radians per unit distance (i.e. the Fourier transform is a 1-D sheet of impulses concentrated on a circle of radius λ_k). Hence, our problem is to determine the number, location and amplitude of the cylindrical impulses in the wave-number spectrum of an isotropic field. This problem differs from the one which was investigated by Lang and McClellan [38] and by Wax and Kailath [84], who extended Pisarenko's method [58] and the MUSIC method [73],[9] respectively, and used them to estimate power spectra which are equal to a weighted sum of multidimensional *point* impulses.

4.1 MOTIVATION AND PROBLEM STATEMENT

Array processing is a popular technique for solving estimation problems involving propagating waves. Such problems arise in geophysics and in ocean acoustics, among other fields. Arrays of receivers often work against a background noise field. Knowledge of the frequency-wave-number power spectrum of the background noise field is important for studying the performance of optimal array processing schemes [13]. In many applications, the medium in which the waves of interest are propagating supports surface waves; for example Rayleigh and Love waves or

internal ocean waves. In such applications, the background noise is often modeled as consisting of a large number of independent waves propagating from all azimuths simultaneously with the same velocity c m/sec and with the same frequency f_0 Hertz. The frequency-wave-number power spectrum of such a background noise has the form [13],[5]

$$S(\omega : \vec{\lambda}) = 2\pi C \frac{\delta(\lambda - \lambda_0)}{\lambda}, \tag{4.1.1}$$

where C is a positive constant and where $\lambda_0 = 2\pi f_0/c$ is the wave-number of the background noise in rad/m. A background noise that has a frequency-wave-number power spectrum of the form (4.1.1) is called a ***circle noise*** in [5] and has a temporal frequency spatial correlation function (i.e. the inverse Fourier transform of $S(\omega : \vec{\lambda})$ with respect to $\vec{\lambda}$) of the form

$$K(\omega : \vec{r}) = C J_0(\lambda_0 r). \tag{4.1.2}$$

Observe that $K(\omega : \vec{r})$ is a cylindrical harmonic. In some situations, the background noise can be a superposition of a number L of circle noises of different wave-numbers, and one is then interested in estimating both the cylindrical frequencies corresponding to the circle noise wave-numbers λ_l and the cylindrical harmonics amplitudes C_l for $1 \leq l \leq L$, in order for example to evaluate the performance of any processing array which is to be used in the presence of such a background noise field.

In other situations, one might be interested in finding the propagating surface modes that an isotropic medium of interest can support. In this case, one can excite the medium with a large number of independent wideband directional sources uniformly distributed over the circumference of a circle whose radius is large compared to any wavelength of interest, and all radiating towards the center of the circle. One then measures the response of the medium close to the center of such a circle. If the isotropic medium can support only a finite number of wave-numbers then the resulting waves will have a frequency-wave-number power spectrum of the form (4.1.1).

In all of the above situations, there is a need to estimate isotropic covariance functions that are equal to a weighted sum of cylindrical harmonics. This is the

problem that we consider in this chapter. We shall assume that we are given some noisy measurements $y(\vec{r})$ of a Gaussian isotropic random field $z(\vec{r})$ whose covariance function is of the form

$$K(r) = \sum_{l=1}^{L} P_l J_0(\lambda_l r), \tag{4.1.3}$$

or equivalently whose power spectrum is of the form

$$S(\lambda) = \sum_{l=1}^{L} P_l \frac{\delta(\lambda - \lambda_l)}{\lambda}. \tag{4.1.4}$$

Specifically, we assume that measurements $y(\vec{r})$ are made on a finite set of concentric circles of radii $\{r_i : 1 \leq i \leq I\}$, and that $y(r_i, \theta)$ is given by

$$y(r_i, \theta) = z(r_i, \theta) + n(r_i, \theta), \ 0 \leq \theta \leq 2\pi. \tag{4.1.5}$$

In (4.1.5) the observation noise $n(r_i, \theta)$ is uncorrelated with $z(\vec{r})$ and is a zero-mean Gaussian white noise process of intensity σ^2 in the discrete radial coordinate r_i and the continuous angle coordinate θ, i.e.

$$E[n(r_i, \theta)] = 0, \tag{4.1.6}$$

$$E[n(r_i, \theta) n(r_j, \phi)] = \sigma^2 \frac{\delta_{i,j}}{r_i} \delta(\theta - \phi). \tag{4.1.7}$$

Our objective is to solve the cylindrical harmonics retrieval problem, i.e. to simultaneously estimate the measurement noise power σ^2 and to reconstruct $S(\lambda)$ by finding the cylindrical harmonics powers P_l and the cylindrical harmonics frequencies λ_l.

Note the parallel between the cylindrical harmonics retrieval problem that we have described above and the 1-D harmonic retrieval problem where one is interested in estimating a *stationary* covariance function that is equal to a weighted sum of cosine functions. In the 1-D case, the objective is to estimate both the location and amplitude of a number of 1-D impulses in the frequency domain while in the 2-D case the goal is to find the location and amplitude of a number of cylindrical impulses in the wave-number plane. It will turn out that the algorithm that we propose in the next section for solving the cylindrical harmonics retrieval problem is very similar in spirit to 1-D harmonic retrieval procedures [73], [9] based on an eigenanalysis of a covariance matrix, even though our algorithm uses samples of the

non-stationary covariance function of the zeroth-order Fourier coefficient process corresponding to the measurements $y(\vec{r})$.

Finally, observe that the harmonic retrieval problem can be solved by using any of the high resolution 2-D spectral estimation methods. As mentioned earlier, these methods are very general and do not exploit any special property of the power spectrum to be estimated. In comparison, our procedure takes explicitly into account the isotropy property of $y(\vec{r})$, as well as the special structure (4.1.4) of the spectrum that we want to estimate.

4.2 THE EIGENSTRUCTURE APPROACH

In one-dimensional signal processing, several eigenstructure methods have been proposed to solve the 1-D harmonic retrieval problem, i.e. the problem of estimating covariance functions of the form

$$r(i,j) = \sigma^2 \delta_{i,j} + \sum_{l=1}^{L} P_l \cos(2\pi f_l (i-j)\Delta t), \tag{4.2.1}$$

where $\delta_{i,j}$ is a Kronecker delta. These techniques require an eigenanalysis of the autocorrelation matrix $R = [r(i,j)]$. In particular, Pisarenko's method [58] uses the value of the smallest eigenvalue of R as an estimate of the noise power σ^2. The locations of the frequencies f_l are then determined by finding the zeros of a polynomial whose coefficients are the elements of the eigenvector corresponding to the smallest eigenvalue. The disadvantage of this method lies in the fact that, as the size of the matrix R grows, (a necessary feature for resolving closely spaced frequencies), the number of close eigenvalues corresponding to the white noise component of the signal becomes large, leading to an ill-conditioned eigenvector determination problem [86]. To overcome this problem, Schmidt [73] and Bienvenu and Kopp [9] noted that the computation of the subspace spanned by the eigenvectors corresponding to the set of smallest eigenvalues is much less sensitive to perturbations in the entries of the matrix R, than the computation of individual eigenvectors. The methods they proposed use the whole eigenspace associated with the cluster of smallest eigenvalues to estimate the frequencies of model (4.2.1). In this section, we shall

develop an algorithm for solving the cylindrical harmonics retrieval problem which is very similar in spirit to those of Schmidt and Bienvenu and Kopp.

A Mathematical Theory

In the remainder of this section we shall assume that we are given a finite number of samples $\{k_0(r_i, r_j) : 1 \leq i, j \leq n\}$ of the covariance function of the zeroth-order Fourier coefficient process $y_0(r)$ corresponding to the measurements $y(r_i, \theta)$ of (4.1.5). A procedure for estimating $k_0(r_i, r_j)$ from the given measurements will be presented in the next section. Note that under the assumptions of Section 4.2 (see (4.1.3)-(4.1.7)), $k_0(r_i, r_j)$ is of the form

$$k_0(r_i, r_j) = \sum_{l=1}^{L} P_l J_0(\lambda_l r_i) J_0(\lambda_l r_j) + \frac{\sigma^2}{2\pi} \frac{\delta_{i,j}}{r_i}. \tag{4.2.2}$$

The first step in our algorithm is to construct a symmetric matrix out of the given sample values of $k_0(r, s)$. Denote by k_{ij}

$$k_{ij} = \sqrt{r_i r_j} k_0(r_i, r_j). \tag{4.2.3}$$

Note that k_{ij} is a normalized version of $k_0(r_i, r_j)$. The normalization is introduced here to make the measurement noise intensity constant instead of inversely proportional to r_i (see (4.2.2)). This will enable us to solve the cylindrical harmonics retrieval problem by performing an eigenanalysis of the normalized zeroth order covariance matrix $R = [k_{ij}]$. In particular, note that the matrix R can be written as

$$R = S + \frac{\sigma^2}{2\pi} I_{n \times n}, \tag{4.2.4}$$

where S is an $n \times n$ symmetric matrix with entries

$$s_{ij} = \sum_{l=1}^{L} \sqrt{r_i r_j} P_l J_0(\lambda_l r_i) J_0(\lambda_l r_j). \tag{4.2.5}$$

Furthermore, it is clear from (4.2.5) that S can be decomposed as

$$S = CDC^T, \tag{4.2.6}$$

where D =diag$\{P_l\}$ is an $L \times L$ diagonal matrix, and where C is an $n \times L$ matrix with entries

$$c_{ij} = \sqrt{r_i} J_0(\lambda_j r_i). \tag{4.2.7}$$

In the sequel, we assume that C has full column rank. Hence, the rank of S is equal to $\min(n, L)$. This key observation will allow us to recover both the measurement noise power σ^2 and the cylindrical harmonics frequencies λ_l by performing an eigenanalysis of the matrix R.

To estimate the noise power σ^2, we note that if $n > L$, then the $(n-L)$ smallest eigenvalues of R are exactly equal to $\sigma^2/2\pi$. Hence, the measurement noise power can be computed as $2\pi\mu_{\min}$, where $\mu_{\min}$ is the smallest eigenvalue of R.

The location of the cylindrical harmonics can be determined by observing that if $\mathbf{y}_m = [y_m(i)]$, $1 \leq m \leq (n-L)$, is an eigenvector of R corresponding to the repeated eigenvalue $\sigma^2/2\pi$, then

$$S\mathbf{y}_m = 0. \tag{4.2.8}$$

Equation (4.2.8) implies that

$$\sum_{l=1}^{L} P_l \sqrt{r_i} J_0(\lambda_l r_i) \sum_{j=1}^{n} \sqrt{r_j} J_0(\lambda_l r_j) y_m(j) = 0, \quad \forall i. \tag{4.2.9}$$

If we denote by $f_m(\lambda)$ the quantity

$$f_m(\lambda) = \sum_{i=1}^{n} \sqrt{r_i} J_0(\lambda r_i) y_m(i), \tag{4.2.10}$$

then equation (4.2.9) is equivalent to

$$CD \begin{bmatrix} f_m(\lambda_1) \\ \vdots \\ f_m(\lambda_L) \end{bmatrix} = 0, \tag{4.2.11}$$

where CD has full rank, so that we must have

$$f_m(\lambda_l) = 0 \tag{4.2.12}$$

for $1 \leq m \leq n - L$ and $1 \leq l \leq L$. Hence, the cylindrical harmonic frequencies λ_l appearing in (4.1.3) must be the roots of the equation

$$f_m(\lambda) = 0. \tag{4.2.13}$$

However, equation (4.2.13) is not useful as a practical way of computing the values of the cylindrical frequencies. The roots of (4.2.13) are very sensitive to perturbations in the entries of the matrix R because the coefficients of (4.2.13) come from a single eigenvector associated with the smallest eigenvalue of R. To avoid this problem, we can use the whole eigenspace associated with the smallest eigenvalue of R, and take our estimates of the cylindrical frequencies to be the roots of the equation

$$\sum_{m=1}^{n-L} f_m^2(\lambda) = 0. \tag{4.2.14}$$

Finally, to compute the amplitudes P_l of the cylindrical harmonics, we use the fact that

$$\sqrt{r_i r_j} k_0(r_i, r_j) = \sum_{l=1}^{L} \sqrt{r_i r_j} P_l J_0(\lambda_l r_i) J_0(\lambda_l r_j) + \frac{\sigma^2}{2\pi} \delta_{i,j}, \ 1 \leq i, j \leq n. \tag{4.2.15}$$

Because of the symmetry of R, there are only $n(n+1)/2$ identities of the form (4.2.15). By properly scanning the indices i and j, these $n(n+1)/2$ identities can be written in matrix form as

$$\mathbf{k} = A\mathbf{p} \tag{4.2.16}$$

where the vectors $\mathbf{k}$ and $\mathbf{p} = [P_l]$ are of size $\frac{n(n+1)}{2} \times 1$ and $p \times 1$ respectively, and where A is a matrix of appropriate dimensions. The estimated cylindrical harmonics amplitudes are then taken to be equal to the entries of the optimal solution $\mathbf{p}^*$ to (4.2.16), i.e. the one that minimizes the Euclidean norm of the error $||\mathbf{k} - A\mathbf{p}||$. Note that it is well known that $\mathbf{p}^*$ is given by the equation [78]

$$\mathbf{p}^* = (A^T A)^{-1} A^T \mathbf{k}. \tag{4.2.17}$$

B Numerical Implementation

The eigenvalues and eigenvectors of R can be computed numerically by first reducing the matrix R to tridiagonal form by means of Householder transformations and then using the QR algorithm to generate the eigenvalues. The eigenvectors of R can be computed by saving and then multiplying together the transformations used in the first step. The above procedure has been implemented as part of a singular

value decomposition routine available through Linpack [17] and was found to be numerically robust and highly accurate. Its only deficiency lies in its complexity; it requires $O(n^3)$ operations where n is the size of the square symmetric matrix R.

In practice, due to inaccuracies in the estimated values of $k_0(r_i, r_j)$, the computed smallest eigenvalues of R are not all exactly equal, and the noise power has to be computed as the average of the $(n - L)$ cluster of smallest eigenvalues of R Furthermore, the separation between the "large" eigenvalues of R and the "small" eigenvalues of R is sometimes not well marked and it is difficult to determine the exact number of harmonics in the given zeroth order Fourier coefficient covariance data. In this case statistical methods can be used to determine the number of cylindrical harmonics. (See [83] for a discussion of how statistical methods can be used to determine the number of signals in the one-dimensional case.)

The search for the cylindrical harmonics is done by plotting $v(\lambda) = 1/\sum_{m=1}^{n-L} f_m^2(\lambda)$, where $f_m(\lambda)$ is defined in (4.2.10). The roots of equation (4.2.14) correspond to peaks of $v(\lambda)$. This step is numerically robust and poses no problems.

Finally, the estimation of the amplitudes of the cylindrical harmonics via equation (4.2.17) is done by performing a QR decomposition of the matrix A. The total number of operations involved in this step is of the order of $\frac{n(n+1)}{2}L^2 - \frac{L^3}{3}$, where L is the number of harmonics to be estimated and $\frac{n(n+1)}{2}$ is the total number of correlations available.

C Summary

In summary, given sample values of the covariance function $k_0(r_i, r_j)$ corresponding to the zeroth-order Fourier coefficient process associated with the measurements $y(\vec{r})$ of (4.1.5), the cylindrical harmonics retrieval problem can be solved by performing the following steps:

1. For a suitably large n, determine all of the eigenvalues and eigenvectors of the $n \times n$ covariance matrix R obtained from the normalized sample values of the zeroth order Fourier process covariance function. The noise power is equal to

$2\pi\overline{\mu}$, where $\overline{\mu}$ is the average of the $(n-L)$ cluster of smallest eigenvalues of R. The number L of larger eigenvalues is equal to the number of cylindrical harmonics.

2. Let $f_m(\lambda)$, $1 \leq m \leq n-L$, be the functions given by (4.2.10) and which are specified by the eigenvectors y_m corresponding to the $(n-L)$ smallest eigenvalues of R. Determine the roots of equation (4.2.14) by plotting $v(\lambda) = 1/\sum_{m=1}^{n-L} f_m^2(\lambda)$. The computed roots are the estimates of the values of the L cylindrical frequencies in (4.1.3).

3. Using equation (4.2.17), determine the amplitudes $\{P_l : \ 1 \leq l \leq L\}$ of the L cylindrical harmonics.

The next section examines the problem of obtaining unbiased and consistent estimates of $k_0(r_i, r_j)$ from the given field measurements.

4.3 ESTIMATION OF THE COVARIANCE FUNCTIONS

The algorithm that we presented in the last section is based on the knowledge of $k_0(r_i, r_j)$, the covariance function of the zeroth order-Fourier coefficient process corresponding to the measurements $y(\vec{r})$. However in practice, one is given the measurements themselves rather than $k_0(r_i, r_j)$. In this section, we present an *unbiased* and *consistent* procedure for estimating the *non-stationary* covariance functions $k_n(r_i, r_j)$ from the measurements. This procedure is well suited for the sampling geometry that we introduced in Section 4.2, where the isotropic random field of interest is measured along a discrete set of concentric circles (see the discussion preceding (4.1.5)).

A Theory

Let us start by assuming that measurements of the field $y(\vec{r})$ are available at all points inside the disk $D_{R^*} = \{\vec{r} : \ 0 \leq r \leq R^*\}$. Then to estimate $k_n(r_i, r_j)$, we

can use a two step procedure. In the first step we estimate $K(r)$ using the given data. In the second step we substitute our estimate of $K(r)$ into (2.1.10) to obtain an estimate of $k_n(r_i, r_j)$.

$K(r)$ can be estimated by using a simple extension of the 1-D techniques that were developed to estimate the covariance function of ergodic stationary processes. Observe that along any line $\phi = \phi_0$ in a tomographic coordinate system[1], $y(\vec{r})$ is stationary. Hence, given the measurements $\{y(t, \phi_0) : \ -R^* \leq t \leq R^*\}$ along this line we can estimate $K(r)$ using a simple extension of the 1-D techniques as

$$\hat{K}(r : \phi_0) = \frac{1}{R^{*2}} \int_{-R^*}^{R^*} y(t, \phi_0) y(r + t, \phi_0) \, |t| \, dt. \tag{4.3.1}$$

Since measurements of $y(\vec{r})$ are assumed to be available all over the disk D_{R^*}, we can compute $\hat{K}(r : \phi_0)$ for all ϕ_0, $0 \leq \phi_0 \leq \pi$, and take $\hat{K}(r)$ to be the average of the $\hat{K}(r : \phi_0)$ over all ϕ_0. In other words, we can estimate $K(r)$ as

$$\hat{K}(r) = \frac{1}{\pi R^{*2}} \int_0^{R^*} ds \int_0^{2\pi} d\theta \, s \, y(s, \theta) y(r + s, \theta). \tag{4.3.2}$$

Note that we have used the weight function $w(t) = |t|$ in (4.3.1) to guarantee that $\hat{K}(r)$ corresponds to a spatial average.

Next, we can use $\hat{K}(r)$ to obtain an estimate of $k_n(r_i, r_j)$ by simply substituting $\hat{K}(r)$ for $K(r)$ into (2.1.10). Thus, we take as our estimate of $k_n(r_i, r_j)$ the quantity

$$\hat{k}_n(r_i, r_j) = \frac{1}{2\pi} \int_0^{2\pi} d\theta \, \hat{K}((r_i^2 + r_j^2 - 2 r_i r_j \cos\theta)^{1/2}) \cos n\theta. \tag{4.3.3}$$

Note that according to (4.3.3) one needs to estimate $K(r)$ for $0 \leq r \leq 2R^*$ in order to be able to estimate $k_n(r_i, r_j)$ for $0 \leq r_i, r_j \leq R^*$.

Let us now study the unbiasedness and consistency properties of the estimates (4.3.2) and (4.3.3). It is a simple matter to show that $\hat{K}(r)$ is an unbiased estimate of $K(r)$. Equation (4.3.2) implies that

$$\begin{aligned} E[\hat{K}(r)] &= \frac{1}{\pi R^{*2}} \int_0^{R^*} ds \int_0^{2\pi} d\theta \, s \, K(r) \\ &= K(r), \end{aligned} \tag{4.3.4}$$

[1] A tomographic coordinate system (t, ϕ) is a modified polar coordinate system where t takes both positive and negative real values, and where ϕ varies from 0 to π.

which proves that $\hat{K}(r)$ is indeed an unbiased estimate of $K(r)$. Furthermore, it can be shown that $\hat{K}(r)$ is a consistent estimate of $K(r)$ under the assumption that the underlying random field is Gaussian. The proof of this fact uses a number of properties of the Bessel functions and can be found in Appendix 4.A.

To show that $\hat{k}_n(r_i, r_j)$ is an unbiased and consistent estimate of $k_n(r_i, r_j)$, we note that (2.1.10) and (4.3.3) imply that $k_n(r_i, r_j)$ and $\hat{k}_n(r_i, r_j)$ are related linearly to $K(r)$ and $\hat{K}(r)$ respectively. Hence, it follows immediately from the unbiasedness and consistency properties of $\hat{K}(r)$ that $\hat{k}_n(r_i, r_j)$ is an unbiased and consistent estimate of $k_n(r_i, r_j)$. Thus, by using (4.3.2) and (4.3.3) we are able to obtain an unbiased and consistent estimate of the non-stationary covariance function $k_n(r_i, r_j)$.

B Numerical Implementation

In practice we are given the values of the field $y(\vec{r})$ at discrete points $\{(r_i, \theta_j) : 1 \leq i \leq I, 1 \leq j \leq J\}$. Let us assume for simplicity that $r_i = i\Delta$ where Δ is a positive number, and that $\theta_j = (j-1)2\pi/J$. The estimate $\hat{K}(r)$ can be computed by approximating the 2-D integral (4.3.2) with a rectangular rule in the radial coordinate s, and with a trapezoidal rule in the angular coordinate θ as

$$\hat{K}(l\Delta) \approx \frac{2}{I^2 J} \sum_{i=1}^{I} \sum_{j=1}^{J} i\, y(i\Delta, (j-1)\frac{2\pi}{J}) y((i+l)\Delta, (j-1)\frac{2\pi}{J}). \qquad (4.3.5)$$

The estimated covariance $k_n(i, j)$ can then be computed by similarly approximating the integral (4.3.3) with a trapezoidal rule as

$$\hat{k}_n(l_1, l_2) \approx \frac{1}{J} \sum_{j=1}^{J} \hat{K}((l_1^2 + l_2^2 - 2l_1 l_2 \cos((j-1)\frac{2\pi}{J}))^{1/2}) \cos(n(j-1)\frac{2\pi}{J}). \qquad (4.3.6)$$

Note that to compute $\hat{k}_n(l_1, l_2)$ via (4.3.6) we may need the value of $\hat{K}(\cdot)$ at points not equal to some multiple of Δ. However, if Δ is chosen small enough, we can interpolate the values of $\hat{K}(r)$ at the required abscissas from the values of $\hat{K}(r)$ at the points $r = k\Delta$ using any of the known interpolation schemes. In our experiments we used a linear interpolation procedure. Better, but computationally more expensive ways of approximating the integrals in (4.3.2) and (4.3.3) can be found in [16].

C Example

We now illustrate the behavior of our estimation procedure for the case where $n = 0$ with an example. The example clearly indicates the fact that as more and more data is available the estimates that we get become better and better. This is to be expected since our estimators are consistent.

Example 4.3.1

In this example we used the method of [76] to generate an isotropic random field with a covariance function consisting of two cylindrical harmonics at 0.1 rad/m and 0.3 rad/m. The amplitudes of both harmonics were taken to be equal to 10 Watts. The field was generated on a circular polar grid at the points $(0.1i, (j-1)\frac{2\pi}{50})$. We then added to this field a 2-D white noise field of intensity 3 Watt.m^2. Thus, over the rectangular grid $r = i$, $s = j$ the covariance function of the zeroth-order Fourier coefficient corresponding to the resulting field has the form

$$k_0(i,j) = 10J_0(0.1i)J_0(0.1j) + 10J_0(0.3i)J_0(0.3j) + \frac{3}{2\pi i}\delta_{i,j}. \tag{4.3.7}$$

The corresponding power spectrum is shown in Fig. 4.1. Using the values of the resulting field over the disk $0 \leq r \leq 30$, we used (4.3.10) to compute $\hat{K}(l\Delta)$ for $\Delta = 0.1$ and for $0 \leq l \leq 200$, and then used these values to compute $k_0(i,j)$ for $1 \leq i,j \leq 10$ via (4.3.11). The estimated covariance function $\hat{K}(l\Delta)$ that we computed is plotted in Fig. 4.2 together with the exact covariance function. Fig. 4.2 shows that the relative error in the estimated values of $K(l\Delta)$ was less than 2.2 percent for $l \leq 10$ and less than 130 percent for all l. The corresponding relative error in the estimated values of $k_0(i,j)$ was on the average around 40 percent and was equal to 246 percent in one case. By using the values of the field over the disk $0 \leq r \leq 100$ we obtained the estimated covariance function $\hat{K}(l\Delta)$ that is plotted in Fig. 4.3. In this case, the relative error in the estimates $\hat{K}(l\Delta)$ was less than 1.6 percent for $l \leq 10$ and less than 70 percent for all l, while the relative error in the estimated values of $k_0(i,j)$ was on the average around 30 percent and was still equal

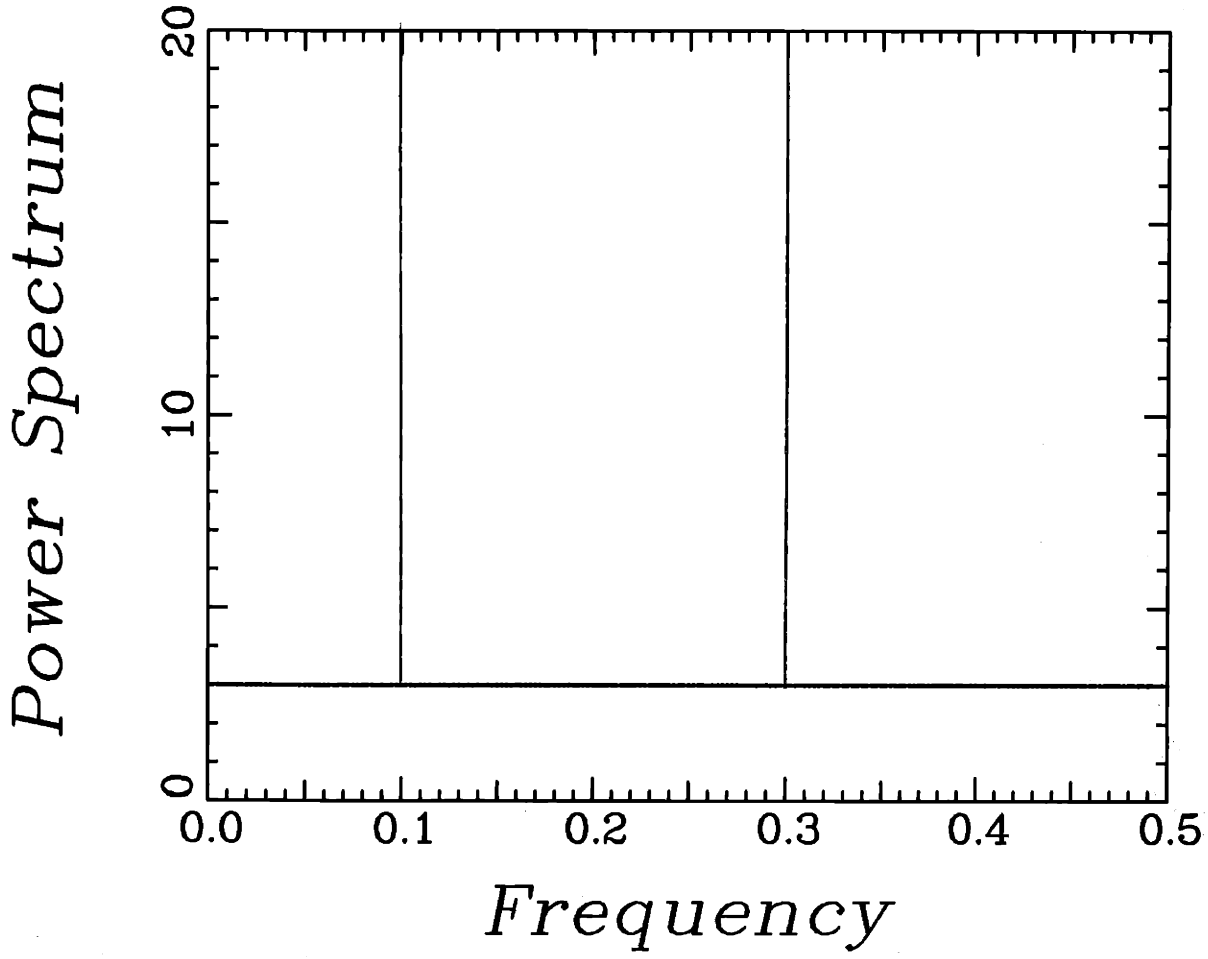

Figure 4.1: True spectrum for Examples 4.3.1 and 4.4.1.

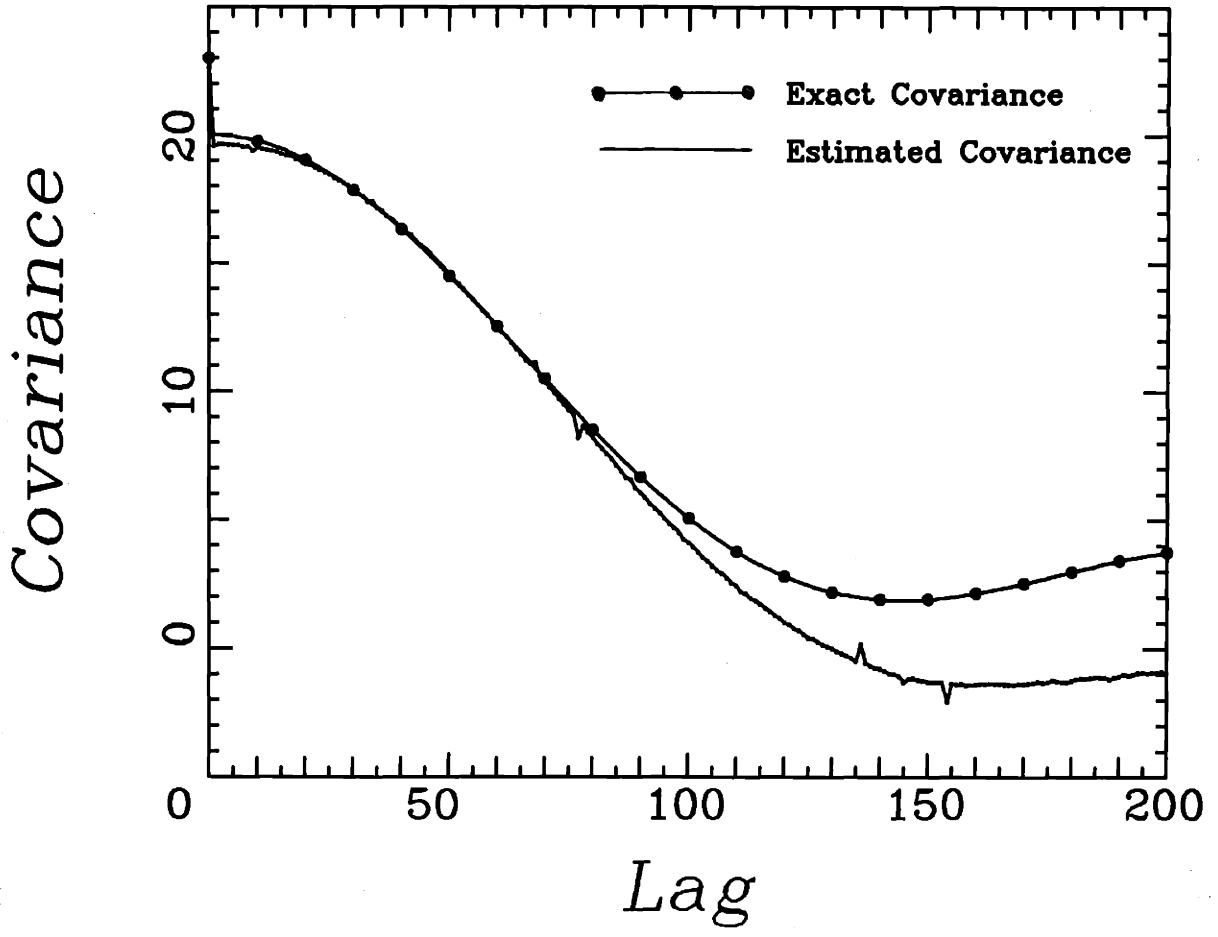

Figure 4.2: Plot of exact and estimated covariance functions for Example 4.4.1 when observations are available over a disk of radius 30.

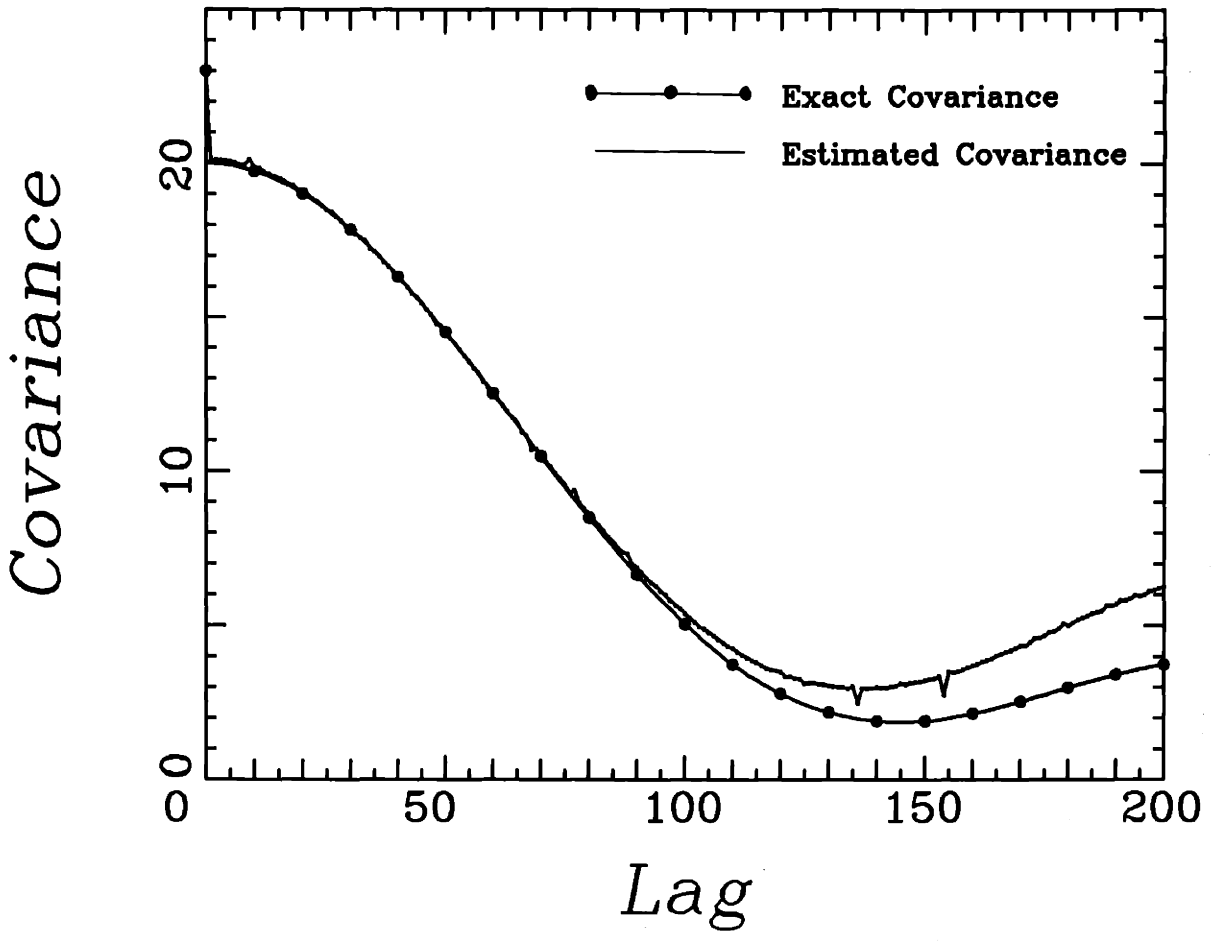

Figure 4.3: Plot of exact and estimated covariance functions for Example 4.4.1 when observations are available over a disk of radius 100.

to 246 percent in one case. Note that as with any covariance estimation method whether in one or two dimensions, one expects a degradation in performance as l increases because of the reduction in the extent of spatial averaging that can be done. This effect can clearly be seen in this example. Indeed, Figs. 4.2 and 4.3 show that the large relative errors in the values of $\hat{K}(l\Delta)$ occur for large lags. Furthermore, by comparing Figs. 4.2 and 4.3 we observe an increase in the range of lags over which $\hat{K}(l\Delta)$ can be estimated accurately when we expand the set of available data. The overall improvement in the accuracy of our estimates when more data is used in (4.3.10) is a direct result of the fact that our estimator is consistent and should not come as a surprise. Note also that the improvement in the estimates of $k_0(i,j)$ when more data is used is relatively smaller than the corresponding improvement in the estimates of $K(l\Delta)$ for large values of l. This seems to be due to the fact that the relative improvement in the accuracy of the estimates of $K(l\Delta)$ for small l is small and that these values tend to be used repeatedly in computing $\hat{k}_0(i,j)$. Furthermore, part of the inaccuracy of the estimated values of $k_0(i,j)$ is due to both the approximation errors and the interpolation errors that occur in the process of computing $\hat{k}_0(i,j)$ via equation (4.3.11). However, as we shall see in the next section, the inaccuracy of the estimated values of $k_0(i,j)$ does not seriously affect the performance of our cylindrical harmonics retrieval algorithm.

D Alternative Estimation Procedures

Other methods for estimating $k_n(r_i,r_j)$ can also be developed. For example, one can imagine estimating $k_n(r_i,r_j)$ as a spatial average of the product of the nth order Fourier coefficients $z_n(r_i)$ and $z_n(r_j)$ computed with respect to different origins. Specifically, if we denote by $z_n(\vec{u};r_i)$ the value of the nth order Fourier coefficient process computed with respect to the origin $\vec{u}$, e.g.

$$z_n(\vec{u};r_i) = \frac{1}{2\pi}\int_0^{2\pi} z(\vec{u}+\vec{r}_i)\, e^{-jn\theta_i}\, d\theta_i, \tag{4.3.8}$$

where $\vec{r}_i = (r_i,\theta_i)$, then we can compute an estimate of $k_n(r_i,r_j)$ as

$$\hat{k}_n(r_i,r_j) = \frac{1}{A}\int_A z_n(\vec{u};r_i) z_n^*(\vec{u};r_j)\, d\vec{u}, \tag{4.3.9}$$

where A denotes the area of the region $\mathcal{A}$ and where a^* denotes the complex conjugate of a.

Once more, it is a simple matter to show that $\hat{k}_n(r_i, r_j)$ is an unbiased estimate of $k_n(r_i, r_j)$. Equation (4.3.9) implies that

$$\begin{aligned} E[\hat{k}_n(r_i, r_j)] &= \frac{1}{4\pi^2 A} \int_{\mathcal{A}} d\vec{u} \int_0^{2\pi} d\theta_i \int_0^{2\pi} d\theta_j \; E[z(\vec{u}+\vec{r}_i) z(\vec{u}+\vec{r}_j)] e^{jn(\theta_i - \theta_j)} \\ &= \frac{1}{A} \int_{\mathcal{A}} d\vec{u} \; k_n(r_i, r_j) \\ &= k_n(r_i, r_j). \end{aligned} \tag{4.3.10}$$

The proof of the consistency of the estimator $\hat{k}_n(r_i, r_j)$ under the assumption that $z(\cdot)$ is Gaussian is more involved and can be found in Appendix 4.B.

In practice, it may be easier to estimate $k_n(r_i, r_j)$ via equations (4.3.2)-(4.3.3) rather then via equation (4.3.9) because of the fact that (4.3.9) requires computing the Fourier coefficient processes $z_n(\vec{u}; r_i)$ taken around different origins. The numerical integration of highly oscillatory functions, such as the integrands in (4.3.3) and (4.3.9), is a difficult operation that has to be implemented with some care [16]. The method that we have presented in this subsection involves computing a large number of such integrals per pair (r_i, r_j) (one for each $z_n(\vec{u}; r_i)$) whereas the method of subsection 4.4.A requires computing only one such integral per pair (r_i, r_j). Furthermore, the method of this subsection requires a much denser sampling scheme.

4.4 EXAMPLES

The objective of this section is to illustrate some properties of the algorithm of Section 3 with two examples. The first example uses the synthetic data generated for Example 4.3.1 while the second example uses exact covariance values. Example 4.4.1 clearly displays the robustness of our procedure and its high resolution properties, even when relatively inaccurate estimates of $k_0(i, j)$ are used as an input to our method. Example 4.4.2 is meant to show the robustness of our algorithm in the presence of modeling errors.

Example 4.4.1

In this example we consider the data generated for Example 4.3.1 and use the resulting estimates of $k_0(i,j)$ as an input to our procedure. Recall that the exact form of $k_0(i,j)$ for the data of Example 4.3.1 is

$$k_0(i,j) = 10J_0(0.1i)J_0(0.1j) + 10J_0(0.3i)J_0(0.3j) + \frac{3}{2\pi i}\delta_{i,j}, \tag{4.4.1}$$

Recall also that this corresponds to a signal with a power spectrum consisting of two cylindrical impulses at 0.1 rad/m and 0.3 rad/m respectively, and which have both an amplitude of 10 watts, i.e. the power spectrum of the signal is of the form

$$S(\lambda) = 10\frac{\delta(\lambda - 0.1)}{\lambda} + 10\frac{\delta(\lambda - 0.3)}{\lambda}. \tag{4.4.2}$$

Observe that the noise intensity in the data of Example 4.3.1 is 3 watts.m^2. Thus, the total noise power in the wave-number band [0,1] rad/m is only 0.25 dB lower than that of either cylindrical impulse. The exact power spectrum of the observations (i.e. of the signal plus noise field) is shown in Fig. 4.1.

When we used the estimates of $k_0(i,j)$, $1 \leq i,j \leq 10$, which were computed in Example 4.3.1 from the data inside the disk $0 \leq r \leq 30$, we obtained the results shown in Table 4.1 and Fig. 4.4.

Table 4.1 lists the eigenvalues of R. It is clear from this list that two eigenvalues are considerably larger than the other ones and must therefore be associated with cylindrical harmonics. However, there exists several intermediate eigenvalues which might correspond to low energy harmonics. To determine whether this is the case, we plot in Fig. 4.4 the function $v(\lambda)$ formed with the eigenvectors of R corresponding to its five smallest eigenvalues. Note that this means that our initial guess for the number of cylindrical harmonics is five. From Fig. 4.4 we see that $v(\lambda)$ has only two peaks corresponding to the presence of cylindrical harmonics at 0.134 rad/m and 0.285 rad/m, so that the intermediate eigenvalues of R do not correspond to low level harmonics. These intermediate eigenvalues can be attributed

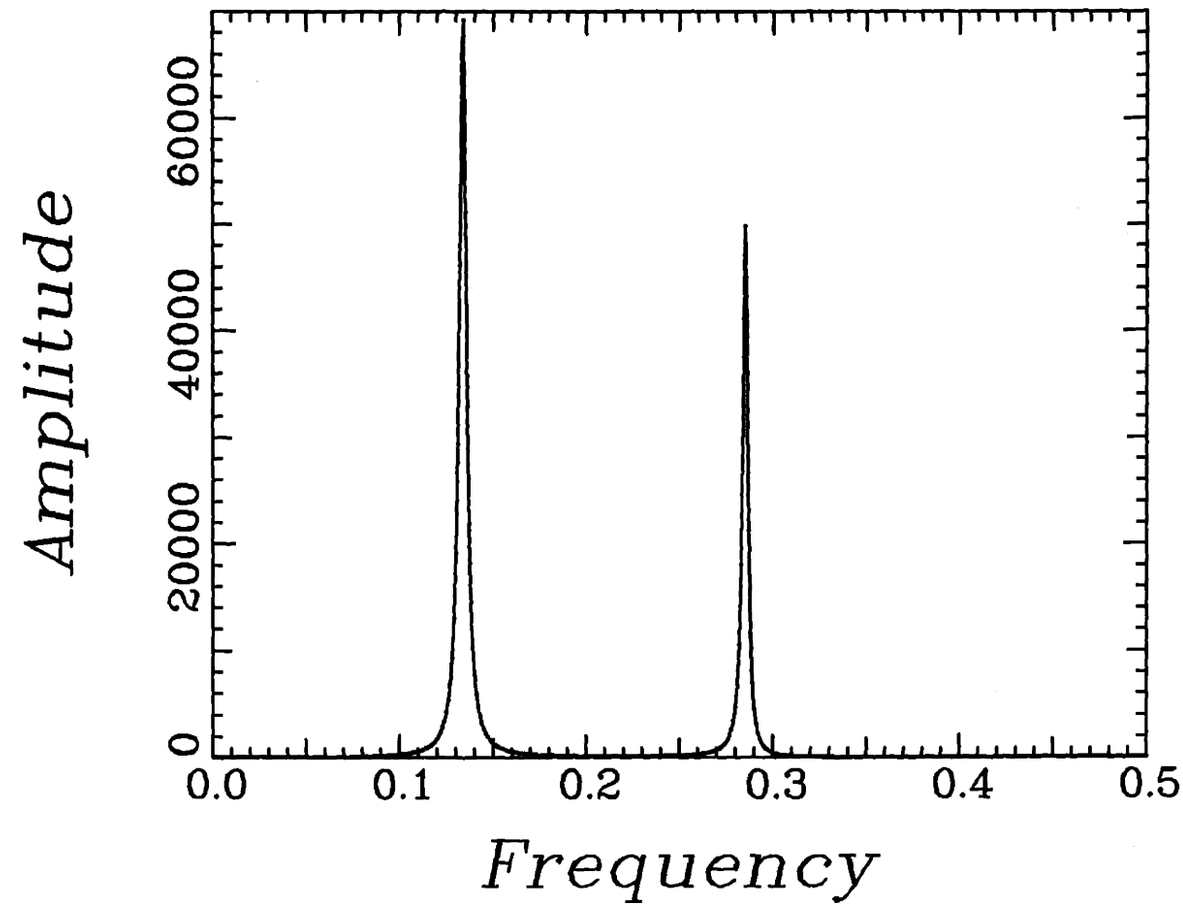

Figure 4.4: Plot of $v(\lambda)$ for Example 4.4.1 when $\hat{k}_0(i,j)$ is computed from the observations inside the disk $0 \leq r \leq 30$.

to the fact that we have used noisy estimates of the covariance $k_0(i,j)$ as an input to our algorithm. The five smallest eigenvalues of R correspond to an estimated noise intensity of 2.50 watt.m^2. Finally, the amplitudes of the cylindrical harmonics were computed via (3.17) to be 10.29 and 9.69 watts respectively. Hence, the estimated signal power spectrum is

$$S(\lambda) = 10.29\frac{\delta(\lambda - 0.134)}{\lambda} + 9.69\frac{\delta(\lambda - 0.285)}{\lambda}. \tag{4.4.3}$$

Note that our algorithm is quite robust since it performed well even though the relative error in the estimated values of $k_0(i,j)$ is relatively large, as was noted in our discussion of Example 4.3.1.

When we used the estimates of $k_0(i,j)$, $1 \leq i,j \leq 10$, which were computed in Example 4.3.1 from the data inside the disk $0 \leq r \leq 100$, we obtained the results shown in Table 4.2 and Fig. 4.5. Table 4.2 lists the eigenvalues of R and Fig. 4.5 is a plot of $v(\lambda)$. In this case, the estimated signal power spectrum that we find is given by

$$S(\lambda) = 9.62\frac{\delta(\lambda - 0.096)}{\lambda} + 10.35\frac{\delta(\lambda - 0.298)}{\lambda}. \tag{4.4.4}$$

The estimated noise strength using the values of the four smallest eigenvalues of R (see Table 4.2) is 2.98 watt.m^2. Note the improvement in the estimates in general, and particularly in the location of the cylindrical frequencies. This improvement is a direct result of the fact that we have used more accurate estimates of $k_0(i,j)$ as an input to our algorithm. In fact, the performance of our algorithm is limited only by the accuracy of the estimated values of $k_0(i,j)$. With the exact values of $k_0(i,j)$, $1 \leq i,j \leq 10$, used as an input to our procedure, the computed eigenvalues of R and a scaled-down version of the corresponding $v(\lambda)$ formed with the eigenvectors corresponding to the five smallest eigenvalues of R, are shown in Table 4.3 and Fig. 4.6 respectively. In this case the estimated noise intensity is 3 watt.m^2 and the estimated signal power spectrum is exactly equal to the actual signal power spectrum.

Finally, to demonstrate the high resolution property of our algorithm, we used a conventional spectral estimation method on the estimated values of the field covariance function $K(r)$ that were computed in Example 4.3.1 using the data inside

395.6348871
48.78233281
1.953576725
1.329499295
0.963524627
0.655533921
0.476865723
0.444997037
0.245708484
0.168982806

Table 4.1: Eigenvalues of R for Example 4.4.1 when $\hat{k}_0(i,j)$ is computed from the observations inside the disk $0 \leq r \leq 30$.

470.6285589
73.75886486
1.966603612
1.321482724
1.079630800
0.825874114
0.717595417
0.545082122
0.416322808
0.217525026

Table 4.2: Eigenvalues of R for Example 4.4.1 when $\hat{k}_0(i,j)$ is computed from the observations inside the disk $0 \leq r \leq 100$.

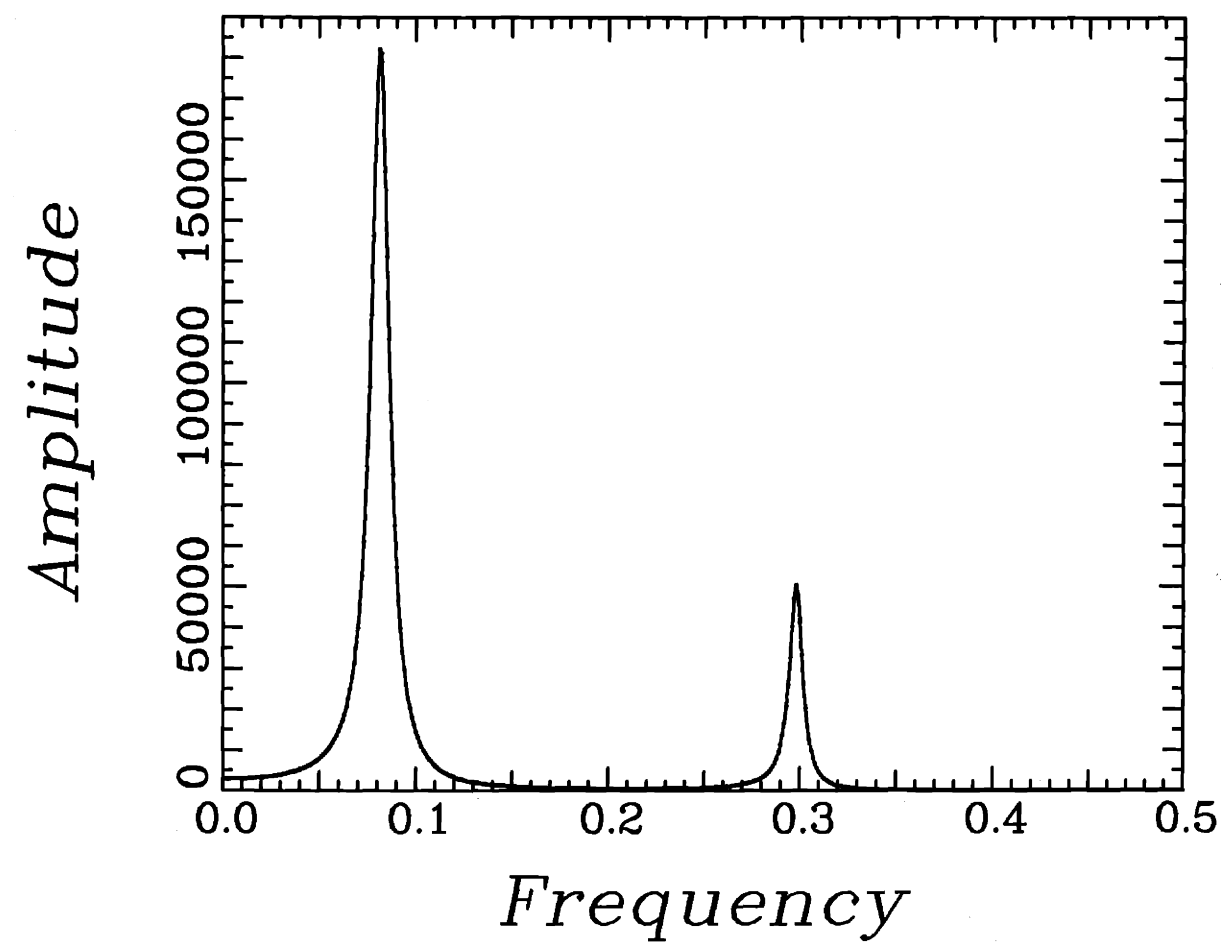

Figure 4.5: Plot of $v(\lambda)$ for Example 4.4.1 when $\hat{k}_0(i,j)$ is computed from the observations inside the disk $0 \leq r \leq 100$.

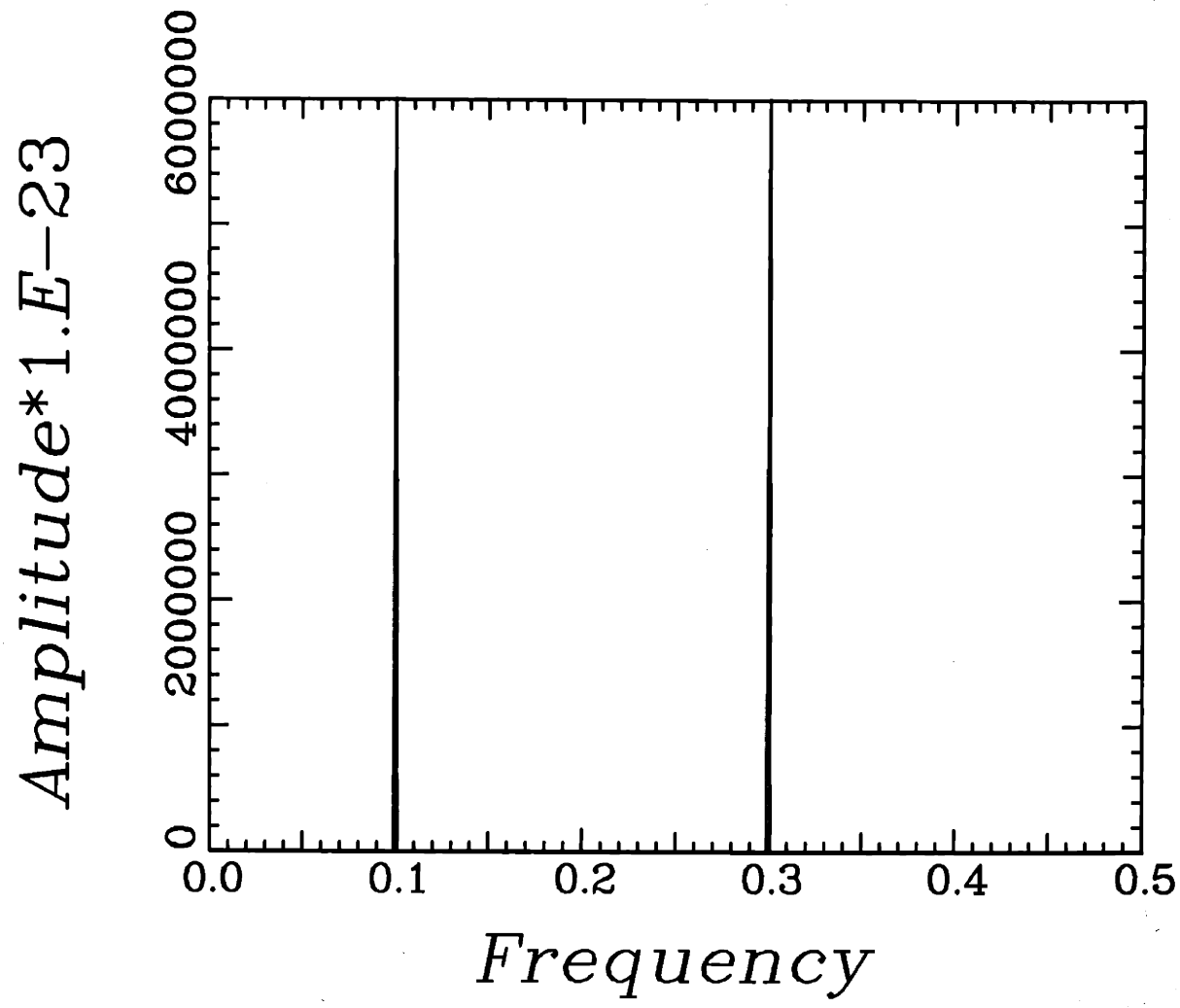

Figure 4.6: Plot of $v(\lambda)$ for Example 4.4.1 when exact values of $\hat{k}_0(i,j)$are used.

the disk $0 \leq r \leq 100$. The conventional power spectral estimate was taken to be equal to a weighted Hankel transform of the estimated field covariance function $K(r)$ [47]. The weighting function that we chose was of the form

$$w(r) = \begin{cases} 2\cos^{-1}\frac{r}{20} - \frac{r}{10}\sqrt{1-(\frac{r}{20})^2} & 0 \leq r \leq 20, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4.5}$$

The Hankel transform of this weighting function is

$$2\pi\frac{J_1^2(10\lambda)}{\lambda^2}, \tag{4.4.6}$$

and is positive for all frequencies. Hence, the expected value of the conventional power spectral estimate obtained by using this window is guaranteed to be positive [47]. The computed estimate is shown in Fig. 4.7. Note that the conventional method does not resolve the two cylindrical harmonics. This should not come as a surprise since the resolution of conventional spectral estimation is always inversely proportional to the spatial extent of the interval over which $K(r)$, or its estimate, is given, *regardless* of the choice of the window [47]. In our case an estimate of $K(r)$ was computed over the interval $[0, 20]$. This implies that the resolution of any conventional spectral estimation method is on the order of 0.3 radians/m, which is much larger than the separation between the cylindrical harmonics of (4.5.2).

Example 4.4.2

In this example we demonstrate the robustness of our algorithm with respect to modeling errors. Since we have already analyzed the effect of errors due to inaccurate covariance estimates in Example 4.4.1, we shall use exact covariance data in order to focus our attention on errors due to inaccuracies in the signal power spectrum model.

Consider a signal power spectrum of the form

$$\begin{aligned} S(\lambda) &= 10\frac{\delta(\lambda-0.1)}{\lambda} + 15\frac{\delta(\lambda-0.2)}{\lambda} + 10\frac{\delta(\lambda-0.3)}{\lambda} \\ &\quad +100(u(\lambda-0.4) - u(\lambda-0.3)), \end{aligned} \tag{4.4.7}$$

where $u(\lambda)$ is a unit step function. Note the presence in (4.5.7) of a relatively strong unmodeled colored noise component whose total power is only 1.76 dB lower than

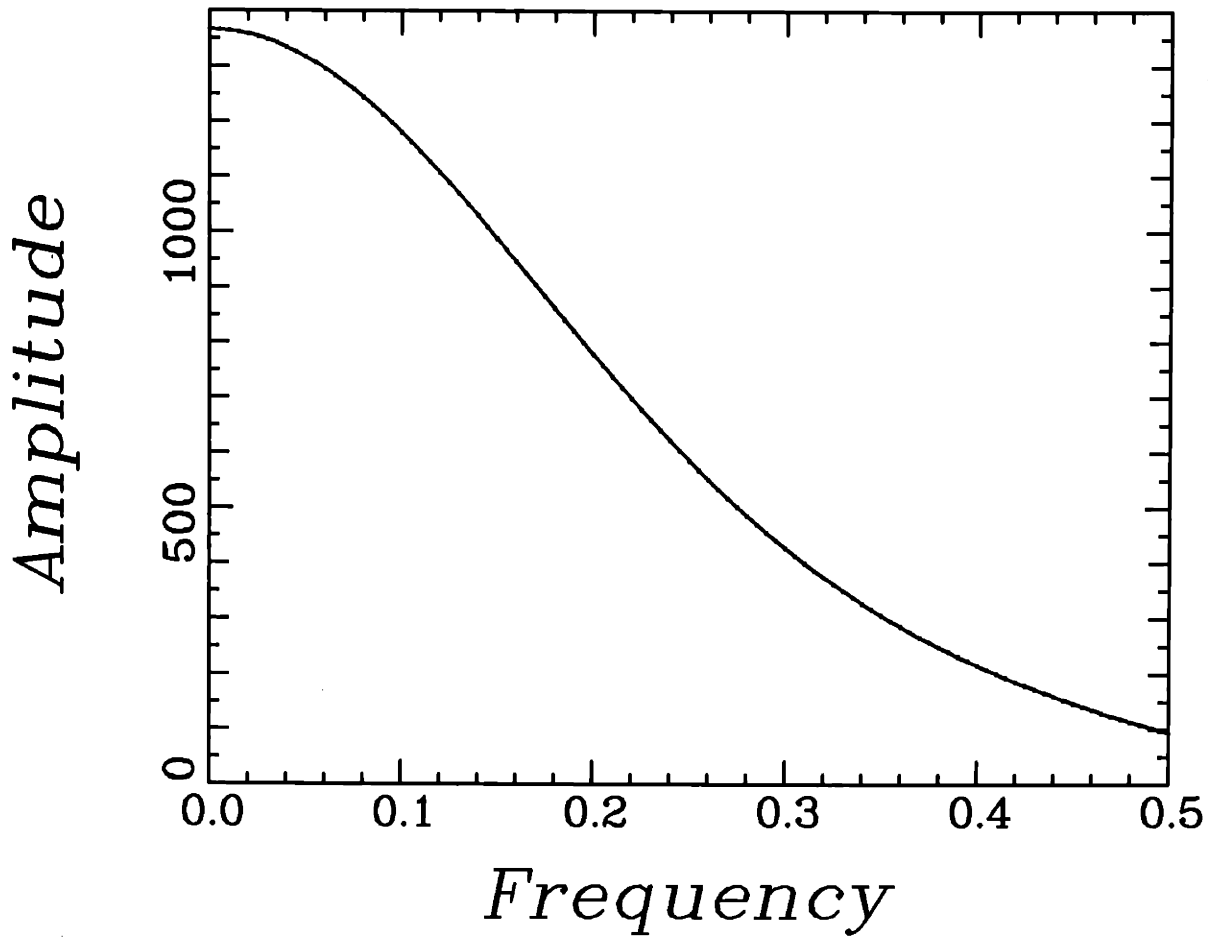

Figure 4.7: Plot of the conventional power spectral estimate of Example 4.4.1.

that of the strongest cylindrical harmonic. The measurement noise power σ^2 is taken to be 1 watt.m^2 corresponding to a total white noise power in the wave-number band [0,1] rad/m which is only 6.7 dB lower than the power of the strongest harmonic. Let us further assume that we are given exact values of $k_0(i,j)$ for $1 \leq i,j \leq 10$. The computed eigenvalues of the 10×10 symmetric matrix $R = [\sqrt{ij}k_0(i,j)]$ are listed in Table 4.4. Examination of these eigenvalues reveals the presence of four cylindrical harmonics, a fact that is confirmed by the scaled-down plot of the $v(\lambda)$ (Fig. 4.8) which is formed with the eigenvectors corresponding to the two smallest eigenvalues of R. In this example, the estimated measurement noise power is 1 watt.m^2 and the reconstructed signal power spectrum is found to be of the form

$$\begin{aligned} S(\lambda) &= 10.95\frac{\delta(\lambda-0.104)}{\lambda} + 15.1976\frac{\delta(\lambda-0.208)}{\lambda} \\ &\quad +10.95\frac{\delta(\lambda-0.31)}{\lambda} + 1.39\frac{\delta(\lambda-0.387)}{\lambda}. \end{aligned} \tag{4.4.8}$$

Note that the presence of a strong unmodeled colored noise component in the signal power spectrum has introduced a small bias in the estimated positions of the cylindrical harmonics, and has led to estimated cylindrical harmonics amplitudes which are slightly higher than their true values. Note also the presence of a spurious cylindrical harmonic at 0.387 radians/m in the reconstructed power spectrum given by (4.5.8). This spurious cylindrical harmonic is solely due to the unmodeled colored noise component. The fact that the unmodeled colored noise component gives rise to a spurious cylindrical harmonic is reminiscent of what happens in the 1-D case, since it was observed in [33] that 1-D harmonic retrieval methods which are based on an eigenanalysis of the covariance matrix do produce spurious 1-D harmonics in the presence of an unmodeled colored noise component.

The above two examples, and others, show that, overall, our algorithm is quite robust, that it has a strong resolution property and that its accuracy is really limited only by the accuracy of the estimated values of $k_0(i,j)$.

446.8880110
66.36065499
0.477464780
0.477464780
0.477464780
0.477464780
0.477464780
0.477464780
0.477464780
0.477464780

Table 4.3: Eigenvalues of R for Example 4.4.1 when exact values of $\hat{k}_0(i,j)$ are used.

732.4243282
99.73272777
0.827748745
0.160029111
0.159159233
0.159155971
0.159154672
0.159154011
0.159153889
0.159151870

Table 4.4: Eigenvalues of R in Example 4.4.2.

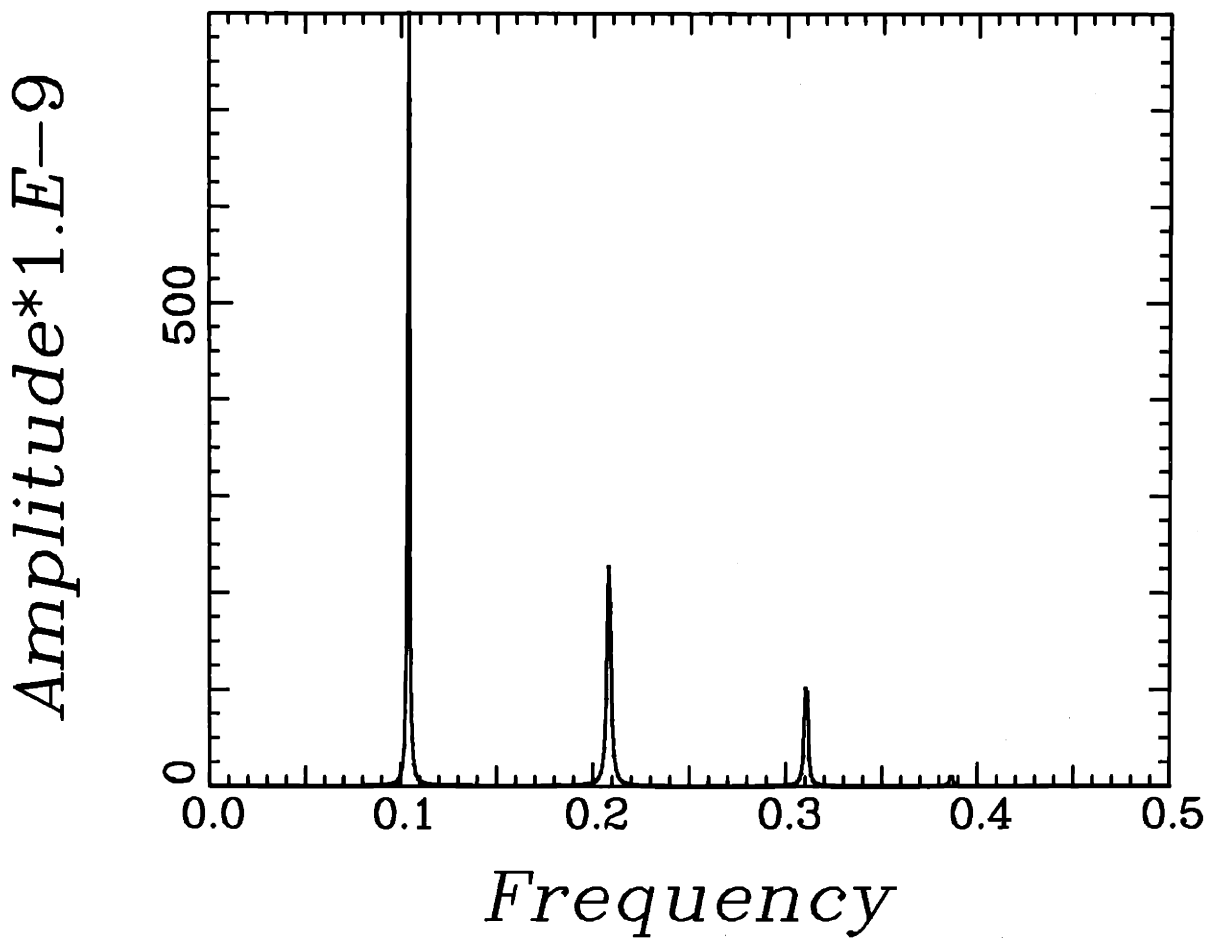

Figure 4.8: Plot of $v(\lambda)$ for Example 4.4.2.

4.5 CONCLUSION

In this chapter, we have presented a high resolution spectral estimation method for isotropic random fields with a covariance function equal to a weighted sum of cylindrical harmonics. Such fields are often used to model some types of background noise in geophysics and in ocean acoustics. The algorithm that we have obtained takes maximal advantage of the symmetries implied by the special structure of covariance functions which are equal to a weighted sum of cylindrical harmonics. Our approach is similar in spirit to the 1-D spectral estimation methods based on harmonic retrieval from an eigenanalysis of the covariance matrix. In the 2-D isotropic context, the spectral estimate is determined by performing an eigenanalysis of the covariance matrix of samples of the zeroth-order Fourier coefficient process corresponding to the given noisy observations of the underlying field. We have also discussed the estimation of this covariance matrix and presented examples to illustrate the high resolution and robustness properties of our procedure.

In the next chapter, we present a new maximum entropy spectral estimation algorithm for isotropic random fields. This algorithm can be used to estimate power spectra of the form (4.1.1.). However, as will become clear in Chapter 5, our maximum entropy spectral estimation algorithm is better suited for estimating smooth power spectra. Furthermore, it requires the estimation of a number of the nth order Fourier coefficient covariance functions $k_n(r_i, r_j)$. In contrast, the procedure of this chapter is based on the knowledge of $k_0(r_i, r_j)$ only, and is specifically designed to estimate impulsive power spectra. Hence, the method of this chapter is to be preferred if one is interested in estimating power spectra of the form (4.1.1).

APPENDIX 4.A

Proof of the consistency of $\hat{K}(r)$.

The variance of the estimator $\hat{K}(r)$ of (4.3.2) is given by

$$\mathrm{var}(\hat{K}(r)) = \frac{1}{(\pi R^{*2})^2}\int_{D_{R^*}} d\vec{s}_1 \int_{D_{R^*}} d\vec{s}_2 \qquad E[y(s_1,\theta_1)y(r+s_1,\theta_1) \\ y(s_2,\theta_2)y(r+s_2,\theta_2)] - K^2(r). \tag{4.A.1}$$

In (4.A.1) $d\vec{s_i}$ denotes the infinitesimal unit of area $d\vec{s_i} = s_i\, ds_i\, d\theta_i$ and D_{R^*} is the disk of radius R^* centered at the origin, so that $D_{R^*} = \{\vec{r}: \quad r \le R^*\}$. Using the moment factoring property of jointly Gaussian random variables, we obtain from (4.A.1)

$$\begin{aligned}\mathrm{var}(\hat{K}(r)) = \quad & \frac{1}{(\pi R^{*2})^2}\int_{D_{R^*}} d\vec{s}_1 \int_{D_{R^*}} d\vec{s}_2 \{K((s_1^2+s_2^2-2s_1s_2\cos(\theta_1-\theta_2))^{1/2}) \\ & \cdot K(((r+s_1)^2+(r+s_2)^2-2(r+s_1)(r+s_2)\cos(\theta_1-\theta_2))^{1/2}) \\ & + K((s_1^2+(r+s_2)^2-2s_1(r+s_2)\cos(\theta_1-\theta_2))^{1/2}) \\ & \cdot K(((r+s_1)^2+s_2^2-2(r+s_1)s_2\cos(\theta_1-\theta_2))^{1/2})\}.\end{aligned} \tag{4.A.2}$$

Substituting the identity [7]

$$K((s_1^2+s_2^2-2s_1s_2\cos(\theta_1-\theta_2))^{1/2}) = \sum_{n=-\infty}^{\infty} k_n(r,s)e^{-jn(\theta_1-\theta_2)} \tag{4.A.3}$$

into (4.A.2) yields

$$\mathrm{var}(\hat{K}(r)) = \sum_{n=-\infty}^{\infty} c_n(r,R^*), \tag{4.A.4}$$

where

$$\begin{aligned}c_n(r,R^*) \quad = \quad & \frac{4}{R^{*4}}\int_0^{R^*} ds_1 \int_0^{R^*} ds_2\, s_1\, s_2 k_n(s_1,s_2)k_n(r+s_1,r+s_2) \\ & + k_n(s_1,r+s_2)k_n(r+s_1,s_2),\end{aligned} \tag{4.A.5}$$

and where we have used the fact that the series in (4.A.3) is uniformly convergent to interchange the orders of summation and integration [72]. To investigate the behavior of $c_n(r, R^*)$ as R^* tends to infinity, let us substitute (2.6) into (4.A.5) to obtain

$$c_n(r,R^*) = \int_0^\infty d\lambda_1 \int_0^\infty d\lambda_2 \quad S(\lambda_1)\,S(\lambda_2)\{d_n^2(\lambda_1,\lambda_2;r,R^*) + d_n(\lambda_1,\lambda_2;r,R^*)d_n(\lambda_2,\lambda_1;r,R^*)\}, \quad (4.A.6)$$

where

$$d_n(\lambda_1,\lambda_2;r,R^*) = \frac{2}{R^{*2}}\int_0^{R^*} ds_1\, s_1 J_n(\lambda_1 s_1) J_n(\lambda_2(s_1+r)). \quad (4.A.7)$$

We now invoke the Bessel addition theorem [7]

$$J_n(\lambda(r+s)) = \sum_{k=-\infty}^{\infty} J_{n-k}(\lambda r) J_k(\lambda s) \quad (4.A.8)$$

and use the fact that (4.A.8) is a uniformly convergent series to rewrite (4.A.7) as

$$d_n(\lambda_1,\lambda_2;r,R^*) = \sum_{k=-\infty}^{\infty} J_{n-k}(\lambda_2 r) e_{n,k}(\lambda_1,\lambda_2;R^*), \quad (4.A.9)$$

where the series (4.A.9) is also uniformly convergent, and where

$$e_{n,k}(\lambda_1,\lambda_2;R^*) = \frac{2}{R^{*2}}\int_0^{R^*} ds_1\, s_1 J_n(\lambda_1 s_1) J_k(\lambda_2 s_1). \quad (4.A.10)$$

As R^* tends to infinity and for any $\lambda_1 \geq 0$ and any $\lambda_2 \geq 0$, there exist a constant $R^{**} < R^*$ such that

$$|J_n(\lambda_1 s_1) J_k(\lambda_2 s_1)| < \frac{2}{\pi s_1 \sqrt{\lambda_1 \lambda_2}} \quad \forall s_1 > R^{**}. \quad (4.A.11)$$

Furthermore, we have

$$\lim_{R^*\to\infty} e_{n,k}(\lambda_1,\lambda_2;R^*) = \lim_{R^*\to\infty}\frac{2}{R^{*2}}\int_0^{R^{**}} ds_1\, s_1\, J_n(\lambda_1 s_1) J_k(\lambda_2 s_1) + \lim_{R^*\to\infty}\frac{2}{R^{*2}}\int_{R^{**}}^{R^*} ds_1\, s_1\, J_n(\lambda_1 s_1) J_k(\lambda_2 s_2). \quad (4.A.12)$$

But

$$\left|\int_0^{R^{**}} ds_1\, s_1\, J_n(\lambda_1 s_1) J_k(\lambda_2 s_1)\right| \leq R^{**2} \tag{4.A.13}$$

since

$$|J_n(\lambda_i s_1)| \leq 1 \quad \forall s_1 \geq 0 \text{ and } \forall \lambda_i \geq 0, \tag{4.A.14}$$

and (4.A.11) implies that

$$\left|\int_{R^{**}}^{R^*} ds_1\, s_1\, J_n(\lambda_1 s_1) J_k(\lambda_2 s_1)\right| < \frac{2(R^{**} - R^*)}{\pi\sqrt{\lambda_1 \lambda_2}}. \tag{4.A.15}$$

Hence

$$\lim_{R^* \to \infty} e_{n,k}(\lambda_1, \lambda_2; R^*) = 0, \quad \forall \lambda_1, \lambda_2 \geq 0, \tag{4.A.16}$$

and since (4.A.9) is uniformly convergent, this means that

$$\lim_{R^* \to \infty} d_n(\lambda_1, \lambda_2; r, R^*) = 0 \quad \forall \lambda_1, \lambda_2 \geq 0 \text{ and } \forall r \geq 0. \tag{4.A.17}$$

If we now assume that $K(0) < \infty$ and if we use the Lebesgue dominated convergence theorem to interchange limit and integration [72], we obtain from (4.A.6) and (4.A.17)

$$\lim_{R^* \to \infty} c_n(r, R^*) = 0. \tag{4.A.18}$$

Finally, the above equation and the fact that the series (4.A.4) is uniformly convergent imply that

$$\lim_{R^* \to \infty} \text{var}(\hat{K}(r)) = 0 \tag{4.A.19}$$

which is the desired result.

APPENDIX 4.B

Proof of the consistency of the estimator (4.3.9) of $k_n(r_i, r_j)$.

Let us assume for simplicity that $\mathcal{A}$ is a disk of radius r^* and that $z(\vec{r})$ is Gaussian. The variance of the estimator (4.3.9) is given by

$$\begin{aligned}\mathrm{var}(\hat{k}_n(r_i,r_j)) &= \frac{1}{\pi^2 r^{*4}} \int_{\mathcal{A}} d\vec{u}_1 \int_{\mathcal{A}} d\vec{u}_2 E[z_n(\vec{u}_1;r_i) z_n(\vec{u}_1;r_j) z_n(\vec{u}_2;r_i) z_n(\vec{u}_2;r_j)] \\ &\quad - k_n^2(r_i,r_j), \end{aligned} \tag{4.B.1}$$

where $\vec{u}_1 = (u_1, \phi_1)$ and $\vec{u}_2 = (u_2, \phi_2)$.

In order to proceed further, we shall need an expression for the cross-correlation $E[z_n(\vec{u}_1;r_i) z_n(\vec{u}_2;r_j)]$. The addition theorem for Bessel functions [1], [7] can be used repeatedly to show that

$$\begin{aligned} J_0(\lambda|\vec{u}_1 + \vec{r}_i - \vec{u}_2 - \vec{r}_j|) = \sum_{m=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} & J_l(\lambda r_i) e^{-jl\theta_i} J_k(\lambda r_j) e^{jl\theta_j} \\ & J_{l+m}(\lambda u_1) e^{j(l+m)\phi_1} J_{k+m}(\lambda u_2) e^{-j(k+m)\phi_2}. \end{aligned} \tag{4.B.2}$$

Hence, if we write $\vec{r}_i = (r_i, \theta_i)$ and $\vec{r}_j = (r_j, \theta_j)$ we find that

$$\begin{aligned} E[z_n(\vec{u}_1;r_i) z_n(\vec{u}_2;r_j)] &= \frac{1}{4\pi^2} \int_0^{2\pi} d\theta_i \int_0^{2\pi} d\theta_j E[z(\vec{u}_1 + \vec{r}_i) z(\vec{u}_2 + \vec{r}_j)] e^{-jn(\theta_i - \theta_j)} \\ &= \frac{1}{4\pi^2} \int_0^{2\pi} d\theta_i \int_0^{2\pi} d\theta_j K(|\vec{u}_1 + \vec{r}_i - \vec{u}_2 - \vec{r}_j|) e^{-jn(\theta_i - \theta_j)} \\ &= \sum_{m=-\infty}^{\infty} k_{n,m}(r_i, r_j, u_1, u_2)\, e^{-jm(\phi_1 - \phi_2)} \end{aligned} \tag{4.B.3}$$

where

$$k_{n,m}(u,v,x,y) = \int_0^{\infty} d\lambda\, \lambda\, S(\lambda) J_n(\lambda u) J_n(\lambda v) J_m(\lambda x) J_m(\lambda y) \tag{4.B.4}$$

and where we have used the identity (4.B.2) and the fact that the series in (4.B.2) is uniformly convergent to interchange the orders of summation and integration.

If we now use the moment factoring property of jointly Gaussian random variables, we obtain from (4.B.1) and (4.B.3)

$$\begin{aligned}\mathrm{var}(\hat{k}_n(r_i,r_j)) &= \frac{4}{r^{*4}}\int_0^{r^*}\int_0^{r^*} du_1\,du_2\,u_1u_2\,\{\sum_{n=-\infty}^{\infty} k^2_{n,m}(r_i,r_j,u_1,u_2) \\ &\quad + k_{n,m}(r_i,r_i,u_1,u_2)k_{n,m}(r_j,r_j,u_1,u_2)\}. \end{aligned} \tag{4.B.5}$$

Since the series in (4.B.3) is uniformly convergent, we can interchange the orders of summation and integration, and use (4.B.5) to obtain

$$\mathrm{var}(\hat{k}_n(r_i,r_j)) = \sum_{m=-\infty}^{\infty} c_{n,m}(r^*,r_i,r_j), \tag{4.B.6}$$

where

$$\begin{aligned}c_{n,m}(r^*,r_i,r_j) &= \int_0^{\infty} d\lambda_1 \int_n^{\infty} d\lambda_2\,\lambda_1\,\lambda_2\,S(\lambda_1)S(\lambda_2)J_n(\lambda_1 r_i)J_n(\lambda_2 r_j) \\ &\quad \{J_n(\lambda_1 r_i)J_n(\lambda_2 r_j) + J_n(\lambda_2 r_i)J_n(\lambda_1 r_j)\}d^2_m(r^*,\lambda_1,\lambda_2),\end{aligned} \tag{4.B.7}$$

and where $d_m(r^*,\lambda_1,\lambda_2)$ is given by

$$d_m(r^*,\lambda_1,\lambda_2) = \begin{cases} \frac{\lambda_2 J_m(\lambda_1 r^*)J_{m-1}(\lambda_2 r^*) - \lambda_1 J_m(\lambda_2 r^*)J_{m-1}(\lambda_1 r^*)}{\pi r^*(\lambda_1^2-\lambda_2^2)} & \lambda_1 \neq \lambda_2 \\ \frac{1}{2\pi}(J^2_m(\lambda_1 r^*) + J_{m-1}(\lambda_1 r^*)J_{m+1}(\lambda_1 r^*)), & \text{otherwise.} \end{cases} \tag{4.B.8}$$

Clearly,

$$\lim_{r^*\to\infty} d_m(r^*,\lambda_1,\lambda_2) = 0. \tag{4.B.9}$$

Hence, if we use the fact that the series in (4.B.6) is uniformly convergent to interchange the operations of limit and summation, and if we use the Lebesgue dominated convergence theorem to interchange limit and integration [72], we obtain from (4.B.6), (4.B.7) and (4.B.9)

$$\lim_{r^*\to\infty} \mathrm{var}(\hat{k}_n(r_i,r_j)) = 0. \tag{4.B.10}$$

Chapter 5

A FAST MAXIMUM ENTROPY SPECTRAL ESTIMATION PROCEDURE

As mentioned in Chapter 4, the need for efficient power spectral estimation techniques arises in a number of practical applications, such as speech processing [43], radar [46], sonar [6], image processing [85] and seismic signal processing [65], to mention a few. For one-dimensional signals, the maximum entropy spectral estimation method (MEM) has become very popular due to the facts that it can provide excellent frequency resolution, and that it can be implemented in a computationally efficient way [33]. Because of the multidimensional nature of the signals arising in many applications (e.g. in geophysical problems, imaging, sonar, etc.) a number of maximum entropy algorithms have been developed over the past ten years (see Chapter 1 and [47] for references) for estimating two-dimensional spectra. These algorithms are very general and do not attempt to exploit any special structure of the power spectrum to be estimated. Since 2-D polynomials do not possess in general a quarter-plane factorization [19],[52], most of the known 2-D MEM algorithms involve solving a *non-linear* optimization problem that cannot be reduced to a linear prediction problem as in the 1-D case [47]. Furthermore, 2-D covariance functions which are positive definite on a subspace of the plane $\mathbf{R}^2$ do not necessarily

have a positive-definite extension to the whole plane [39], [70]. Thus, for any given set of stationary covariance data, the 2-D MEM problem is not guaranteed in general to have a solution. This can constitute a major problem in practice, since the covariance values that are usually used as an input to the direct 2-D MEM spectral estimation algorithms are estimates, rather than exact values, of the true covariance values, and thus may not correspond to an extendible positive-definite 2-D function. A good review of the various 2-D MEM algorithms and of the extendibility issue can be found in [47].

In this chapter we present a new *linear* MEM algorithm for 2-D isotropic random fields. An important property of isotropic covariance functions which are positive-definite over a disk is that they *always* have *positive-definite isotropic extensions* to the whole plane [71]. Here, we develop a computationally efficient linear procedure for computing the maximum entropy isotropic power spectral estimate corresponding to a covariance function that is given over a disk of radius $2R$. The maximum entropy power spectral estimate is the one that maximizes the entropy of the underlying random field. Our 2-D isotropic MEM algorithm is similar in spirit to the 1-D MEM procedure as will become clear from what follows. By using a nonsymmetric half plane spectral factorization and the properties of radially symmetric functions which are zero outside a disk in the space domain, we show that the isotropic MEM problem is equivalent to the problem of constructing the optimal linear filter for estimating the value of the underlying isotropic field at a point on the boundary of a disk of radius R given noisy observations of the field inside the disk. We then present a computationally efficient and robust procedure for computing the isotropic MEM spectral estimate. Our procedure is based on a Fourier expansion of the optimal linear estimation filter in terms of the angle θ in a polar representation of the underlying 2-D space, and on the fast recursions that were derived in [41] for solving filtering problems for isotropic random fields. These recursions are very similar to the Levinson's equations of one-dimensional prediction. The computational complexity of our procedure is $O(BRL^2)$ where B is the bandwidth in the wave-number plane of the spectrum that we want to estimate, and where L is the number of points used to discretize the interval $[0, R]$.

Note that our results show that the isotropic MEM spectral estimation problem has a linear solution. It was previously shown in the 2-D discrete space case that the MEM spectral estimation problem has a linear solution whenever the underlying field is Gauss-Markov [75]. However, there is no contradiction between our results and those of [75], since the condition of [75] is only sufficient but not necessary.

5.1 ISOTROPIC MEM SPECTRAL ESTIMATE

Consider the following spectral estimation problem. Suppose that we are given the value of the covariance function $K_y(|\vec{r}-\vec{s}|) = E[y(\vec{r})y(\vec{s})]$, of an isotropic random field $y(\vec{r})$ for $|\vec{r}-\vec{s}| \leq 2R$, and suppose that we wish to estimate the power spectrum of the "most random" *isotropic* field $y(\cdot)$ whose covariance function is consistent with the set of known values of $K_y(r)$. Furthermore, assume that $y(\vec{r})$ is given by

$$y(\vec{r}) = z(\vec{r}) + v(\vec{r}), \quad \vec{r} \in \mathbf{R}^2 \tag{5.1.1}$$

where $z(\vec{r})$ is an isotropic zero-mean Gaussian random field with a covariance function $K_z(|\vec{r}-\vec{s}|) = E[z(\vec{r})z(\vec{s})]$, and where $v(\vec{r})$ is a two-dimensional white Gaussian noise of strength P which is uncorrelated with $z(\vec{r})$.

Our problem is really that of extending a radial positive definite function given its values inside a disk of radius $2R$. It is well known [39], [70] that in general, 2-D positive definite functions defined over some finite domain, do not always have a positive definite extension on $\mathbf{R}^2$. However, it was shown in [71] that every radial positive definite function $K(r)$ defined over a disk is *extendible*. Specifically, it is proved in [71] that for every radial positive definite function $K(r)$ defined on a disk of radius $2R$ there exists *radially symmetric positive definite* functions $\hat{K}(r)$ on $\mathbf{R}^2$ such that $\hat{K}(r) = K(r)$ for $r \leq 2R$. Among all such extensions $\hat{K}_y(r)$ of $K_y(r)$, we are looking here for the one whose 2-D Fourier transform $\hat{S}_y(\vec{\lambda})$ maximizes the entropy H of the field $y(\cdot)$ where

$$\begin{aligned} H &= \frac{1}{4\pi^2}\int_{\mathbf{R}^2} d\vec{\lambda}\, \ln \hat{S}_y(\vec{\lambda}) \\ &= \frac{1}{2\pi}\int_0^\infty d\lambda\, \lambda\, \ln \hat{S}_y(\lambda), \end{aligned} \tag{5.1.2}$$

and where we have used the fact that $\hat{S}_y(\vec{\lambda}) = \hat{S}_y(\lambda)$ since $y(\cdot)$ is an isotropic random field (cf. (2.1.3)). The exact form of the power spectrum $\hat{S}_y(\lambda)$ that we seek is given in the following theorem.

Theorem 5.1 *The estimated power spectrum $\hat{S}_y(\lambda)$ which maximizes H in (5.1.2) subject to the positive definiteness constraint*

$$\hat{S}_y(\lambda) \geq 0 \quad \forall \lambda \geq 0 \tag{5.1.3}$$

and the correlation matching constraint

$$\begin{aligned} \frac{1}{2\pi}\int_0^\infty \hat{S}_y(\lambda)\, J_0(\lambda r)\, \lambda\, d\lambda &= \hat{K}_y(r) \\ &= K_y(r) \quad \textit{for } r \leq 2R \end{aligned} \tag{5.1.4}$$

is given by

$$\hat{S}_y(\lambda) = \frac{P}{|e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|^2} \tag{5.1.5}$$

where $G(R,\vec{\lambda})$ is the 2-D Fourier transform of the function $g(R,\vec{r})$ defined by the integral equation

$$K_z(|\vec{R}_0 - \vec{r}|) = \int_{u\leq R} d\vec{u}\, K_z(|\vec{r} - \vec{u}|)g(R,\vec{u}) + Pg(R,\vec{r}), \quad r \leq R \tag{5.1.6}$$

and where $\vec{R}_0 = (R,0)$.

The proof of Theorem 5.1 is based on the Lagrange multiplier method for solving constrained optimization problems [77], and on a non-symmetric half plane (NSHP) factorization that we obtain for power spectra corresponding to positive-definite radially symmetric functions that are zero outside a disk of radius $2R$ in the space domain. However, unlike in the 2-D discrete space case [19], [44], the NSHP spectral factor that we find has a bounded support in the space domain, and its spatial support is in fact a disk of radius R. The details of the proof of Theorem 5.1 can be found in Appendix 5.A.

Several comments have to be made at this point. First, note that even though $e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})$ in (5.1.5) is a function of $\vec{\lambda}$, its magnitude

$|e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|$ is by construction a function of $\lambda = |\vec{\lambda}|$ only, which is consistent with the fact that $\hat{S}_y(\lambda)$ is an isotropic power spectrum. (See Appendix 5.A for details.) Second, if we assume that $K_z(|\vec{r}-\vec{s}|) \in L_2(d\vec{r}\,d\vec{s}, [0,R]^2 \times [0,2\pi]^2)$ then the solution of (5.1.6) exists and is unique. This follows from the fact that the operator

$$\mathbf{K}: \quad a(\vec{r}) \rightarrow b(\vec{r}) = \int_{u\leq R} d\vec{u}\; K_z(|\vec{r}-\vec{u}|)a(\vec{u}) \tag{5.1.7}$$

defined over $L_2(r\,dr\,d\theta, [0,R]\times[0,2\pi])$ is self-adjoint and non-negative definite, so that $\mathbf{K}+\mathbf{I}$ is invertible. Third, observe that $g(R,\vec{r})$ is just the *optimal linear filter for estimating* $z(\vec{R}_0)$ given the observations $y(\vec{r})$ on the disk of radius R centered at the origin. Given the observations $y(\vec{r})$ of (5.1.1) for $0 \leq r \leq R$, we can express the linear least-squares estimate of $z(\vec{R}_0)$ as

$$\hat{z}(\vec{R}_0) = \int_{u\leq R} d\vec{u}\, y(\vec{u})h(\vec{u}). \tag{5.1.8}$$

Using the orthogonality principle of linear least squares estimation [32], we find that the optimal filter $h(\vec{u})$ satisfies the integral equation

$$K_z(|\vec{r}-\vec{R}_0|) = \int_{u\leq R} K_z(|\vec{r}-\vec{u}|)h(\vec{u})\,d\vec{u} + Ph(\vec{r}) \quad r \leq R. \tag{5.1.9}$$

It then follows from the uniqueness of the solution to (5.1.6) that

$$h(\vec{u}) = g(R,\vec{u}). \tag{5.1.10}$$

Hence, solving the 2-D isotropic MEM spectral estimation problem is equivalent to solving a *filtering* problem for the underlying signal field. This is analogous to the 1-D *continuous* time case where the MEM spectral estimation problem is equivalent to a filtering problem for the underlying signal process [18]. In contrast, the MEM spectral estimation problem is equivalent to a *prediction* problem for the underlying signal process [33] in the 1-D *discrete* time case. In the next section we use (5.1.10) to compute $g(R,\vec{r})$, and hence $\hat{S}_y(\lambda)$, *recursively* as a function of R via the efficient recursions developed in [41] to solve a filtering problem for 2-D isotropic random fields. The notation $g(R,\vec{r})$ (as opposed to $g(\vec{r})$) is used here to stress the dependence of the filter $g(R,\vec{r})$ on the radius $2R$ of the disk over which

$K_y(\cdot)$ is given. It is this dependence that will be exploited below to compute $g(R,\vec{r})$ recursively for increasing values of R. In this respect, our method is similar to the 1-D MEM algorithms that use the Levinson equations of 1-D prediction [33] to compute spectral estimates recursively as a function of the size of the interval over which correlation lags are given. Observe also that the choice of the point $\vec{R}_0$ in (5.1.6) is not restrictive. In fact, we can chose $\vec{R}_0$ to be any point on the boundary of the disk of radius R centered at the origin. Specifically, by using the fact that $z(\cdot)$ is an isotropic random field and the theory of [55], it is shown in the next section that $\hat{S}(\lambda)$ in (5.1.5) is invariant under rotations of the vector $\vec{R}_0$. Finally, note that as mentioned earlier, it was previously shown in the 2-D discrete space case that the MEM extension problem has a linear solution whenever the underlying field is a Gauss-Markov random field [75]. According to Theorem 5.1, the highly non-linear MEM covariance extension problem has a *linear* solution whenever the underlying field is a Gaussian *isotropic* random field regardless of whether it is Gauss-Markov or not. This is not inconsistent with the results of [75], since the condition of [75] is sufficient but not necessary. The existence of a linear solution to the isotropic MEM covariance extension problem should not come as a surprise given that the 1-D MEM stationary covariance extension problem is known to have a linear solution and that isotropy in higher dimensions is the natural extension of stationarity in 1-D. Finally, note that we have so far assumed that $K_y(r)$ is known exactly for $r \leq 2R$. In practice, one is given the observations $y(\cdot)$ over a finite disk, rather than exact values of $K_y(r)$ itself. However, $K_y(r)$ can be estimated directly from the observed data $y(\vec{r})$ by using the procedure of Section 4.3. The use of this procedure is illustrated in the second part of Example 5.4.1 where we compute MEM isotropic power spectral estimates starting from the observations $y(\cdot)$.

5.2 A FAST ALGORITHM FOR COMPUTING $\hat{S}_y(\lambda)$

In order to use (5.1.5) to compute the MEM spectral estimate $\hat{S}_y(\lambda)$, we need to know the optimal linear estimation filter $g(R,\vec{r})$ and the noise intensity P. In the 1-D discrete time case where the MEM spectral estimation problem is equivalent to a prediction problem, the constant P appearing in the numerator of the MEM spectral estimate is equal to the variance of the prediction error, and can be computed directly from the known lags of the signal covariance function [33]. In contrast, in the continuous 2-D isotropic case the constant P in (5.1.5) is equal to the intensity of the observation white Gaussian noise process and cannot be reliably computed from the known values of $K_y(r)$.

A Estimation of the Noise Intensity P

Given the measurements $y(\vec{r})$, the noise intensity P can be estimated by passing $y(\vec{r})$ through a 2-D filter whose wave-number response is zero, or almost zero, within the region of the wave-number plane that contains the spectral support of $z(\vec{r})$. The noise intensity can then be computed from the knowledge of the wave-number response of the filter and from the total power of P_f of the filtered signal. Specifically, if $F(\vec{\lambda})$ is the wave-number response of the filter then P can be estimated as

$$P = \frac{P_f}{\int_{\mathbf{R}^2} |F(\vec{\lambda})|^2 \, d\vec{\lambda}}. \tag{5.2.1}$$

Note that this approach is analogous to estimating the intensity of a 1-D additive white Gaussian noise process by passing the noisy measurements of a signal of interest through a 1-D bandpass filter followed by an output power measurement stage, with the bandpass filter specifically designed to block the signal.

B Efficient Computation of $g(R,\vec{r})$

Next, the filter $g(R,\vec{r})$ can obviously be computed by discretizing the integral equation (5.1.6) using any of the rules outlined in [16, Chapter 5] and by solving the resulting linear equation. Such an approach has two major drawbacks. The first is that it is *computationally* very *expensive* since it requires $O(M^3N^3)$ operations, where M and N are the number of discretization steps used to approximate the integral (5.1.6) in the angular variable and the radial variable respectively. Secondly, the accumulation of rounding errors and approximation errors made during the numerical computation of $g(R,\vec{r})$ and of its 2-D Fourier transform $G(R,\vec{\lambda})$ can destroy the circular symmetry of the quantity $|e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|^2$, so that the estimated power spectrum $\hat{S}_y(\lambda)$ can turn out to be *non-isotropic*. Let us now present a computationally efficient procedure for computing $\hat{S}_y(\lambda)$ that has the additional feature of guaranteeing that $\hat{S}_y(\lambda)$ is an isotropic power spectrum. As mentioned earlier, our procedure exploits the relationship between the 2-D isotropic MEM problem and a filtering problem for isotropic random fields to compute $\hat{S}_y(\lambda)$ recursively as a function of the radius $2R$ of the disk over which $K_y(r)$ is known, much in the same spirit as the 1-D MEM algorithms that compute 1-D spectral estimates recursively as a function of the number of the known covariance lags. Our approach is based on a Fourier series expansion of $g(R,\vec{r})$ in the space domain as

$$g(R,\vec{r}) = \sum_{n=-\infty}^{\infty} g_n(R,r)\, e^{jn\theta}, \tag{5.2.2}$$

and on a corresponding Fourier series expansion for $e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})$ in the wave-number plane. In the remainder of this section we shall show how to compute the coefficients $g_n(R,r)$ efficiently and then use the Hankel transform of those coefficients to compute $\hat{S}_y(\lambda)$ in a robust fashion.

B.1 Interpretation of the Fourier coefficients $g_n(R,r)$

Substituting (5.2.2) and (2.1.8) into (5.1.6) and equating the Fourier coefficients on both sides of the resulting equation yields the countably infinite set of integral

equations

$$k_n(r,R) = 2\pi \int_0^\infty k_n(r,u)\, g_n(R,u)\, u\, du + P g_n(R,r) \quad 0 \le r \le R. \tag{5.2.3}$$

Equation (5.2.3) is quite interesting because it also arises in the context of filtering for isotropic random fields [41]. In particular, the Fourier series expansions (2.1.15) are used in [41] to convert the 2-D problem of estimating the value of $z(R,\theta)$ on the boundary of a disk of radius R given the observations $y(\vec{r})$ inside the disk, into a countably infinite number of 1-D estimation problems where the objective is to estimate each of the signal Fourier coefficient processes $z_n(R)$ given the corresponding observations Fourier coefficient processes $y_n(r)$ on the interval $0 \le r \le R$. By comparing (5.2.3) with equation (2.21) of [41], it becomes clear that the coefficient $g_n(R,r)$ is only a scaled version of the optimum linear filter for estimating $z_n(R)$ given $\{y_n(s) : 0 \le s \le R\}$. Furthermore, it is shown in [41] that the optimum linear filter for estimating $z_n(R)$ given $\{y_n(s) : 0 \le s \le R\}$ obeys a quasilinear hyperbolic system [40], [3] of partial differential equations which when properly scaled take the form

$$\begin{aligned} (\frac{\partial}{\partial R} - \frac{n}{R}) g_n(R,r) + (\frac{\partial}{\partial r} + \frac{(n+1)}{r}) g_{n+1}(R,r) &= \\ -\rho_n(R) g_n(R,r) & \qquad (5.2.4) \\ (\frac{\partial}{\partial r} - \frac{n}{r}) g_n(R,r) + (\frac{\partial}{\partial R} + \frac{(n+1)}{R}) g_{n+1}(R,r) &= \\ \rho_n(R) g_n(R,r), & \qquad (5.2.5) \end{aligned}$$

with

$$\rho_n(R) = \frac{R}{2\pi}(g_n(R,R) - g_{n+1}(R,R)) \tag{5.2.6}$$

and with the initial conditions

$$\begin{aligned} \frac{\partial}{\partial r} g_0(R,r)|_{r=0} &= 0, & (5.2.7) \\ g_n(R,0) &= 0, \quad \text{for } n \neq 0. & (5.2.8) \end{aligned}$$

Note that as claimed earlier the coefficients $g_n(R,r)$, and hence the filter $g(R,\vec{r})$ and the power spectral estimate $\hat{S}_y(\lambda)$, can be computed recursively as a function of the radius $2R$ of the disk over which $K_y(r)$ is given via (5.2.4)-(5.2.5). In this respect,

equations (5.2.4)-(5.2.5) are similar to the Levinson recursions of 1-D prediction. Equations (5.2.4)-(5.2.5) can be derived by exploiting the special structure of $k_n(r,s)$ as displayed by equation (2.1.9), and by using the properties of Bessel function (see [41] for details). The numerical computation of $g_n(R,r)$ via (5.2.3)-(5.2.5) has to be performed with some care. In particular, one has to study carefully the stability and convergence properties of any numerical method used to solve the coupled partial differential equations (5.2.4)-(5.2.5) [40], [3]. In Section 5 we present a stable and convergent numerical method for computing $g_n(R,r)$. Our method is computationally very efficient and requires $O(L^2)$ operations where L is the number of discretization points in the interval $[0,R]$ where we want to compute $g_n(R,r)$.

B.2 Fourier expansions in the wave-number plane

Next, we expand $e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})$ in terms of the angle ϕ defined by $\vec{\lambda} = (\lambda, \frac{\pi}{2} - \phi)$ in a polar representation of the wave-number space. Then, by using the theory of [55, Chapter 5] and the expansion

$$e^{-j\vec{\lambda}\cdot\vec{R}_0} = \sum_{n=-\infty}^{\infty} J_n(\lambda R)\, e^{-jn\phi} \tag{5.2.9}$$

we can write

$$e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda}) = \sum_{n=-\infty}^{\infty} \left(J_n(\lambda R) - 2\pi G_n(R,\lambda)\right) e^{-jn\phi}. \tag{5.2.10}$$

In (5.2.10), $G_n(R,\lambda)$ is the nth- order Hankel transform of $g_n(R,r)$ [55], i.e.

$$G_n(R,\lambda) = \int_0^{\infty} g_n(R,r) J_n(\lambda r)\, r\, dr. \tag{5.2.11}$$

Since the magnitude of $e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})$ is a function of λ only, it follows from (5.2.10) that

$$|e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|^2 = \sum_{n=-\infty}^{\infty} |J_n(\lambda R) - 2\pi G_n(R,\lambda)|^2. \tag{5.2.12}$$

Equation (5.2.12) is a little surprising at first sight because it claims that the square magnitude of a function of the variable ϕ is equal to the sum of the square magnitudes of its Fourier coefficients in a Fourier expansion in terms of ϕ. However, the

functions that we consider here have a very special structure since their magnitude is *not* a function of ϕ by construction. An example of such functions is provided by the function $\lambda \sin\phi - j\sqrt{\lambda^2 \cos^2\phi + 1}$ whose squared magnitude $\lambda^2 + 1$ depends on λ only. Observe also that (5.2.12) implies that $e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})$ is an "all-pass" function of the variable ϕ. Finally, equation (5.2.12) can be used to show that the MEM isotropic spectral estimate $\hat{S}_y(\lambda)$ is invariant under rotations of the vector $\vec{R}_0$. Specifically, if we rotate the vector $\vec{R}_0$ by an angle θ_0 and since $K_z(\cdot)$ is invariant under rotations, then the solution to (5.1.6) is also rotated by an angle θ_0. Hence, if we denote by $g^{(\theta_0)}(R,\vec{r})$ the new solution to (5.1.6), then we have

$$g^{(\theta_0)}(R,\vec{r}) = g(R,(r,\theta+\theta_0)). \tag{5.2.13}$$

Equation (5.2.13) implies that under a rotation of the vector $\vec{R}_0$ by the angle θ_0 $\hat{S}_y(\lambda)$ is changed to $\hat{S'}_y(\lambda)$ where

$$\hat{S'}_y(\lambda) = \frac{P}{|e^{-j\vec{\lambda}\cdot\vec{R}'_0} - G^{(\theta_0)}(R',\vec{\lambda})|^2}, \tag{5.2.14}$$

where $G^{(\theta_0)}(R,\vec{\lambda})$ is the 2-D Fourier transform of $g^{(\theta_0)}(R,\vec{r})$ and where $\vec{R'}_0 = (R,\theta_0)$ is the rotated vector $\vec{R}_0$. Furthermore, it is shown in [55, Chapter 3] that if the 2-D function $f(\vec{r})$ is rotated by an angle θ_0 then its 2-D Fourier transform $F(\vec{\lambda})$ is rotated by the same angle. Hence, we can write

$$e^{-j\vec{\lambda}\cdot\vec{R}'_0} - G^{(\theta_0)}(R,\vec{\lambda}) = \sum_{n=-\infty}^{\infty} (J_n(\lambda R) - 2\pi G_n(R,\lambda))\, e^{-jn(\phi+\theta_0)}, \tag{5.2.15}$$

so that

$$|e^{-j\vec{\lambda}\cdot\vec{R}'_0} - G^{(\theta_0)}(R,\vec{\lambda})|^2 = \sum_{n=-\infty}^{\infty} |J_n(\lambda R) - 2\pi G_n(R,\lambda)|^2. \tag{5.2.16}$$

Equations (5.1.5), (5.2.12), (5.2.14) and (5.2.16) imply that

$$\hat{S'}_y(\lambda) = \hat{S}_y(\lambda) \tag{5.2.17}$$

which proves that the spectral estimate $\hat{S}_y(\lambda)$ is invariant under rotations of the vector $\vec{R}_0$.

A further simplification of (5.2.12) is possible by noting that (2.1.11) together with the uniqueness of the solution of (5.2.3) [41] imply that

$$g_n(R,r) = g_{-n}(R,r) \tag{5.2.18}$$

Hence, it follows from the fact that

$$J_{-n}(\lambda r) = (-1)^n J_n(\lambda r) \tag{5.2.19}$$

that

$$G_{-n}(R,\lambda) = (-1)^n G_n(R,\lambda). \tag{5.2.20}$$

By combining (5.2.12) and (5.2.19)-(5.2.20) we can rewrite (5.2.12) as

$$\begin{aligned} |e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|^2 &= |J_0(\lambda R) - 2\pi G_0(R,\lambda)|^2 \\ &+ 2\sum_{n=1}^{\infty} |J_n(\lambda R) - 2\pi G_n(R,\lambda)|^2. \end{aligned} \tag{5.2.21}$$

In practice, of course, one would compute only a finite number $N+1$ of the coefficient functions $G_n(R,\lambda)$ and one would obtain an approximation to the estimated power spectrum $\hat{S}_y(\lambda)$ as

$$\hat{S}_y(\lambda) \approx \frac{P}{|C_N(R,\lambda)|^2}, \tag{5.2.22}$$

where

$$\begin{aligned} |C_N(R,\lambda)|^2 &= |J_0(\lambda R) - 2\pi G_0(R,\lambda)|^2 \\ &+ 2\sum_{n=1}^{N} |J_n(\lambda R) - 2\pi G_n(R,\lambda)|^2. \end{aligned} \tag{5.2.23}$$

The number N can be determined by noting that

$$J_n(x) \approx 0 \quad \text{for } x \gg 1 \text{ and } n > x. \tag{5.2.24}$$

Hence, if we are interested in computing $\hat{S}_y(\lambda)$ over the disk $\lambda \leq B$ in the wave-number plane, we can take $N = BR$ provided that $BR \gg 1$, and in this case (5.2.22)-(5.2.23) give a very good approximation to $\hat{S}_y(\lambda)$.

Let us now make a few comments. First, note that (5.2.22)-(5.2.23) guarantee that $\hat{S}_y(\lambda)$ is isotropic since (5.2.23) involves a sum of positive terms that depend

on λ only. Second, observe that the nth-order Hankel transforms in (5.2.11) can be implemented efficiently by using any of the existing fast Hankel transform algorithms [12], [51], [26]. These techniques require $O(L \ln L)$ operations, where L is the number of discretization points at which $g_n(R,r)$ is available. Hence, our procedure for constructing $\hat{S}_y(\lambda)$ requires $O(L^2)$ per coefficient and its complexity in practice is $O(BRL^2)$ operations. Finally, note that in our procedure, the coefficients $g_n(R,r)$ are computed recursively as a function of R via (5.2.4)-(5.2.5), so that the spectral estimate $\hat{S}_y(\lambda)$ can be easily updated whenever new measurements become available, i.e. as the disk radius R is increased.

C Summary

The procedure for computing $\hat{S}_y(\lambda)$ approximately can therefore be summarized as follows:

1. Estimate $K_y(r)$ for $0 \leq r \leq 2R$ and $k_n(r,s)$ for $0 \leq r,s \leq R$ and for $|n| \leq N$, from the given data using the procedure of Section 4.3.

2. Use a stable and convergent numerical method, such as the one appearing in the next section, to compute $g_n(R,r)$ recursively from equations (5.2.3)-(5.2.5) for $n \leq N$ and for a suitably chosen N.

3. Evaluate the nth-order Hankel transforms $G_n(R,\lambda)$ by using a fast Hankel transform method.

4. Compute an approximation to $\hat{S}_y(\lambda)$ via equations (5.2.22)-(5.2.23).

5.3 NUMERICAL COMPUTATION OF THE COEFFICIENTS $g_n(R,r)$

Recall that the fast algorithm that we proposed in the last section for computing $\hat{S}_y(\lambda)$ involves the solution of the quasilinear hyperbolic system of partial differential equations (5.2.4)-(5.2.5). It is quite possible to discretize a system of partial differential equations in an apparently natural way and yet obtain completely erroneous computational results. This is especially true for propagation problems described by parabolic and hyperbolic equations. The reason for this numerical ill-behavior is that round-off and other computational errors coupled with a bad choice of discretization scheme may lead to both numerical instability and convergence problems. In this section, we present a stable and convergent method for computing $g_n(R,r)$ via (5.2.3)-(5.2.5). Our approach is based on the *method of characteristics* for solving hyperbolic partial differential equations [40], [3]. The basic idea is to replace the original system of hyperbolic partial differential equations with an equivalent system of differential equations each involving differentiation in only one of the variables of an appropriate coordinate system. The resulting system can then be solved in a well-behaved, stable and convergent manner. Specifically, let us consider a new coordinate system α, β defined by

$$\alpha = R + r \tag{5.3.1}$$

$$\beta = R - r \tag{5.3.2}$$

Equations (5.2.4)-(5.2.5) can now be rewritten in the new coordinate system as

$$\begin{aligned} \frac{\partial}{\partial\alpha}g_n(\alpha,\beta) + \frac{\partial}{\partial\alpha}g_{n+1}(\alpha,\beta) &= \left(\frac{4\alpha n}{\alpha^2-\beta^2} - \rho_n\left(\frac{\alpha+\beta}{2}\right)\right)g_n(\alpha,\beta) \\ &+ \left(\rho_n\left(\frac{\alpha+\beta}{2}\right) - \frac{4\alpha(n+1)}{\alpha^2-\beta^2}\right)g_{n+1}(\alpha,\beta) \end{aligned} \tag{5.3.3}$$

$$\begin{aligned}\frac{\partial}{\partial\beta}g_n(\alpha,\beta)-\frac{\partial}{\partial\beta}g_{n+1}(\alpha,\beta) &= -(\frac{4\beta n}{\alpha^2-\beta^2}+\rho_n(\frac{\alpha+\beta}{2}))g_n(\alpha,\beta)\\ &- (\rho_n(\frac{\alpha+\beta}{2})+\frac{4\beta(n+1)}{\alpha^2-\beta^2})g_{n+1}(\alpha,\beta)\end{aligned} \tag{5.3.4}$$

Note that in the new coordinate system each partial differential equation involves differentiation with respect to only one of the independent variables α and β. Referring to Fig. 5.1, we see that given the values of $g_n(R,r)$ and $g_{n+1}(R,r)$ on the line AB we can compute $g_n(R,r)$ and $g_{n+1}(R,r)$ within the triangle ABC by integrating (5.3.3) and (5.3.4) along the *characteristic* directions α = constant (for (5.3.4)) and β = constant (for (5.3.3)), or equivalently along lines of slope $\pm 45°$ in the (R,r) plane. Specifically, if the values of $g_n(R,r)$ and $g_{n+1}(R,r)$ have been computed inside the triangle OAB (see Fig. 5.1), and in particular on the line AB, then by integrating equation (5.3.3) along β = constant lines starting on AB, we can compute the sum $g_n(R,r)+g_{n+1}(R,r)$ inside the parallelogram ABGF. Similarly, by integrating (5.3.4) along α = constant directions starting on AB, we can compute the difference $g_n(R,r) - g_{n+1}(R,r)$ inside the region ABED. Thus, $g_n(R,r)$ and $g_{n+1}(R,r)$ can be uniquely determined within the triangle ABC (the intersection of regions ABED and ABGF). The values of $g_n(R,r)$ and $g_{n+1}(R,r)$ which are outside triangle ABC, will have to be computed using the integral equation (5.2.3). Our numerical procedure is based on equations (5.2.3) and (5.3.3)-(5.3.4). To compute $g_n(R^*,r)$ and $g_{n+1}(R^*,r)$ for $0 \le r \le R^*$, we divide the interval $[0,R^*]$ into L subintervals of length $\Delta = R^*/L$. If we denote by $G_n(k,l) = g_n(k\Delta,l\Delta)$, and if at stage k we assume that $G_n(k,l)$ and $G_{n+1}(k,l)$ have been computed for $0 \le l \le k$ (i.e. on the line AB of Fig. 5.1), then $G_n(k,l)$ and $G_{n+1}(k+1,l)$ can be evaluated for $0 < l \le k-1$ by integrating equations (5.3.3)-(5.3.4) along the characteristic directions R = constant $\pm\, r$. For $l = k, k+1$ (i.e. outside of the triangle ABC), $G_n(k+1,l)$ and $G_{n+1}(k+1,l)$ can be computed by solving a two by two linear system obtained by discretizing the integral equation (5.2.3) (see Fig. 5.1). Specifically, if we use a simple Euler difference method to integrate (5.3.3)-(5.3.4), and solving for

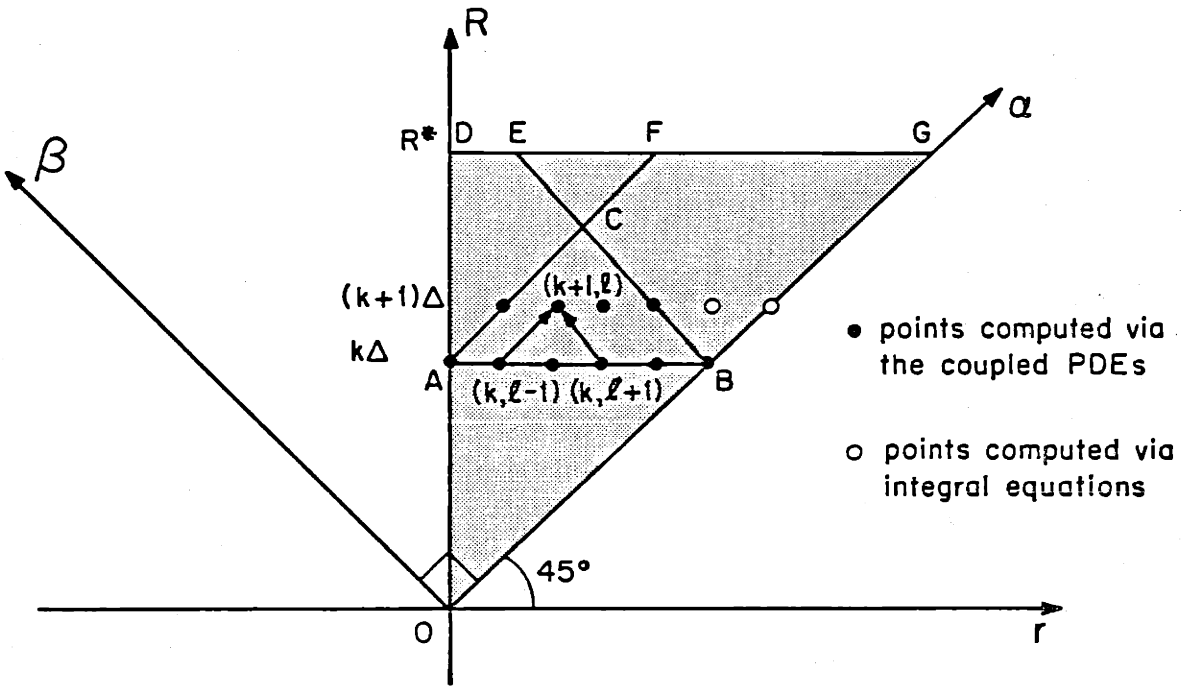

Figure 5.1: Discretization scheme and numerical implementation of the recursions for $g_n(R, r)$.

$G_n(k+1,l)$ and $G_{n+1}(k+1,l)$, we obtain the following recursions for $0 < l \leq k-1$,

$$\begin{aligned} G_n(k+1,l) = \quad & [\frac{1}{2} + \frac{n}{\sqrt{2}k} + \frac{n}{\sqrt{2}(l-1)} - \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_n(k,l-1)) \\ + \ & [\frac{1}{2} - \frac{n+1}{\sqrt{2}k} - \frac{n+1}{\sqrt{2}(l-1)} + \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_{n+1}(k,l-1)) \\ + \ & [\frac{1}{2} + \frac{n}{\sqrt{2}k} - \frac{n}{\sqrt{2}(l+1)} - \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_n(k,l+1)) \\ + \ & [-\frac{1}{2} + \frac{n+1}{\sqrt{2}k} - \frac{n+1}{\sqrt{2}(l+1)} - \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_{n+1}(k,l+1)) \end{aligned} \tag{5.3.5}$$

$$\begin{aligned} G_{n+1}(k+1,l) = \quad & [\frac{1}{2} + \frac{n}{\sqrt{2}k} + \frac{n}{\sqrt{2}(l-1)} - \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_n(k,l-1)) \\ + \ & [\frac{1}{2} - \frac{n+1}{\sqrt{2}k} - \frac{n+1}{\sqrt{2}(l-1)} + \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_{n+1}(k,l-1)) \\ + \ & [-\frac{1}{2} - \frac{n}{\sqrt{2}k} + \frac{n}{\sqrt{2}(l+1)} + \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_n(k,l+1)) \\ + \ & [\frac{1}{2} - \frac{n+1}{\sqrt{2}k} + \frac{n+1}{\sqrt{2}(l-1)} + \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_{n+1}(k,l+1)) \end{aligned} \tag{5.3.6}$$

where

$$\rho_n(k) = \frac{k\Delta}{2\pi}(G_n(k,k) - G_{n+1}(k,k)). \tag{5.3.7}$$

Similarly, if we discretize equation (5.2.3) using the trapezoidal rule, we obtain for $l = k,\ k+1$

$$\begin{aligned} PG_n(k+1,l) = k_n((k+1)\Delta, l\Delta) \ & - \ 2\pi \sum_{i=1}^{k} k_n(l\Delta, i\Delta)G_n(k+1,i)\, i\,\Delta^2 \\ & - \ \pi k_n(l\Delta,(k+1)\Delta)G_n(k+1,k+1)(k+1)\Delta^2. \end{aligned} \tag{5.3.8}$$

Other integration rules can be used as well, instead of the ones we have chosen. Note that our algorithm involves only numerical integration of ordinary differential equations and thus can be implemented in a well behaved, stable and convergent manner. Furthermore, it can be checked that this approach requires $O(L^2)$ operations per Fourier coefficient $g_n(R,r)$.

5.4 EXAMPLES

In this section we present three examples to illustrate the behavior of our 2-D isotropic MEM procedure. The first example is meant to illustrate the high resolution property of our procedure using both exact and estimated covariance data for the case of a signal power spectrum consisting of two cylindrical impulses in an additive white Gaussian noise. In particular, we use exact covariance values in the first part of this example to demonstrate the high resolution property of our algorithm and to study the effect of increasing the radius of the disk over which the covariance function is given on the spectral estimates that we obtain. In the second part of this example, we generate a random field with the desired covariance function and use the covariance estimation procedure of Chapter 4 and the method of Section 5.2 to compute MEM spectral estimates. The results that we obtain show that our procedure does not seem to suffer from the spectral line splitting problem observed in 1-D MEM spectral estimates ([33], Section 2.E). In the second example we use exact covariance data corresponding to a smooth signal spectrum to study the effect of varying the number $N+1$ of terms used in (5.2.22)-(5.2.23) to compute the 2-D isotropic MEM estimate. Finally, the third example illustrates the behavior of our algorithm when dealing with a signal that has a power spectrum consisting of both a smooth and an impulsive component.

Example 5.4.1

To demonstrate the resolution capability of our algorithm let us consider a signal power spectrum consisting of two cylindrical impulses which are spaced closer than the classical Fourier resolution limit of π/R, where $2R$ is the radius of the disk over which the covariance function is given. Specifically, consider the signal covariance function

$$K_z(r) = 10J_0(0.2r) + 10J_0(0.4r) \tag{5.4.1}$$

given over a disk of radius 20 meters. Covariance functions of the form $AK_0(Br)$ are often used in seismology [13] and in ocean acoustics [5], [11] to model some types of background noise fields. Note that $K_z(r)$ corresponds to a power spectrum

consisting of two cylindrical impulses at 0.2 rad/m and 0.4 rad/m, i.e.

$$S_z(\lambda) = 50\delta(\lambda - 0.2) + 25\delta(\lambda - 0.4). \tag{5.4.2}$$

Observe also that the separation between the two cylindrical impulses is smaller than the resolution limit of any classical spectral estimation method, which is of the order of 0.3 rad/m in this case. Furthermore, let us assume that the additive white noise intensity P is 3 Watt.m^2 . Thus the total noise power in the wave-number band [0,0.5] rad/m is 6.27 dB lower than that of either impulses. The true power spectrum of the observations (i.e. of the signal plus noise field) is shown in Fig. 5.2. Fig. 5.3, 5.4 and 5.5 show the estimated power spectra that we obtain when the order N of the highest Fourier coefficients that we use in (5.2.23) is 1, 5 and 10 respectively. Note that even with $N = 1$ we can clearly see two peaks at the correct impulse locations (Fig. 5.3). Note also that the power spectrum that we obtain with $N = 1$ is more 'peaky' than those obtained with $N = 5$ and $N = 10$ (Figs. 5.4 and 5.5). In general, we have observed that when we used only two terms in (5.2.23) the power spectra that we obtained were highly peaked and that sometimes spurious peaks tended to appear (e.g. in Example 5.4.2). Hence, even if one is interested in estimating a highly peaked spectrum consisting of pure cylindrical impulses in an additive white Gaussian noise, one should use a higher value of N to make sure that none of the peaks that appear when $N = 1$ are spurious in nature. Next note that in Fig. 5.4 where we used $N = 5$, the amplitude of the peak around 0.4 rad/m is larger than that of the peak around 0.2 rad/m. By looking at the form of the true signal power spectrum (eq. (5.4.2)) one would expect the amplitude of the peak at 0.2 rad/m to be twice as large as the one of the peak at 0.4 rad/m. As seen in Fig. 5.5, this actually happens when we use a value of N which is large enough to make the approximate spectrum that we compute via (5.2.22)-(5.2.23) very close to the true MEM spectral estimate. Finally, observe that the estimated spectrum of Fig. 5.5 is relatively smooth. This is to be expected since the MEM power spectral estimate is the smoothest of all possible spectra that satisfy the correlation matching constraint. While MEM does a good job of resolving the peaks of the power spectrum of this example, one might prefer to use the method of Chapter 4

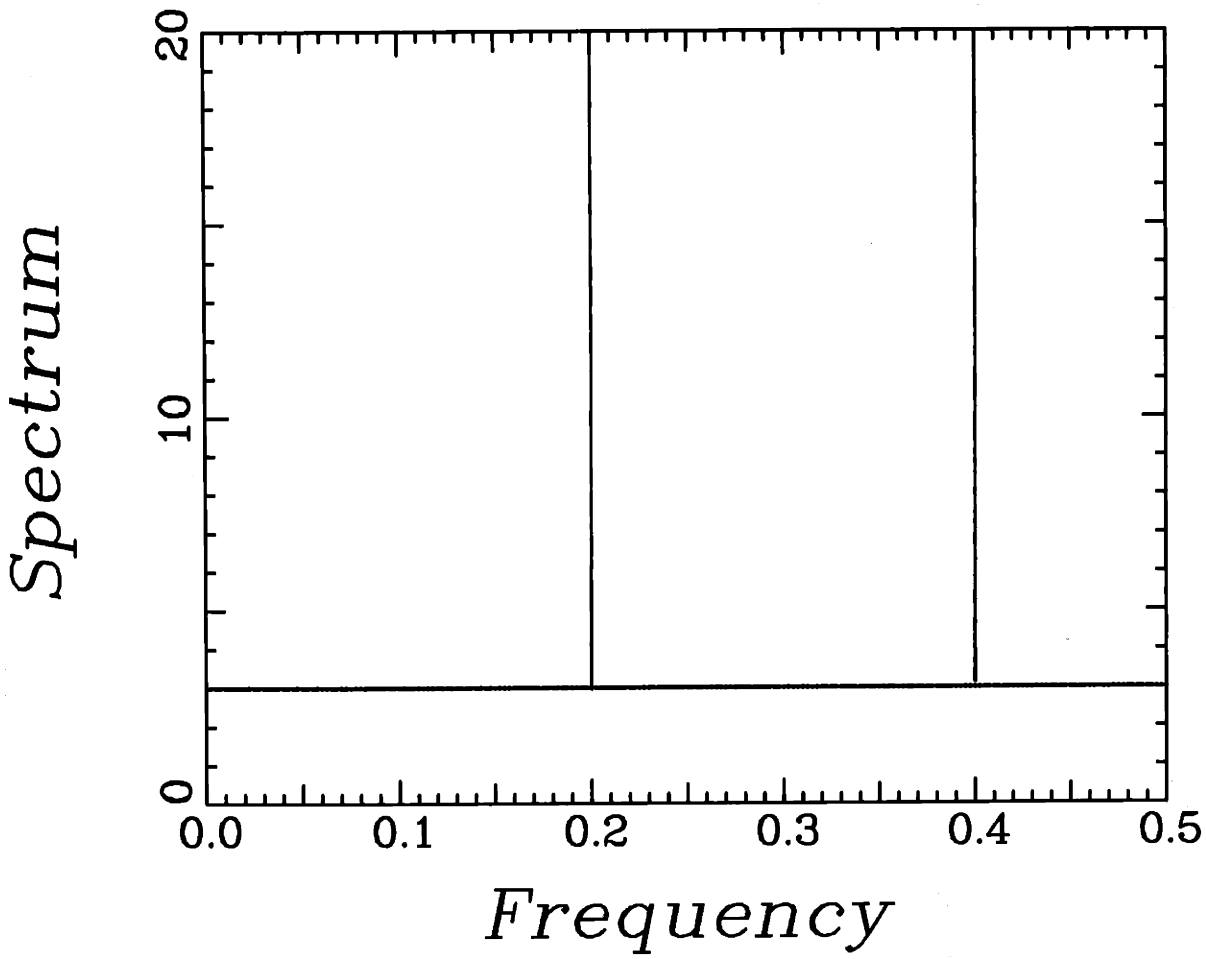

Figure 5.2: True observation power spectrum for Example 5.4.1.

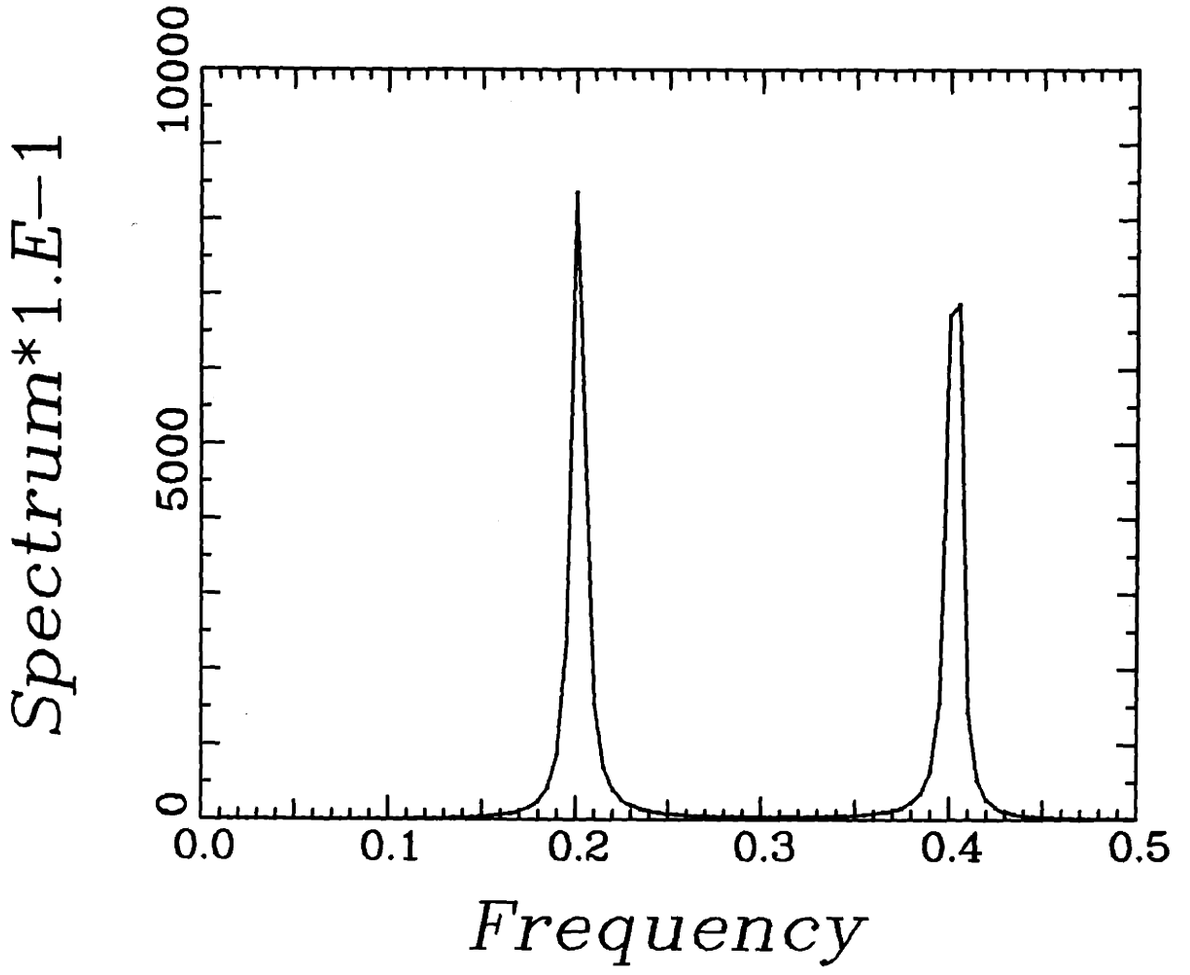

Figure 5.3: Plot of the estimated power spectrum in Example 5.4.1 when exact covariance data is given over a disk of radius 20 m and with $N = 1$.

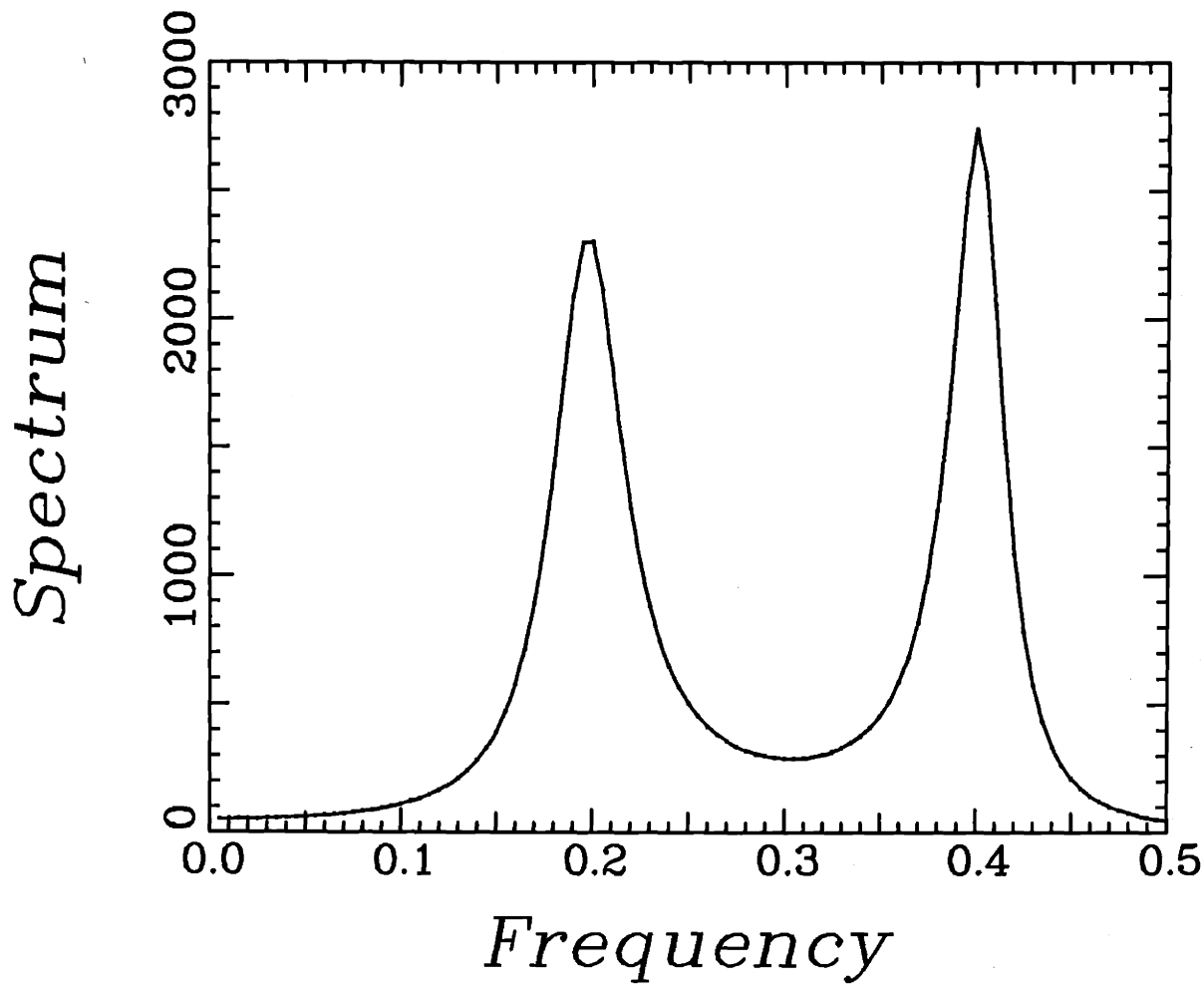

Figure 5.4: Plot of the estimated power spectrum in Example 5.4.1 when exact covariance data is given over a disk of radius 20 m and with $N = 5$.

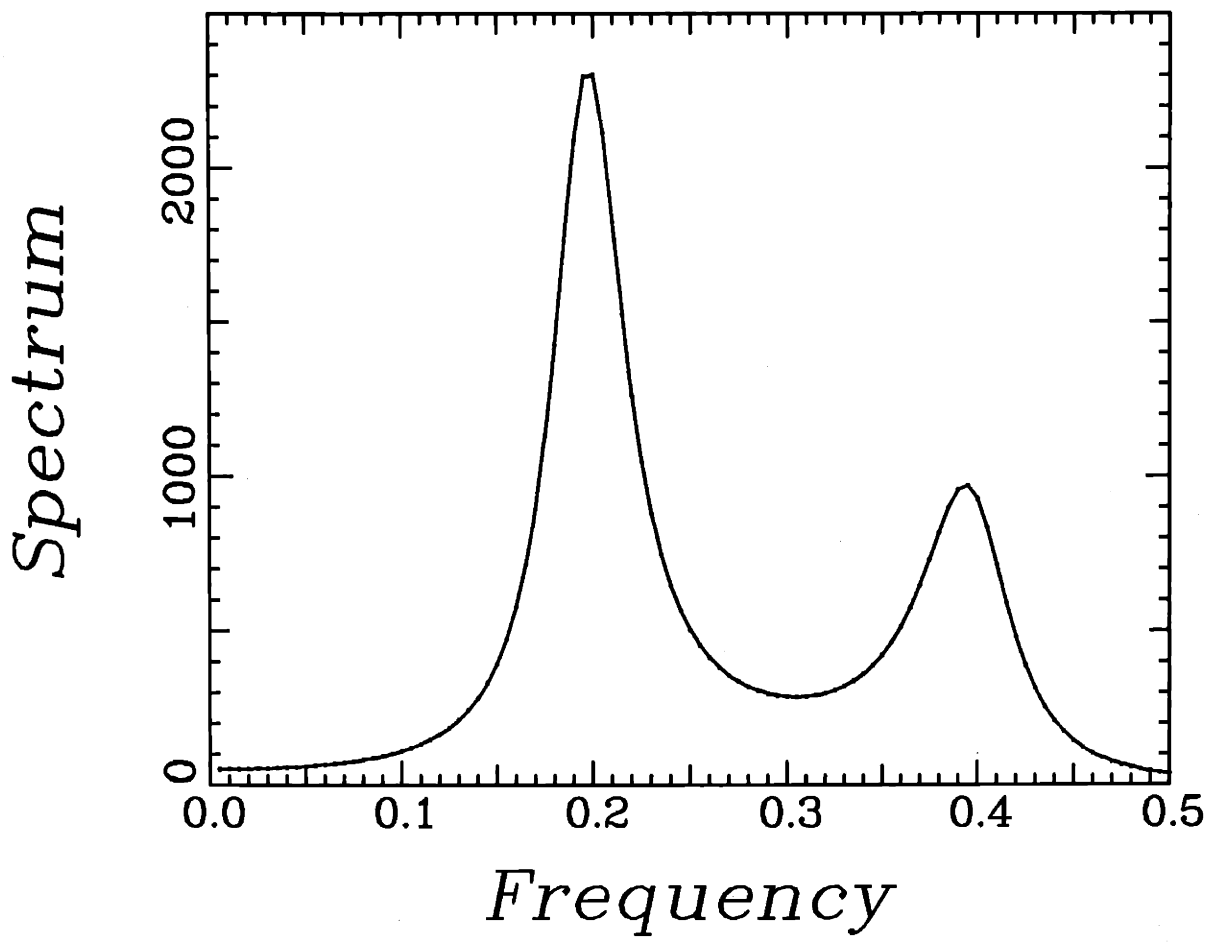

Figure 5.5: Plot of the estimated power spectrum in Example 5.4.1 when exact covariance data is given over a disk of radius 20 m and with $N = 10$.

if the spectra of interest are exclusively of the impulsive form of (5.4.2). This corresponds to using Pisarenko's method [58] or the MUSIC method [73], [9] rather than the MEM method in the 1-D case to estimate power spectra corresponding to a sum of sinusoids in a white Gaussian noise.

Next, to study the effect of the radius $2R$ of the disk over which the covariance $K_y(r)$ is assumed to be given, we double the value of $2R$ from 20 meters to 41 meters. The power spectrum that we obtain in this case using 21 terms in (5.2.23) is plotted in Fig. 5.6. Note that this spectral estimate is quite peaky and that the peak at 0.2 rad/m is twice as large as the one at 0.4 rad/m. This improvement is quite natural and in fact as $2R$ tends to infinity, one expects to be able to reconstruct the power spectrum exactly.

Finally, to study the behavior of our algorithm when data measurements, rather than exact correlation measurements, are given, we synthesized an isotropic random field with a power spectrum of the form (5.4.2) using the method described in [76]. We then added to the resulting field a white Gaussian noise field of intensity 3 Watt.m^2. Using the value of the observations $y(\cdot)$ over disks of various radii, we obtained estimates of the covariances $K_y(r)$ and $k_n(r,s)$ using the spatial averaging procedure of Section 4.3. Particular attention was given in this step to the numerical computation of an estimate of $k_n(r,s)$ via (4.3.3) to avoid the possible errors that may have resulted from the highly oscillatory nature of the integrand. In our experiments we used an integration rule based on Filon's procedure [16] to implement (4.3.3) numerically. Fig. 5.7 is a plot of the power spectrum that we obtain when we use the observations available over a disk of radius 100 m to estimate $K_y(r)$ for $0 \leq r \leq 20$ and $k_n(r,s)$ for $0 \leq r,s \leq 10$ and for $0 \leq n \leq N = 5$ and then feed those estimates as an input to our algorithm. Note the small bias in the position of the spectral peaks which are now located at 0.215 rad/m and 0.40 rad/m respectively. Fig. 5.8 shows the power spectrum that we obtain when we use the observations available over a disk of radius 30 m to estimate $K_y(r)$ for $0 \leq r \leq 20$ and $k_n(r,s)$ for $0 \leq r,s \leq 10$ and for $0 \leq n \leq N = 5$. Observe that the peak at 0.4 rad/m is now barely visible and that the peak at 0.2 rad/m is displaced to about 0.185 rad/m. This degradation in the quality of our spectral estimate is

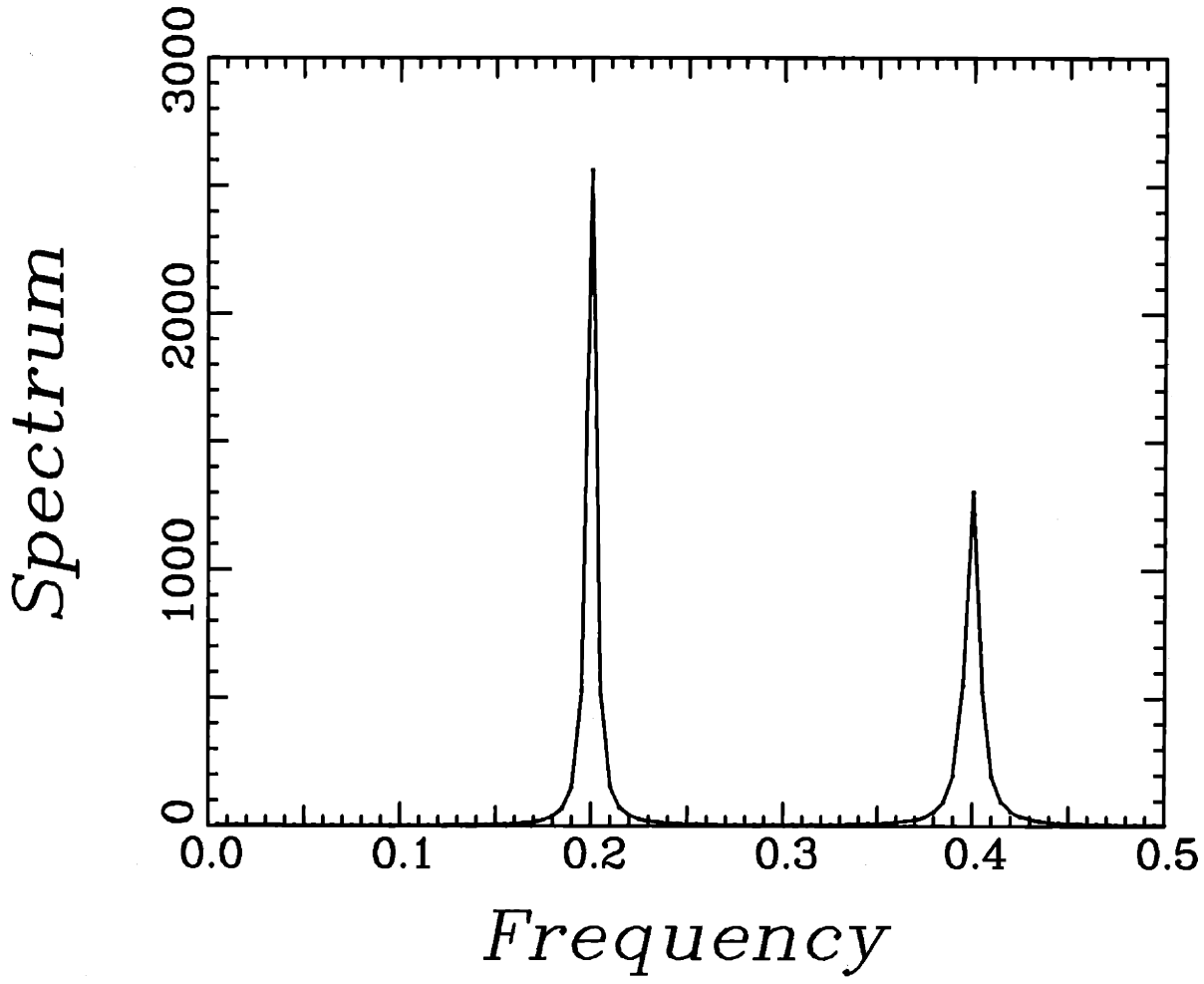

Figure 5.6: Plot of the estimated power spectrum in Example 5.4.1 when exact covariance data is given over a disk of radius 40 m and with $N = 20$.

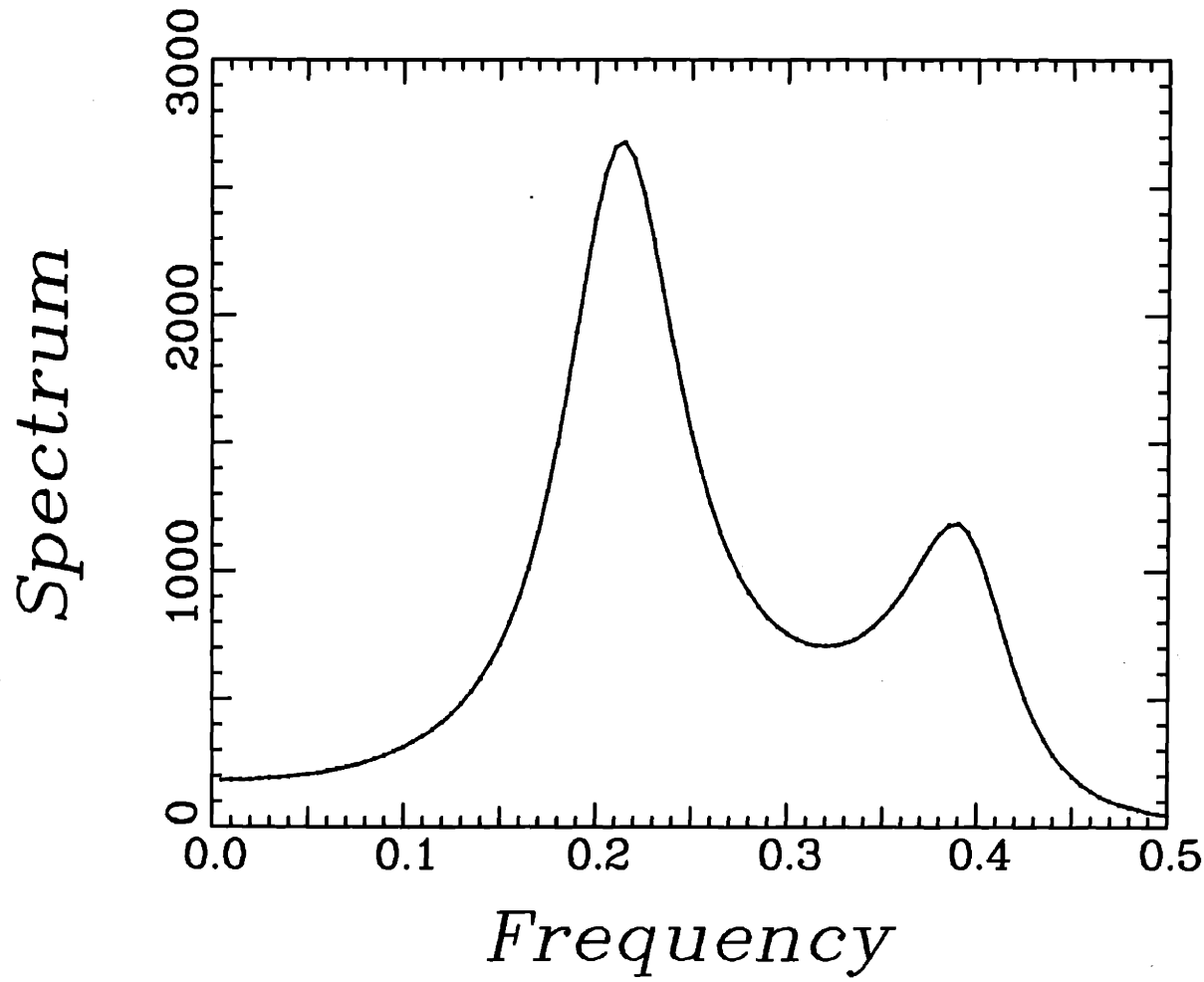

Figure 5.7: Plot of the estimated power spectrum in Example 5.4.1 when estimates of the covariance function are computed over a disk of radius 20 m given the data over a disk of radius 100 m and with $N = 5$.

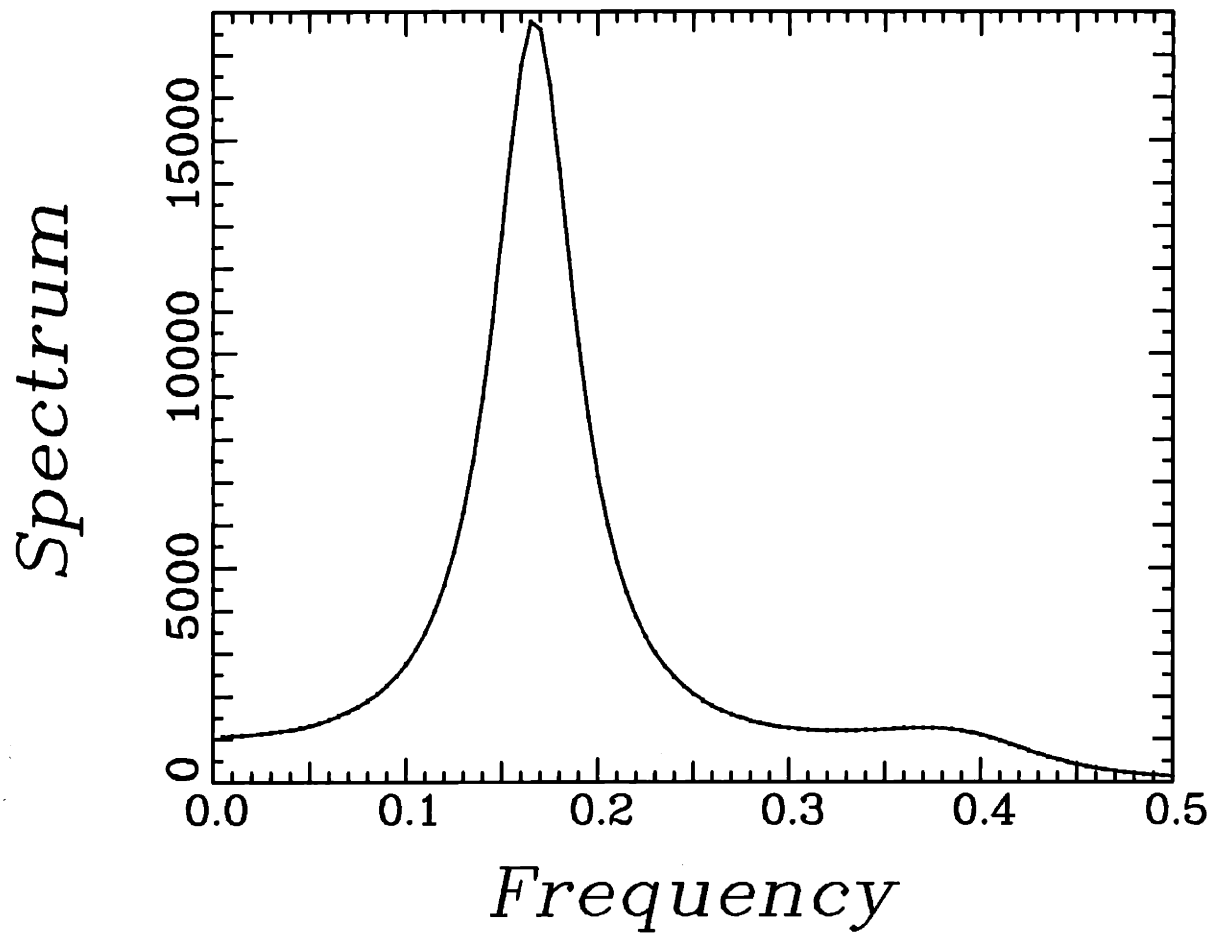

Figure 5.8: Plot of the estimated power spectrum in Example 5.4.1 when estimates of the covariance function are computed over a disk of radius 20 m given the data over a disk of radius 30 m and with $N = 5$.

not surprising since we are now using less accurate estimates of $K_y(r)$ as an input to our procedure.

We conclude this example by computing a power spectral estimate using the values of $K_y(r)$ for $0 \leq r \leq 20$ estimated from the observations inside a disk of radius 30 when the noise intensity is only 0.0001 Watt.m^2 instead of 3 Watt.m^2. Note that the total noise power in the interval [0,0.5] rad/m is now 51 dB lower than that of each of the two cylindrical impulses. The spectrum that we obtain in this case is plotted in Fig. 5.9. Note the definite presence of the two peaks which are now displaced to about 0.18 rad/m and 0.408 rad/m respectively. However, no line splitting is observed. In the 1-D case, MEM algorithms have been observed to yield two close peaks where only one is present whenever the underlying signal is a pure sinusoid with a weak additive noise component [33]. This phenomenon is called the line splitting problem and is more pronounced when the initial phase of the sinusoid is an odd multiple of $\pi/4$ and when the signal-to-noise ratio is high. In the 2-D isotropic case, there is no corresponding initial phase effect. Furthermore, our procedure is based on the computation of the filters $g_n(R, r)$ and thus conceptually involves minimizing the estimation error in all possible directions. Hence, our procedure corresponds to the 1-D MEM algorithms based on minimizing the forward and backward prediction errors [33]. Such approaches are known to alleviate the line splitting problem in the 1-D case.

Example 5.4.2

In this example we study the effect of varying the number of terms in the series (5.2.23). Consider a signal covariance function of the form

$$K_z(r) = 0.25rK_1(0.25r) \tag{5.4.3}$$

where $K_1(x)$ is a modified Bessel function of first order [7], and assume that the noise intensity is 1 Watt.m^2. Covariance functions of the form $ArK_1(Ar)$ have been used in hydrology to model the correlation structure of rainfall [67]. The power

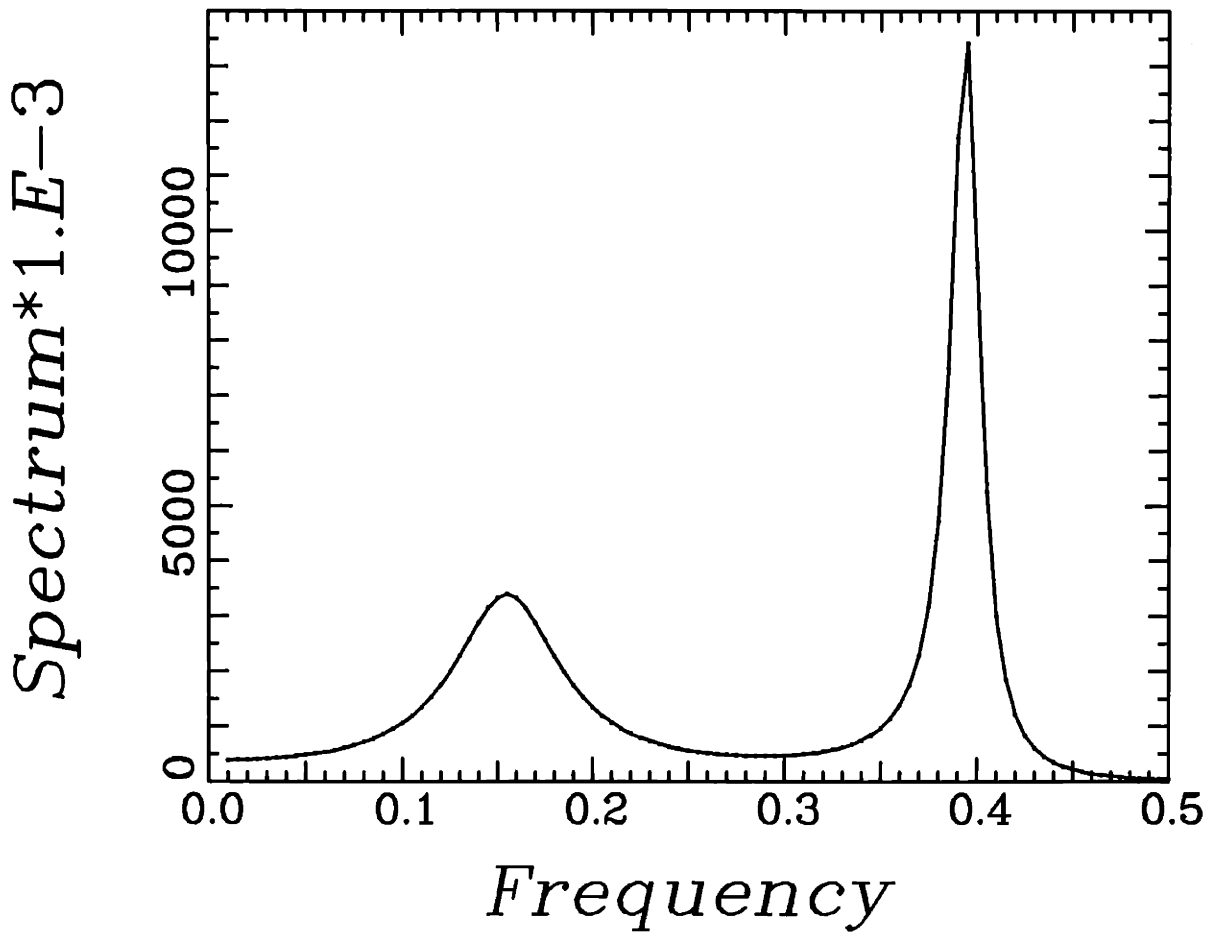

Figure 5.9: Plot of the estimated power spectrum in Example 5.4.1 when estimates of the covariance function are computed over a disk of radius 20 m given the data over a disk of radius 100 m and when $P = 10^{-3}$ Watt.m^2 and $N = 5$.

spectrum corresponding to such covariance functions is smooth and is given by

$$S_z(\lambda) = \frac{2A^2}{(\lambda^2 + A^2)^2}. \tag{5.4.4}$$

The true power spectrum corresponding to the observations (i.e. the signal plus noise fields) for this example is shown in Fig. 5.10.

To study the effect of N in (5.2.23) on the shape of the estimated power spectrum we fixed $2R$ to be 20 meters. Fig. 5.11 shows the power spectrum that we obtain when we pick $N = 1$. Note the presence of ripples in this case. Such ripples can easily be mistaken for cylindrical impulses of the type discussed in Example 5.4.1 . With $N = 3$ we obtain the power spectrum of Fig. 5.12. Note that this estimate is quite smooth. However, a spurious small and broad peak is still visible around 0.56 rad/m. If we pick $N = 10$ we obtain the power spectrum shown in Fig. 5.13. Comparing Fig. 5.10 and 13 we see that this estimate is good except around the origin of the wave-number plane. Experimental results indicate that the quality of our spectral estimates close to the origin improves with the number of discretization points used. In this example we used 100 discretization points and the quality of the spectral estimate that we obtained is good for $\lambda > 0.1$ rad/m. To get better spectral estimates close to the origin one needs to use a very large number of discretization points. For example, when we increased the number of discretization points from 100 to 150 we obtained only a slight improvement over the case that we show here. Finally, note that in this case $B = 1$ and $R = 10$ so that $N = BR = 10$.

To conclude, one should compute 2-D isotropic MEM spectral estimates via (5.2.22)-(5.2.23) by gradually increasing the number $N + 1$ of terms used until the resulting estimates stop changing noticeably as N is increased. In general, this requires computing roughly $BR + 1$ terms as mentioned in Section 5.2.

Example 5.4.3

In this last example, we illustrate the behavior of our algorithm when the signal power spectrum consists of both a smooth component and an impulsive component.

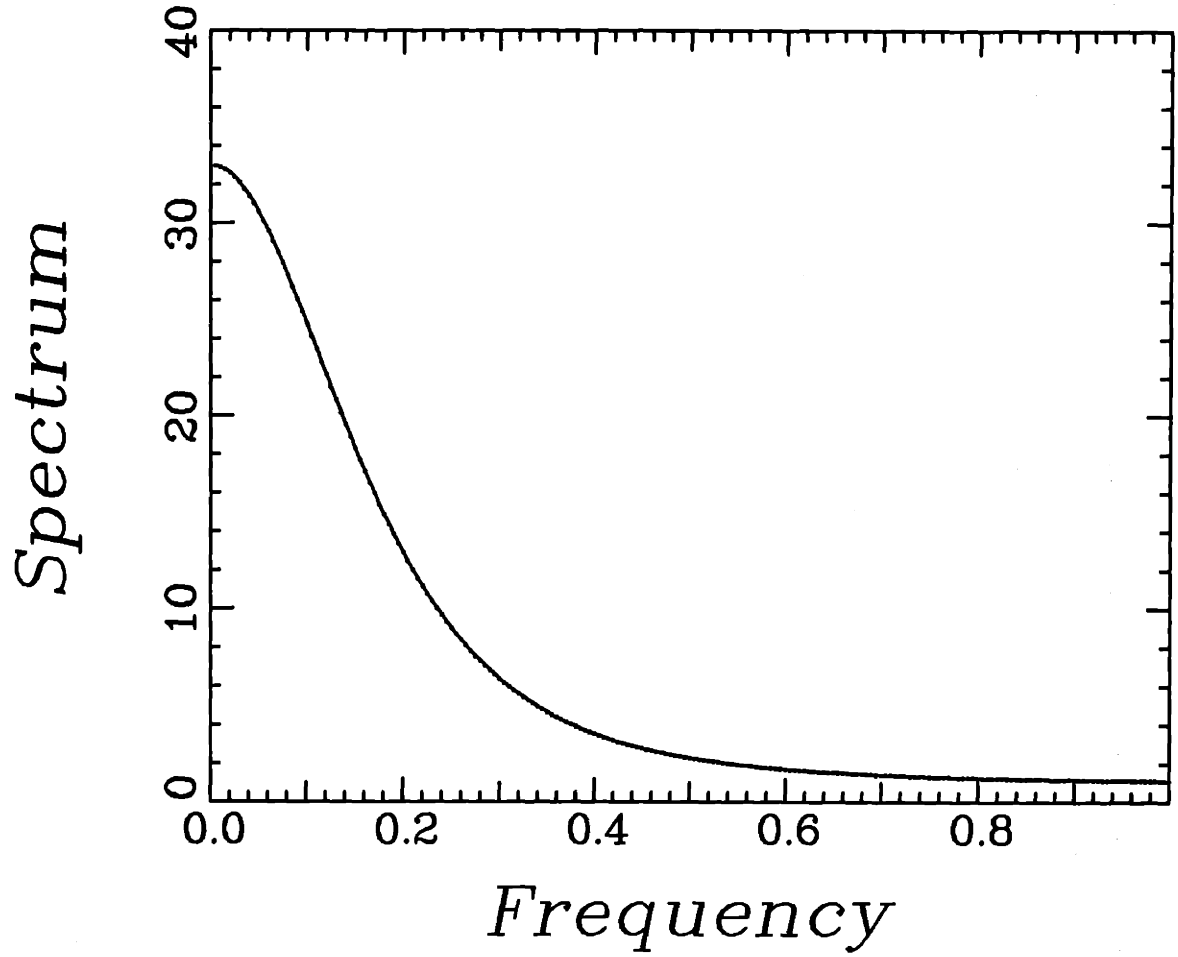

Figure 5.10: True observation power spectrum for Example 5.4.2.

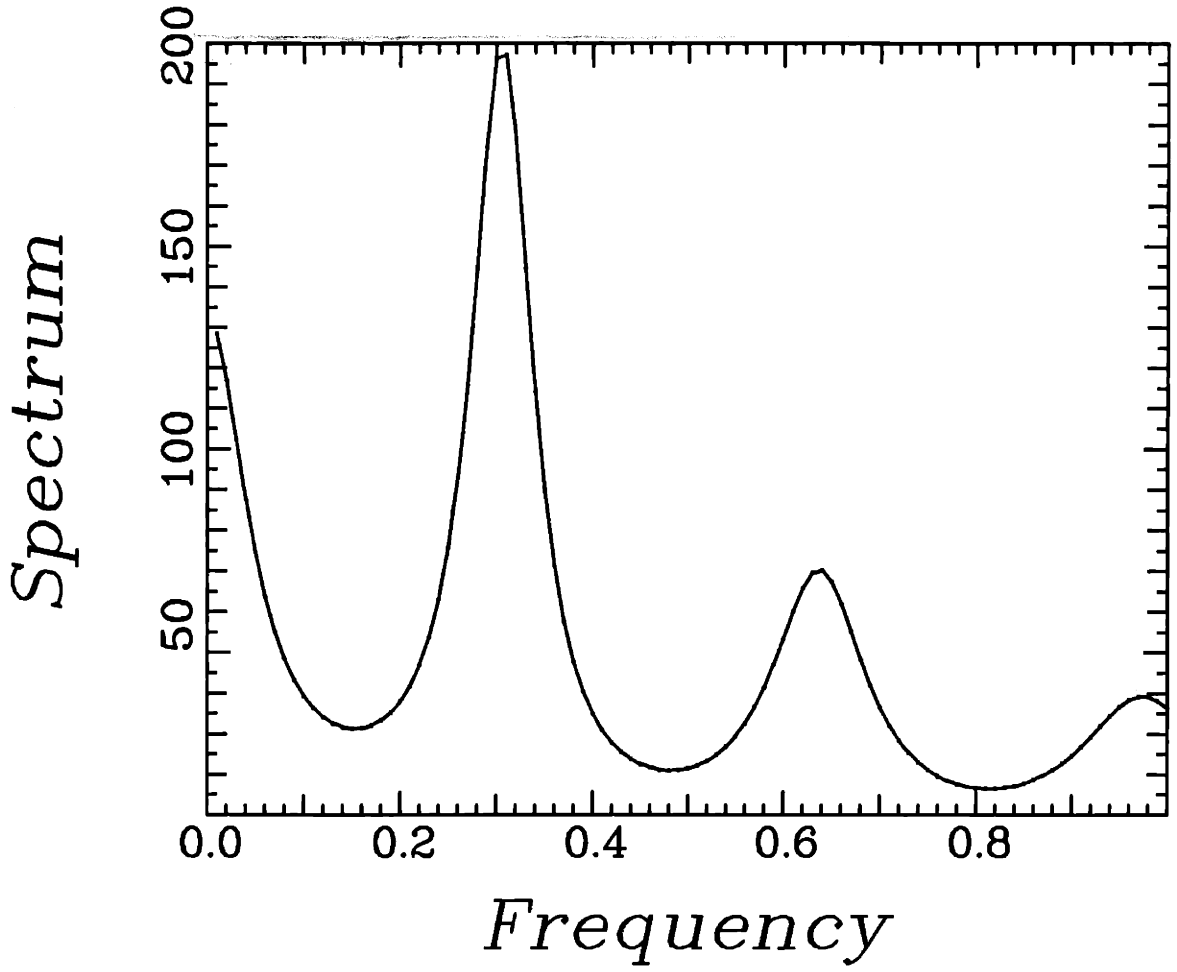

Figure 5.11: Plot of the estimated power spectrum in Example 5.4.2 when exact covariance data is given over a disk of radius 20 m and with $N = 1$.

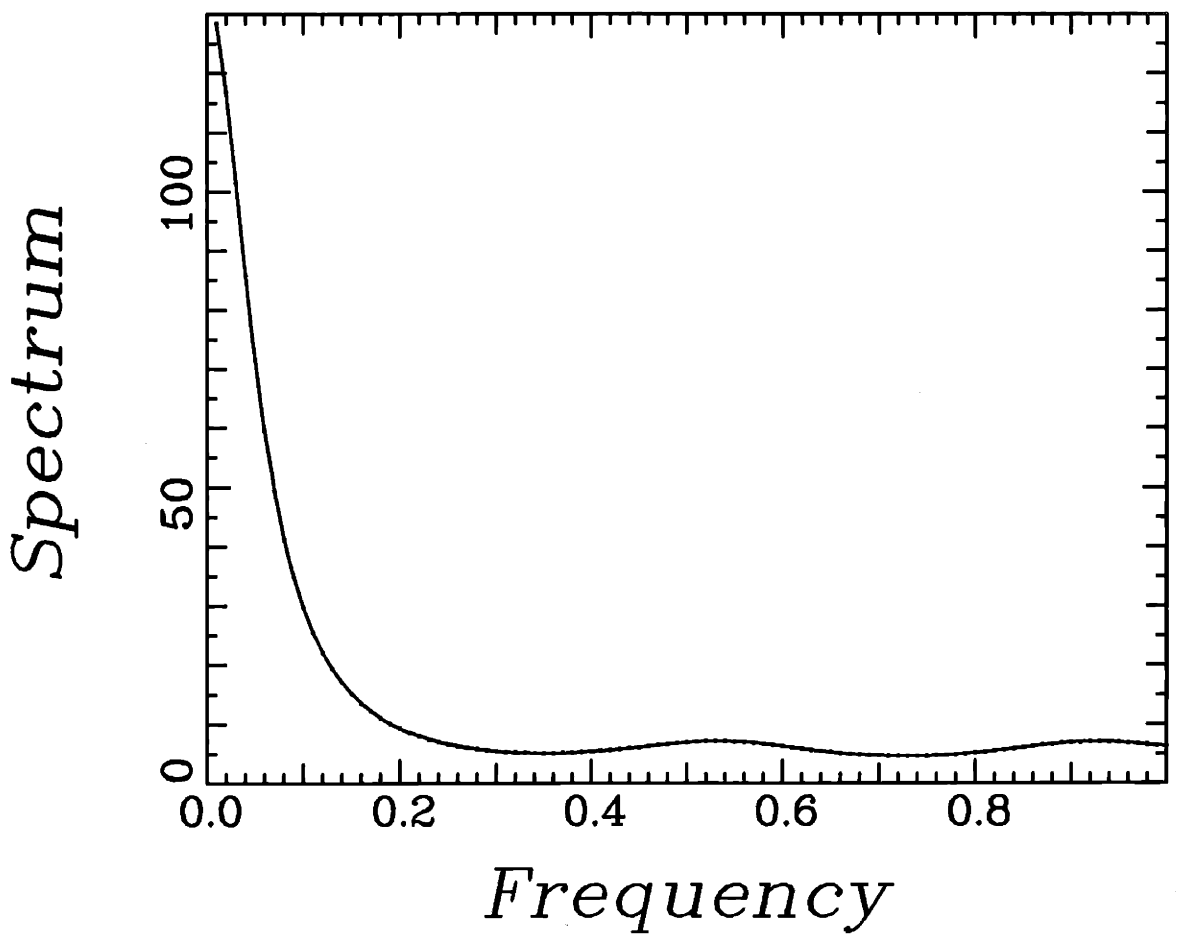

Figure 5.12: Plot of the estimated power spectrum in Example 5.4.2 when exact covariance data is given over a disk of radius 20 m and with $N = 3$.

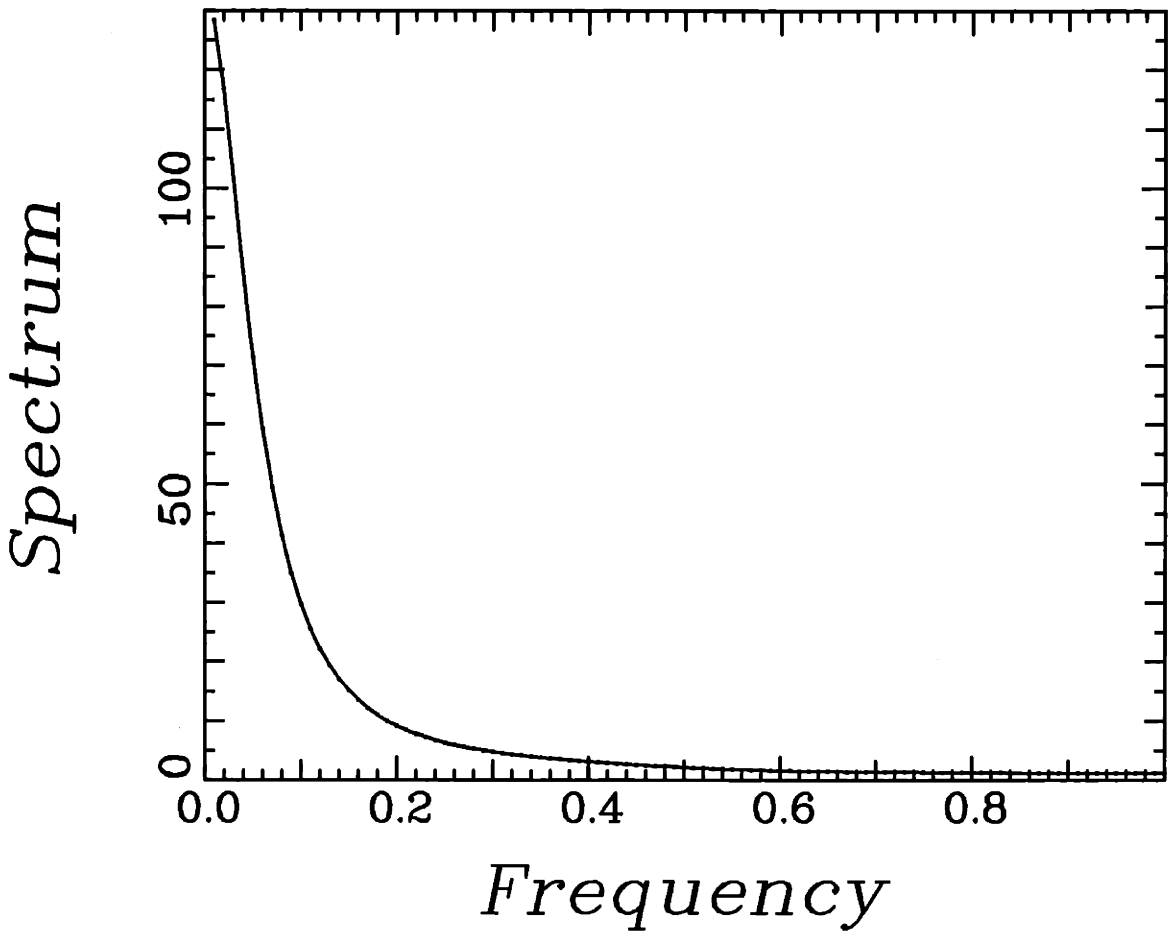

Figure 5.13: Plot of the estimated power spectrum in Example 5.4.2 when exact covariance data is given over a disk of radius 20 m and with $N = 10$.

Assume that we are given the values of the observations covariance function

$$K_y(r) = 0.25rK_1(0.25r) + AJ_0(0.6r) + \delta(\vec{r}) \tag{5.4.5}$$

over the disk $r \leq 20$. Note that the noise intensity is equal to 1 Watt.m ^{2} and that the signal power spectrum consists of both a smooth part of the form (5.4.4) and a cylindrical impulse at 0.6 rad/m. The true power spectrum of the observations is shown in Fig. 5.14.

With $A = 1$, i.e. with the total power in the smooth part of the signal power spectrum equal to that of the impulsive part, we obtain the estimate shown in Fig. 5.15 when we pick $N = 10$ in (5.2.23). Note that the presence of the impulsive component around 0.6 rad/m is barely visible.

When $A = 10$, i.e. when the total power in the impulsive component of the signal power spectrum is 10 dB higher than that of the smooth part, we obtain the estimate plotted in Fig. 5.16 when we choose $N = 10$ in (5.2.23). Observe that the presense of the impulsive component is now well marked.

In conclusion, these experimental results, and others we have obtained, indicate that the spectral estimate computed via the technique that we propose depends strongly on the size of the interval over which the observations covariance function $K_y(r)$ is known and on the accuracy of the estimates of $K_y(r)$ that are used. The number of terms that have to be used in (5.2.23) is on the order of $BR + 1$ where $2R$ is the radius of the disk over which $K_y(r)$, or its estimate is known, and where B is the bandwidth in the wave-number plane of the spectrum that we want to estimate. Finally, our procedure does not seem to suffer from the line splitting problem observed with 1-D MEM algorithms.

5.5 CONCLUSION

In this chapter, we have presented a new linear MEM algorithm for 2-D isotropic random fields. Our procedure differs from previous 2-D MEM algorithms by the fact that we take maximal advantage of the symmetries implied by isotropy which is the natural generalization to several dimensions of the 1-D notion of stationarity.

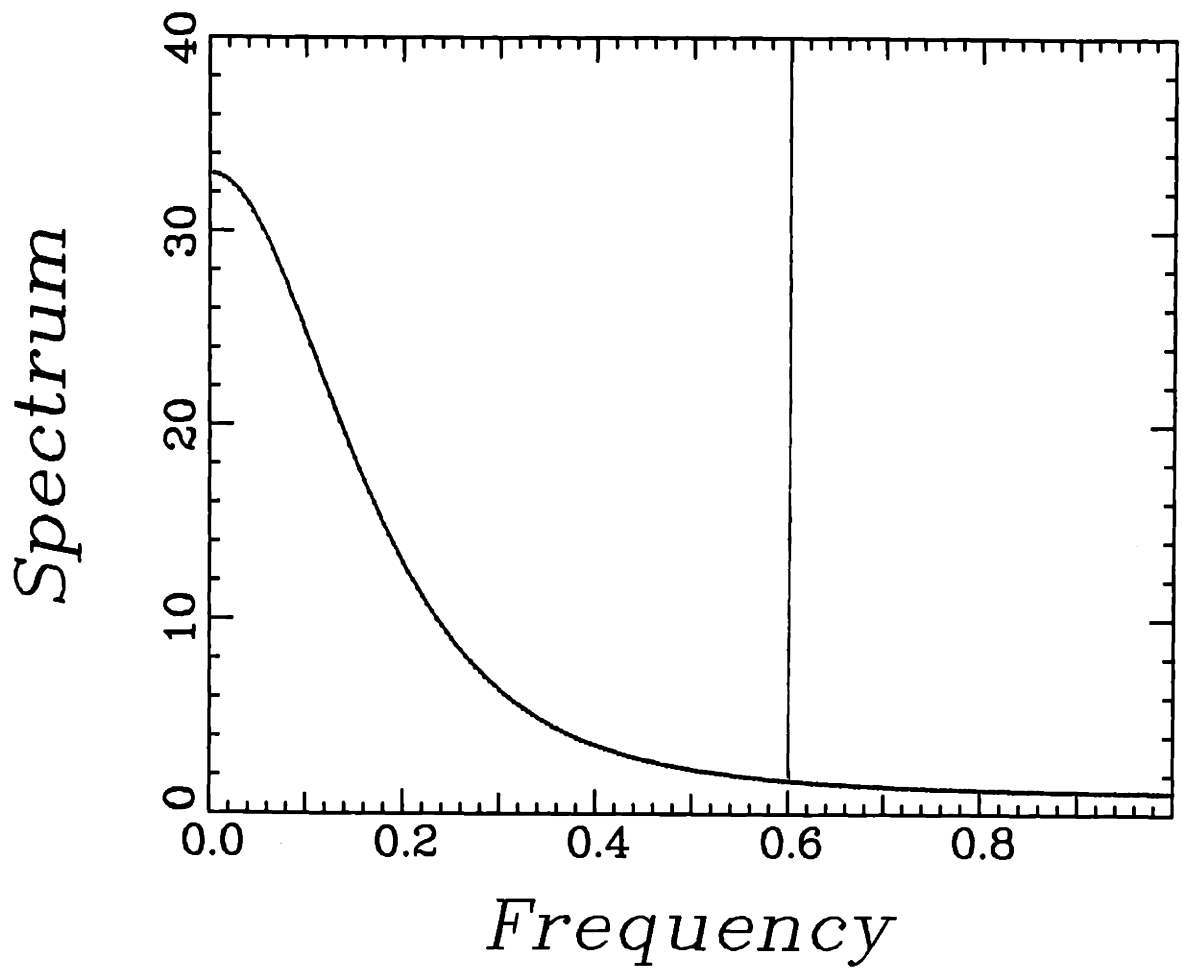

Figure 5.14: True observation power spectrum for Example 5.4.3.

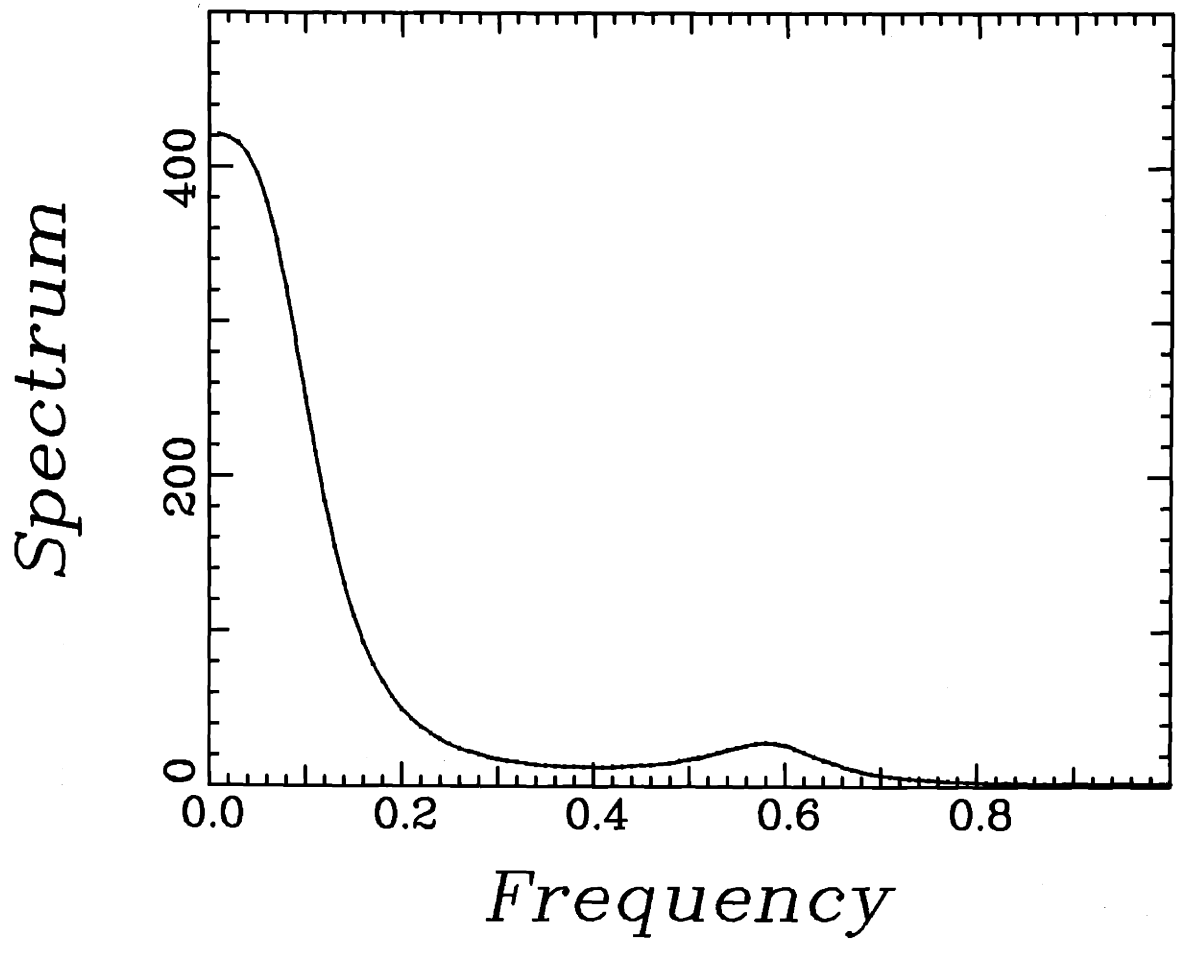

Figure 5.15: Plot of the estimated power spectrum in Example 5.4.3 when exact covariance data is given over a disk of radius 20 m and with $A = 1$ and $N = 10$.

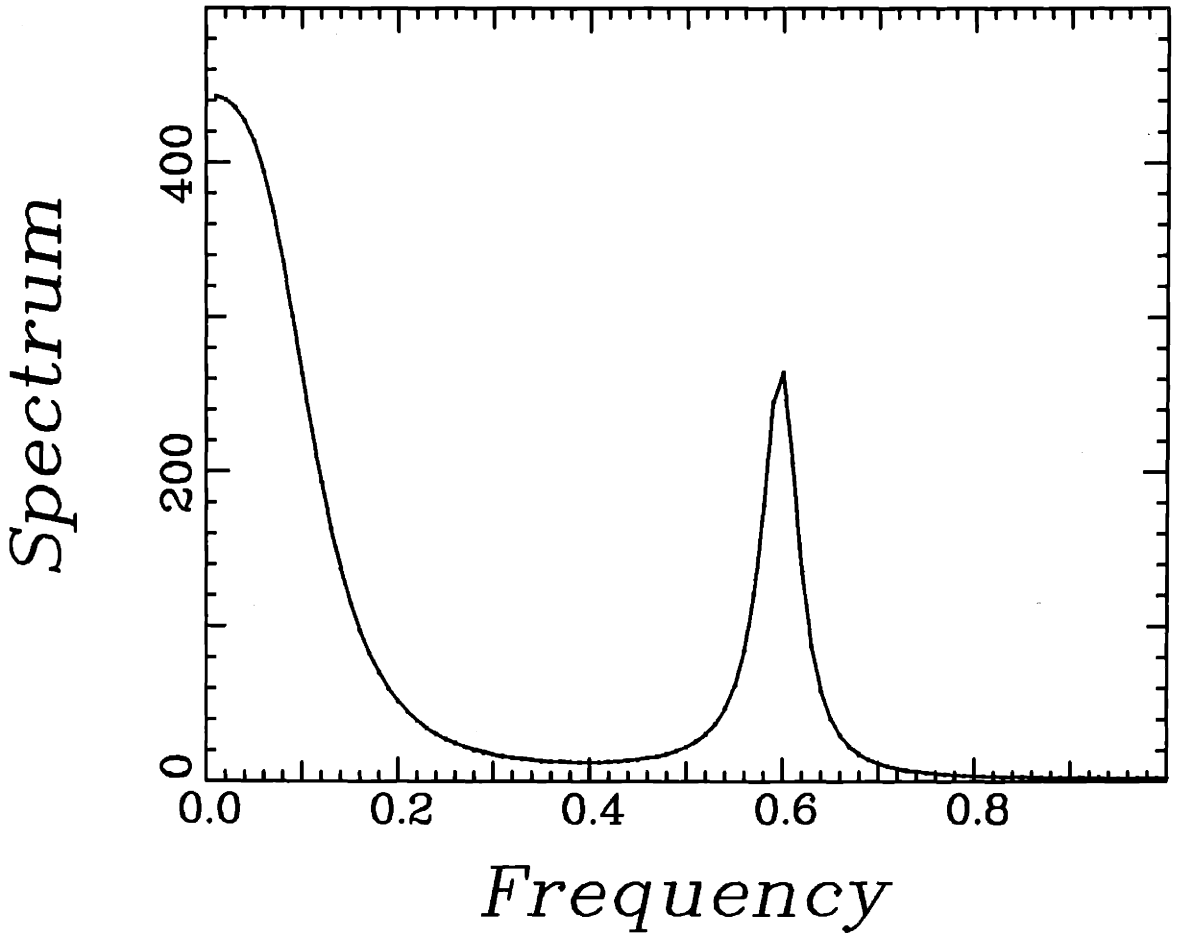

Figure 5.16: Plot of the estimated power spectrum in Example 5.4.3 when exact covariance data is given over a disk of radius 20 m and with $A = 10$ and $N = 10$.

Unlike general 2-D covariances, isotropic covariance functions which are positive definite on a disk are known to be extendible. Here, we have developed a computationally efficient procedure for computing the MEM isotropic spectral estimate corresponding to an isotropic covariance function which is given over a finite disk of radius $2R$. We have shown that the isotropic MEM problem has a linear solution which can be obtained by constructing the optimal linear filter for estimating the underlying isotropic field at a point on the boundary of a disk of radius R given noisy measurements of the field inside the disk. Our procedure is based on Fourier series expansions in both the space and wave-number domains of the inverse of the MEM spectral estimate. Furthermore, our method is guaranteed to yield a valid isotropic spectral estimate and it is computationally efficient since it requires only $O(BRL^2)$ operations where L is the number of points used to discretize the interval $[0, R]$, and where B is the bandwidth of the spectrum that we want to estimate. Finally, we have presented examples to illustrate the behavior of our algorithm and its high resolution property.

There are several directions in which one can try to generalize this work. For example, 2-D covariance functions which are constant along ellipses rather than along circles, arise in some cases of practical interest. Such covariance functions become radially symmetric under an appropriate scaling and rotation of the underlying coordinate axes, and the techniques of this chapter can then be used to estimate a warped version of the power spectrum of the underlying random field. A challenging problem here is to develop an algorithm for finding the correct scaling and rotation operations to be performed given limited measurements of the random field. More generally, another interesting problem is to extend some of the ideas that appear throughout this chapter to homogeneous, but not necessarily isotropic, covariance functions which are defined continuously over the plane. This will require a study of filtering problems for homogeneous fields aimed at determining whether the homogeneity property can be exploited in higher dimensional spaces to develop fast filtering algorithms.

APPENDIX 5.A

Proof of Theorem 5.1.

The problem that we consider in Section 5.2 is mathematically the problem of finding the $\hat{S}_y(\lambda)$ that maximizes the entropy H

$$H = \frac{1}{2\pi}\int_0^\infty d\lambda\, \lambda\, \ln S_y(\lambda), \tag{5.A.1}$$

subject to the positive definiteness and correlation matching constraints

$$\text{(i)} \quad \hat{S}_y(\lambda) \geq 0 \quad \text{for } \lambda \geq 0, \tag{5.A.2}$$

$$\text{(ii)} \quad \frac{1}{2\pi}\int_0^\infty \hat{S}_y(\lambda)\, J_0(\lambda r)\, \lambda\, d\lambda = K_y(r) \quad \text{for } r \leq 2R. \tag{5.A.3}$$

By using the approach outlined in [77] for solving optimization problems with global pointwise inequality constraints, we find that the MEM power spectral estimate $\hat{S}_y(\lambda)$ is given by

$$\hat{S}_y(\lambda) = \frac{1}{A(2R,\lambda)} \tag{5.A.4}$$

where

$$A(2R,\lambda) = \int_0^{2R} a(2R,r)\, J_0(\lambda r)\, r\, dr, \tag{5.A.5}$$

and where $a(2R,r)$ is the Lagrange multiplier function associated with the constraints (5.A.2)-(5.A.3). Observe that $A(2R,\lambda)$ can be interpreted as being the zeroth-order Hankel transform of the function $a(2R,r)$ which is zero outside the disk $r \leq 2R$. Note also that

$$\hat{S}_y(\lambda) = P + \hat{S}_z(\lambda) \tag{5.A.6}$$

where $\hat{S}_z(\lambda)$ is the estimated power spectrum of the process $z(\cdot)$. Hence, if we assume that $K_z(0)$ is finite we must have

$$\lim_{\lambda\to\infty} \hat{S}_z(\lambda) = 0, \tag{5.A.7}$$

for otherwise the integral

$$\int_0^\infty \hat{S}_z(\lambda)\,\lambda\,d\lambda = K_z(0) \tag{5.A.8}$$

would fail to converge. Taking (5.A.7) into account, we can rewrite (5.A.4) as

$$\hat{S}_y(\lambda) = \frac{P}{1 - B(2R,\lambda)} \tag{5.A.9}$$

where

$$\lim_{\lambda\to\infty} B(2R,\lambda) = 0, \tag{5.A.10}$$

and

$$B(2R,\lambda) < 1 \tag{5.A.11}$$

since $\hat{S}_y(\lambda)$ is strictly positive for all λ. Note that (5.A.4)-(5.A.5) and (5.A.9) imply that $B(2R,\lambda)$ is the Hankel transform of a function $b(2R,\vec{r})$ that is zero outside the disk of radius $2R$ centered at the origin of the plane. Now let $\vec{\lambda} = (\lambda_1,\lambda_2)$ in a Cartesian representation of the wave-number plane and consider the function $1 - B(2R,\sqrt{\lambda_1^2+\lambda_2^2})$ viewed as a function of λ_1 only (i.e. with λ_2 fixed). Then (5.A.11) implies that $1 - B(2R,\sqrt{\lambda_1^2+\lambda_2^2})$ is strictly positive for all values of λ_1, so that we can use the results of [22] to factor $1 - B(2R,\lambda)$ as

$$(1 - B(2R,\lambda)) = (1 - F(R,\lambda_1,\lambda_2))(1 - F^*(R,\lambda_1,\lambda_2)) \tag{5.A.12}$$

where $F(R,\lambda_1,\lambda_2)$ is the 2-D Fourier transform of a real function $f(R,\vec{r})$ [1] that is causal in the Cartesian coordinate r_1, where $\vec{r} = (r_1,r_2)$, i.e. where

$$f(R,\vec{r}) = 0, \quad \text{for } r_1 < 0. \tag{5.A.13}$$

Substituting (5.A.12) into (5.A.9) we obtain

$$\hat{S}_y(\lambda) = \frac{P}{(1 - F(R,\lambda_1,\lambda_2))(1 - F^*(R,\lambda_1,\lambda_2))}. \tag{5.A.14}$$

Equation (5.A.14) is the continuous space version of the well known result in the discrete space case [19], [44] that any power spectral density function $S(e^{ju},e^{jv})$ which is strictly positive for all $(u,v) \in [-\pi,\pi]^2$ can be written in factored form as

$$S(z_1,z_2) = \frac{\sigma^2}{A(z_1,z_2)A^*(z_1,z_2)} \tag{5.A.15}$$

[1] The reason for the notation $f(R,\vec{r})$ as opposed to $f(2R,\vec{r})$ will become clear in the sequel.

where the filter $A(z_1, z_2)$ has a nonsymmetric half plane support. In fact (5.A.14) could have been derived by using (5.A.4), (5.A.15) and the transformations

$$z_1 = \frac{1+j\lambda_1}{1-j\lambda_1} \tag{5.A.16}$$

$$z_2 = \frac{1+j\lambda_2}{1-j\lambda_2} \tag{5.A.17}$$

where λ_1 and λ_2 are allowed to take complex values. The transformations (5.A.16)-(5.A.17) are analogous to those which are used in the 1-D context to map the continuous time case into the discrete time case and vice-versa. However, unlike in the discrete space case [19], [44], where $A(z_1, z_2)$ often corresponds to a filter with an unbounded spatial support, the filter $f(R, \vec{r})$ has a finite support in the spatial domain. According to Theorem 3.4.2 in [68] and which is originally due to Plancherel and Polya [60], $B(2R, |\vec{\mu}|)$ where $\vec{\mu} \in \mathbf{C}^2$, must be an entire function of exponential type since it is equal to the Fourier transform of a function that is zero outside the disk of radius $2R$ in $\mathbf{R}^2$. Furthermore, (5.A.12) implies that

$$B(2R, \vec{\mu}) = F(R, \vec{\mu}) + F^*(R, \vec{\mu}) - F(R, \vec{\mu})F^*(R, \vec{\mu}), \tag{5.A.18}$$

so that $F(R, \vec{\mu})$ must also be an entire function of exponential type and must therefore be the Fourier transform of a function that is zero outside a bounded domain by the above mentioned theorem. Let us now study the spatial support $\mathcal{D}$ of the filter $f(R, \vec{r})$. Equation (5.A.12) implies that

$$b(2R, \vec{r}) = f(R, \vec{r}) + f(R, -\vec{r}) - \int_{\mathcal{D}} f(R, \vec{r'}) f(R, \vec{r} + \vec{r'}) d\vec{r'}. \tag{5.A.19}$$

Since $b(2R, \vec{r})$ is zero for $r > 2R$, and since $f(R, \vec{r})$ and $f(R, -\vec{r})$ appear on the right hand side of (5.A.19), then $\mathcal{D}$ must lie inside the half disk $\{\vec{r} : \quad r < 2R \quad \text{and} \quad \frac{-\pi}{2} < \theta < \frac{\pi}{2}\}$. Equation (5.A.19) implies also that the convolution

$$\int_{\mathcal{D}} f(R, \vec{r'}) f(R, \vec{r} + \vec{r'}) d\vec{r'} \tag{5.A.20}$$

has to be zero outside the disk $\{\vec{r} : r < 2R\}$. Hence, the product $f(R, \vec{r'}) f(R, \vec{r} + \vec{r'})$ must vanish identically for all $\vec{r} : r > 2R$, except maybe on a set of measure zero and on which it must remain finite. From the above discussion, we conclude that $\mathcal{D}$ must satisfy the following two constraints:

(i) $\mathcal{D} \subset \{\vec{r}: \quad r < 2R \quad \text{and} \quad \frac{-\pi}{2} < \theta < \frac{\pi}{2}\}$

(ii) $\mathcal{D} \bigcap \{\vec{s}: \quad \vec{s} + \vec{r} \in \mathcal{D} \quad \text{and} \quad r > 2R\} = \phi.$

A simple geometrical picture shows that the only subset of $\mathbf{R}^2$ that satisfies the above two constraints is a disk of radius R centered at the point $\vec{R_0} = (R, 0)$, i.e.

$$\mathcal{D} = \{\vec{r}: \quad |\vec{r} - \vec{R_0}| < R\}. \tag{5.A.21}$$

Alternatively identity (5.A.21) can be derived by noting that if $b(2R, \vec{r}) \in L_2([0, R] \times [0, 2\pi], d\vec{r})$ and since $b(2R, \vec{r})$ is zero outside of the disk $\{\vec{r}: \quad r \leq R\}$, then [68]

$$|B(2R, \vec{\mu})| \leq e^{2R|\mu_1|} e^{2R|\mu_2|} \tag{5.A.22}$$

where $\vec{\mu} = (\mu_1, \mu_2)$ with $\mu_1, \mu_2 \in \mathbf{C}$. Furthermore, (5.A.18) implies that

$$|F(R, \vec{\mu})| \leq e^{2R|\mu_1|} e^{2R|\mu_2|}. \tag{5.A.23}$$

Hence, the spatial support of $f(R, \vec{r})$ must lie inside the square $[0, 2R] \times [-2R, 2R]$ in a Cartesian representation of the plane [68]. Consider now the function

$$e(R, \vec{r}) = \delta(\vec{r} + \vec{R_0}) - f(R, \vec{r} + \vec{R_0}) \tag{5.A.24}$$

which has a finite spatial support centered at the origin that is a shifted version of that of $f(R, \vec{r})$. Note that (5.A.24) implies that

$$E(R, \vec{\lambda}) = e^{j\vec{\lambda} \cdot \vec{R_0}} (1 - F(R, \vec{\lambda})) \tag{5.A.25}$$

and it follows from (5.A.12) that

$$A(R, \lambda) = |E(R, \vec{\lambda})|^2. \tag{5.A.26}$$

Let us now expand $e(R, \vec{r})$ into a Fourier series of the form

$$e(R, \vec{r}) = \sum_{n=-\infty}^{\infty} e_n(R, r) e^{jn\theta}. \tag{5.A.27}$$

Correspondingly, the 2-D Fourier transform of $e(R,\vec{r})$, $E(R,\vec{\lambda})$, must have a Fourier series expansion of the form [55]

$$E(R,\vec{\lambda}) = \sum_{n=-\infty}^{\infty} E_n(R,\lambda)e^{-jn\phi} \tag{5.A.28}$$

where $E_n(R,\lambda)$ is the nth order Hankel transform of $e_n(R,r)$, and where $\vec{\lambda} = (\lambda, \frac{\pi}{2} - \phi)$. Equations (5.A.26) and (5.A.28) imply that (see the discussion preceding (5.2.12))

$$A(R,\lambda) = \sum_{n=-\infty}^{\infty} |E_n(R,\lambda)|^2. \tag{5.A.29}$$

Now, recall that $A(2R,\lambda)$ can be interpreted as being the zeroth-order Hankel transform of the function $a(2R,r)$ which is zero outside the disk $r \leq 2R$. (See (5.A.5).) Hence, according to the results of [23] we must have

$$|A(R,\mu)| \leq e^{2R|\mu|} \tag{5.A.30}$$

where $\mu \in \mathbf{C}$, so that

$$|E_n(R,\mu)|^2 \leq e^{2R|\mu|} \tag{5.A.31}$$

for all n. We now adapt a proof in [71] to show that

$$|E_n(R,\mu)| \leq e^{R|\mu|} \tag{5.A.32}$$

for all n, so that [23]

$$e_n(R,r) = 0 \qquad \text{for } r > R \tag{5.A.33}$$

for all n. Equation (5.A.33) proves that the spatial support of $e(R,\vec{r})$ is a disk of radius R centered at the origin and since $e(R,\vec{r}) = \delta(\vec{r} + \vec{R}_0) - f(R,\vec{r} + \vec{R}_0)$ then (5.A.21) must be true. To prove (5.A.32), we denote by μ_l the zeros of $|E_n(R,\mu)|^2$ counted according to multiplicity and so arranged that $|\mu_l|$ is non-decreasing with l. Since (5.A.31) is true, Jensen's formula [79] can be used to show that $|\mu_l| > cl$, for some $c > 0$. Furthermore, since $|E_n(R,\mu)|^2$ is even, the Hadamard factorization theorem [79] gives the representation

$$|E_n(R,\mu)|^2 = \prod_{l=1}^{\infty}(1 - \mu_l^{-2}\mu^2). \tag{5.A.34}$$

Now, every subproduct of (5.A.34) is dominated by $\prod(1+|\mu/cl|^2)$. Comparison with the well-known product representation of $\mu^{-1}\sin\mu$ therefore shows that $E_n(R,\mu)$ is a function of exponential type. Furthermore, since (5.A.31) is true then $e^{j\mu R}E_n(R,\mu)$ is a function of exponential type which is bounded on the real axis and on the upper half of the imaginary axis since $A(R,\mu)$ is bounded on the real and upper half of the imaginary axis by (5.A.30). The Phragmen-Lindelof theorem [79] now implies that $e^{j\mu R}E_n(R,\mu)$ is bounded in the first and second quadrants. A similar argument applies in the lower half plane. Hence (5.A.32) must be true.

Next denote by $\mathcal{C}$ the causal space of functions of $\vec{\lambda}$ which are the Fourier transforms of functions that are zero for $r_1 < 0$ in a Cartesian coordinate representation of the spatial domain (i.e. where $\vec{r} = (r_1, r_2)$), and denote by $\mathcal{A}$ the anticausal space of functions of $\vec{\lambda}$ which are the Fourier transforms of functions that are zero for $r_1 > 0$. Since $F^*(R,\vec{\lambda})$ is the Fourier transform of a function that is zero for $r_1 > 0$, then

$$D(\vec{\lambda}) = (1 - F^*(R,\vec{\lambda}))^{-1} - 1 \tag{5.A.35}$$

must also correspond to the Fourier transform of a function that is zero for $r_1 > 0$. To see why this has to be true, factor $D(\vec{\lambda})$ as [14]

$$D(\vec{\lambda}) = D_c(\vec{\lambda}) + D_a(\vec{\lambda}), \tag{5.A.36}$$

where $D_c(\vec{\lambda})$ and $D_a(\vec{\lambda})$ belong respectively to $\mathcal{C}$ and $\mathcal{A}$. Equations (5.A.35)-(5.A.36) imply that

$$F^*(R,\vec{\lambda}) = D_c(\vec{\lambda}) + D_a(\vec{\lambda}) - D_c(\vec{\lambda})F^*(R,\vec{\lambda}) - D_a(\vec{\lambda})F^*(R,\vec{\lambda}) \tag{5.A.37}$$

and since $F^*(R,\vec{\lambda})$ is the Fourier transform of a function that is zero for $r_1 > 0$, we must have

$$D_c(\vec{\lambda}) = D_c(\vec{\lambda})F^*(R,\vec{\lambda}) \tag{5.A.38}$$

or

$$D_c(\vec{\lambda}) = 0 \tag{5.A.39}$$

which proves our assertion. Combining (5.A.35) with (5.A.14) we obtain

$$\int_{-\infty}^{\infty} \hat{S}_y(\vec{\lambda})(1 - F(R,\vec{\lambda}))e^{j\lambda_1 r_1}\,d\lambda_1 = P\delta(r_1) \quad \text{for} \quad r_1 > 0. \tag{5.A.40}$$

Furthermore, if we take the inverse Fourier transform of (5.A.40) with respect to λ_2 we get

$$\int_{\mathcal{D}} \hat{K}_y(|\vec{r}-\vec{s}|)(\delta(\vec{s}) - f(R,\vec{s}))\, d\vec{s} = P\delta(\vec{r}) \quad \text{for} \quad r_1 > 0. \tag{5.A.41}$$

To compute $f(R,\vec{r})$ from the above integral equation, we note that (5.A.21) implies that for any $\vec{r} \in \mathcal{D}$ and any $\vec{s} \in \mathcal{D}$ we have $|\vec{r}-\vec{s}| < 2R$, so that

$$\hat{K}_y(|\vec{r}-\vec{s}|) = K_y(|\vec{r}-\vec{s}|) \quad \forall \vec{r} \in \mathcal{D} \quad \text{and} \quad \forall \vec{s} \in \mathcal{D} \tag{5.A.42}$$

by the correlation matching constraint (5.A.3). Since $K_y(\vec{r}) = P\delta(\vec{r}) + K_z(|\vec{r}|)$ is known by assumption for $r < 2R$ then $f(R,\vec{r})$ can be computed as the solution of the following integral equation

$$K_z(r) = \int_{\mathcal{D}} K_z(|\vec{r}-\vec{s}|) f(R,\vec{s})\, d\vec{s} + Pf(R,\vec{s}) \quad \forall \vec{r} \in \mathcal{D}. \tag{5.A.43}$$

Once $f(R,\vec{r})$ has been computed via (5.A.43), then (5.A.41) can be used to extend $K_z(r)$ beyond the disk $r < 2R$. Finally, if we make the change of variables

$$\vec{r'} = \vec{R}_0 - \vec{r} \tag{5.A.44}$$

$$\vec{s'} = \vec{R}_0 - \vec{s} \tag{5.A.45}$$

we obtain from (5.A.43)

$$K_z(|\vec{R}_0 - \vec{r'}|) = \int_{s'<R} K_z(|\vec{r'}-\vec{s'}|) g(R,\vec{s'})\, d\vec{s'} + Pg(R,\vec{r'}) \quad \forall r' < R, \tag{5.A.46}$$

where we have defined

$$g(R,\vec{s'}) = f(R,\vec{R}_0 - \vec{s'}). \tag{5.A.47}$$

Note that (5.A.46) is just (5.1.6) with $\vec{r}$ and $\vec{u}$ replaced by $\vec{r'}$ and $\vec{s'}$ respectively. Furthermore, observe that (5.A.47) implies that

$$G(R,\vec{\lambda}) = F^*(R,\vec{\lambda}) e^{-j\vec{\lambda}\cdot\vec{R}_0} \tag{5.A.48}$$

where $G(R,\vec{\lambda})$ is the 2-D Fourier transform of $g(R,\vec{r})$. Hence,

$$\begin{aligned} |1 - F(R,\vec{\lambda})|^2 &= |1 - F^*(R,\vec{\lambda})|^2 \\ &= |1 - G(R,\vec{\lambda}) e^{j\vec{\lambda}\cdot\vec{R}_0}|^2 \\ &= |e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|^2. \end{aligned} \tag{5.A.49}$$

Combining (5.A.14) with (5.A.49) we obtain

$$\hat{S}_y(\lambda) = \frac{P}{|e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|^2} \tag{5.A.50}$$

which is equation (5.1.5).

Chapter 6

INTRODUCTION TO ISOTROPIC RANDOM VECTOR FIELDS

In this thesis we have considered only scalar zero-mean 2-D isotropic random fields. In practice, isotropic random *vector* fields are also of interest. Examples of such fields are the velocity field in turbulence [50] and the electromagnetic field of a black body radiator [49]. Furthermore, as we shall see, the definition of an isotropic random vector field is less restrictive than that of an isotropic random scalar field and their structure leads to a more general analysis of both spectral estimation problems and recursive estimation problems in higher dimensions. The study of such fields is therefore important in the understanding of the special issues that arise in multidimensional signal processing problems. So far estimation problems for such fields have not been considered in the literature. The objective of this chapter is to introduce the reader to the mathematical theory of isotropic random vector fields and to discuss some estimation problems for such vector fields.

6.1 CHARACTERIZATION OF ISOTROPIC RANDOM VECTOR FIELDS

We begin this section by giving a formal definition of an isotropic random vector field.

Definition 6.1 *An isotropic random vector field is characterized by a mean vector and a covariance matrix which are unaffected by any translation, rotation and/or reflection accompanied by a simultaneous rotation and/or reflection of the coordinate system with respect to which the components of the random vector are determined.*

This definition may look strange at first but becomes more intuitive once we recognize that vector fields are meant to be integrated along curves whereas scalar fields are meant to be integrated over areas. For example, consider the random variable a obtained by integrating the isotropic random scalar field $z(\vec{r})$ over some area A

$$a = \int_A z(\vec{r}) d\vec{r} \tag{6.1.1}$$

where $d\vec{r} = dx\, dy$ denotes an element of area. If we apply an arbitrary rigid motion T to the coordinate system we obtain the new random variable

$$a' = \int_{T(A)} z'(\vec{r'}) d\vec{r'} \tag{6.1.2}$$

where

$$z'(\vec{r'}) = z(T(\vec{r})). \tag{6.1.3}$$

Since $z(\vec{r})$ is isotropic, $z(\vec{r})$ and $z'(\vec{r'})$ have the same first and second order moments by definition, and hence a and a' have the same first and second order statistics. Now consider the random variable b obtained by integrating an isotropic random vector field $\vec{v}(\vec{r})$ along a curve C

$$b = \int_C \vec{v}(\vec{r}) \cdot d\vec{l} \tag{6.1.4}$$

where $\vec{u} \cdot \vec{v}$ denotes the inner product of the vectors $\vec{u}$ and $\vec{v}$ and where $d\vec{l}$ denotes an element of length along C. If we apply an arbitrary rigid motion T to the coordinate system we obtain a new random variable

$$b' = \int_{T(C)} \vec{v'}(\vec{r'}) \cdot d\vec{l'} \tag{6.1.5}$$

where

$$\vec{v'}(\vec{r'}) = T(\vec{v}(T(\vec{r}))). \tag{6.1.6}$$

The random variables b and b' will have the same first and second order statistics for any curve C if and only if $\vec{v}(\vec{r})$ and $\vec{v}'(\vec{r}')$ have the same first and second order moments, i.e. if and only if $\vec{v}(\vec{r})$ is an isotropic random vector field according to the above definition.

Another advantage of the above definition is that the isotropy property is preserved under all elementary vector differential operators such as the gradient, the curl or the divergence operators and is also preserved under scaling, addition and rotation by a multiple of $\pi/2$ as we shall see.

By using Definition 6.1 we obtain the following characterization of isotropic random vector fields [93], [27], [54].

Theorem 6.1 *An isotropic random vector field $\vec{v}(\vec{r})$ is characterized by a mean vector which is identically zero,*

$$E[\vec{v}(\vec{r})] = 0 \tag{6.1.7}$$

and a correlation matrix $\mathbf{K}(\vec{r})$ which has the structure

$$\begin{aligned} \mathbf{K}(\vec{r}) &= E[\vec{v}(\vec{r}+\vec{s})\vec{v}(\vec{s})] \\ &= [K_{ij}(\vec{r})] \\ &= [(K_l(r) - K_t(r))\frac{r_i r_j}{r^2} + K_t(r)\delta_{ij}] \end{aligned} \tag{6.1.8}$$

where $K_{ij}(\vec{r})$ is the ijth component of the correlation matrix $\mathbf{K}(\vec{r})$, and where $K_l(r)$ and $K_t(r)$ are two scalar isotropic covariance functions which are called respectively the longitudinal and transverse covariance functions. Furthermore, the power spectral density matrix of $\vec{v}(\vec{r})$, $\mathbf{S}(\vec{\lambda})$, can be represented in the form

$$\begin{aligned} \mathbf{S}(\vec{\lambda}) &= [S_{ij}(\vec{\lambda})] \\ &= [(S_p(\lambda) - S_s(\lambda))\frac{\lambda_i \lambda_j}{\lambda^2} + S_s(\lambda)\delta_{ij}] \end{aligned} \tag{6.1.9}$$

where $S_p(\lambda)$ and $S_s(\lambda)$ are two scalar isotropic power spectra that are called respectively the potential and solenoidal spectrum.

The covariance functions $K_l(r)$ and $K_t(r)$ and the power spectra $S_p(\lambda)$ and $S_s(\lambda)$ are related via the equations

$$\begin{bmatrix} K_l(r) \\ K_t(r) \end{bmatrix} = 2\pi \int_0^\infty \overline{\mathbf{J}}_1(\lambda r) \begin{bmatrix} S_p(\lambda) \\ S_s(\lambda) \end{bmatrix} \lambda \, d\lambda, \tag{6.1.10}$$

$$\begin{bmatrix} S_p(\lambda) \\ S_s(\lambda) \end{bmatrix} = \frac{1}{2\pi} \int_0^\infty \overline{\mathbf{J}}_1(\lambda r) \begin{bmatrix} K_l(r) \\ K_t(r) \end{bmatrix} r \, dr, \tag{6.1.11}$$

where the matrix $\overline{\mathbf{J}}_1(x)$ is defined by

$$\overline{\mathbf{J}}_1(x) = \begin{bmatrix} \frac{d}{dx}J_1(x) & \frac{1}{x}J_1(x) \\ \frac{1}{x}J_1(x) & \frac{d}{dx}J_1(x) \end{bmatrix}. \tag{6.1.12}$$

Equation (6.1.8) was originally derived by Robertson [64] using the theory of invariants. Theorem 6.1 as a whole was independently derived by Yaglom [93], Ito [27] and Ogura [54]. The proof of this theorem can be found in any of the above cited references.

The covariance functions $K_l(r)$ and $K_t(r)$ have an interesting interpretation. Consider Fig. 6.1 and let L be the line defined by the pair $(\vec{n}, p)$ where $\vec{n}$ is the unit vector normal to L and where p is the distance from the origin to the line L. Furthermore, define $\vec{t}$ to be a unit vector along L. Now let $\vec{r} = p\vec{n} + x\vec{t}$ be a particular point on the line L characterized by the scalar p which is defined above, and the scalar x which is the length of the projection of $\vec{r}$ along L (see Fig. 6.1), and decompose the isotropic random vector field $\vec{v}(\vec{r})$ as

$$\vec{v}(\vec{r}) = v_\perp(x)\vec{n} + v_{||}(x)\vec{t} \tag{6.1.13}$$

where

$$v_\perp(x) = v(p\vec{n} + x\vec{t})\vec{n} = \vec{v}(\vec{r}) \cdot \vec{n} \tag{6.1.14}$$

and

$$v_{||}(x) = v(p\vec{n} + x\vec{t})\vec{t} = \vec{v}(\vec{r}) \cdot \vec{t}. \tag{6.1.15}$$

Then we have the following result.

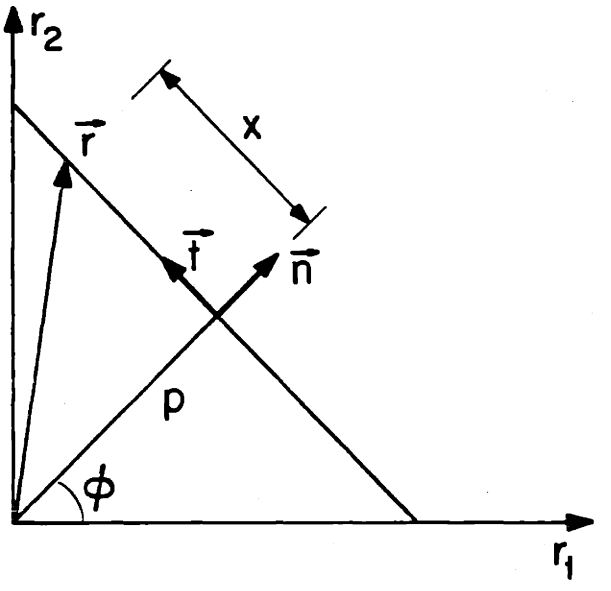

Figure 6.1: Coordinate system associated to the line L.

Theorem 6.2 *The processes $v_\perp(x)$ and $v_{||}(x)$ are zero-mean uncorrelated, stationary processes with covariances.*

$$E[v_\perp(x)v_\perp(y)] = K_t(|x-y|) \tag{6.1.16}$$

$$E[v_{||}(x)v_{||}(y)] = K_l(|x-y|). \tag{6.1.17}$$

Proof

If $\vec{r}$ is on the line L then from the geometry of Fig. 6.1 we have

$$v_\perp(x) = v_1(\vec{r})cos\phi + v_2(\vec{r})sin\phi \tag{6.1.18}$$

and

$$v_{||}(x) = -v_1(\vec{r})sin\phi + v_2(\vec{r})cos\phi, \tag{6.1.19}$$

where $v_1(\vec{r})$ and $v_2(\vec{r})$ are the components of the vector $\vec{v}(\vec{r})$ along the r_1 and r_2 axes respectively, in a Cartesian representation of the underlying plane. Hence,

$$E[v_\perp(x)v_\perp(y)] = K_{11}(\vec{s})\cos^2\phi + 2K_{12}(\vec{s})\sin\phi\cos\phi + K_{22}(\vec{s})\sin^2\phi \tag{6.1.20}$$

$$E[v_\perp(x)v_{||}(y)] = (K_{22}(\vec{s}) - K_{11}(\vec{s}))\sin\phi\cos\phi + K_{12}(\vec{s})(\cos^2\phi - \sin^2\phi) \tag{6.1.21}$$

$$E[v_{||}(x)v_{||}(y)] = K_{11}(\vec{s})\sin^2\phi - 2K_{12}(\vec{s})\sin\phi\cos\phi + K_{22}(\vec{s})\cos^2\phi \tag{6.1.22}$$

where $\vec{s} = (x-y)\vec{t}$. Combining (6.1.8) and (6.1.20)-(6.1.22), we obtain

$$E[v_\perp(x)v_\perp(y)] = K_t(|x-y|) \tag{6.1.23}$$

$$E[v_\perp(x)v_{||}(y)] = 0 \tag{6.1.24}$$

$$E[v_{||}(x)v_{||}(y)] = K_l(|x-y|). \tag{6.1.25}$$

□□

An isotropic random vector field can also be decomposed into two uncorrelated vector components namely its solenoidal part and its potential part. A vector field

$\vec{v}(\vec{r})$ is a solenoidal field if

$$\begin{aligned} \text{div } \vec{v}(\vec{r}) &= \nabla \cdot \vec{v}(\vec{r}) \\ &= \frac{\partial}{\partial r_1} v_1(\vec{r}) + \frac{\partial}{\partial r_2} v_2(\vec{r}) \\ &= 0 \end{aligned} \tag{6.1.26}$$

and it is a potential field if

$$\begin{aligned} \text{curl } v(\vec{r}) &= \nabla \times \vec{v}(\vec{r}) \\ &= \frac{\partial}{\partial r_1} v_2(\vec{r}) - \frac{\partial}{\partial r_2} v_1(\vec{r}) \\ &= 0. \end{aligned} \tag{6.1.27}$$

Note that according to (6.1.27) the curl of a vector field in two dimensions is a scalar field. It is well known that any vector field $\vec{v}(\vec{r})$ can be divided into a solenoidal part $\vec{v}_s(\vec{r})$ and a potential part $\vec{v}_p(\vec{r})$ [34]. We shall now show that the solenoidal and potential parts of an isotropic random vector field are uncorrelated and that their power spectral density matrices $\mathbf{S}^p(\lambda)$ and $\mathbf{S}^s(\lambda)$ are characterized respectively by $S_s(\lambda)$ and $S_p(\lambda)$. Let us first prove the following useful lemma.

Lemma 6.1 *The solenoidal and potential parts $\vec{v}_s(\vec{r})$ and $\vec{v}_p(\vec{r})$ respectively, of a 2-D vector field $\vec{v}(\vec{r})$ can be computed from $\vec{v}(\cdot)$ as*

$$\begin{aligned} \vec{v}_s(\vec{r}) &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \nabla \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \nabla \cdot \vec{w}(\vec{r}) \\ &= \begin{bmatrix} \frac{\partial^2}{\partial r_2^2} & -\frac{\partial^2}{\partial r_1 \partial r_2} \\ -\frac{\partial^2}{\partial r_1 \partial r_2} & \frac{\partial^2}{\partial r_1^2} \end{bmatrix} \vec{w}(\vec{r}) \end{aligned} \tag{6.1.28}$$

$$\begin{aligned} \vec{v}_p(\vec{r}) &= \nabla \nabla \cdot \vec{w}(\vec{r}) \\ &= \begin{bmatrix} \frac{\partial^2}{\partial r_1^2} & \frac{\partial^2}{\partial r_1 \partial r_2} \\ \frac{\partial^2}{\partial r_1 \partial r_2} & \frac{\partial^2}{\partial r_2^2} \end{bmatrix} \vec{w}(\vec{r}) \end{aligned} \tag{6.1.29}$$

where

$$\vec{w}(\vec{r}) = \frac{1}{2\pi} \int_{\mathbf{R}^2} \ln(|\vec{r} - \vec{r}'|) \vec{v}(\vec{r}')\, d\vec{r}'. \tag{6.1.30}$$

Proof

By direct substitution we can check that

$$\nabla \times \vec{v}_p(\vec{r}) = 0 \tag{6.1.31}$$

and

$$\nabla \cdot \vec{v}_s(\vec{r}) = 0. \tag{6.1.32}$$

Next, we note that

$$\begin{aligned} \vec{v}_p(\vec{r}) + \vec{v}_s(\vec{r}) &= \nabla^2 \vec{w}(\vec{r}) \\ &= \vec{v}(\vec{r}) \end{aligned} \tag{6.1.33}$$

and where we have used the fact that $-\frac{1}{2\pi}\ln(|\vec{r} - \vec{r'}|)$ is the Green's function associated to the 2-D Laplacian operator ∇^2 defined over the whole plane, i.e.

$$\frac{1}{2\pi}\nabla^2 \ln(|\vec{r} - \vec{r'}|) = \delta(\vec{r} - \vec{r'}). \tag{6.1.34}$$

□□

The solenoidal and potential parts of the isotropic random vector field are characterized in the next theorem.

Theorem 6.3 *The solenoidal and potential parts $\vec{v}_s(\vec{r})$ and $\vec{v}_p(\vec{r})$ of an isotropic random field $\vec{v}(\vec{r})$ are uncorrelated. Their respective power density matrices $\mathbf{S}^s(\vec{\lambda})$ and $\mathbf{S}^p(\vec{\lambda})$ can be written as*

$$\begin{aligned} \mathbf{S}^s(\vec{\lambda}) &= [(\delta_{i,j} - \frac{\lambda_i \lambda_j}{\lambda^2})S_s(\lambda)] \\ &= \begin{bmatrix} \frac{\lambda_2^2}{\lambda^2}S_s(\lambda) & -\frac{\lambda_1\lambda_2}{\lambda^2}S_s(\lambda) \\ -\frac{\lambda_1\lambda_2}{\lambda^2}S_s(\lambda) & \frac{\lambda_1^2}{\lambda^2}S_s(\lambda) \end{bmatrix} \end{aligned} \tag{6.1.35}$$

$$\begin{aligned} \mathbf{S}^p(\vec{\lambda}) &= [\frac{\lambda_i \lambda_j}{\lambda^2}S_p(\lambda)] \\ &= \begin{bmatrix} \frac{\lambda_1^2}{\lambda^2}S_p(\lambda) & \frac{\lambda_1\lambda_2}{\lambda^2}S_p(\lambda) \\ \frac{\lambda_1\lambda_2}{\lambda^2}S_p(\lambda) & \frac{\lambda_2^2}{\lambda^2}S_p(\lambda) \end{bmatrix} \end{aligned} \tag{6.1.36}$$

where $S_s(\lambda)$ and $S_p(\lambda)$ are the scalar solenoidal and potential spectral densities associated with $\vec{v}(\vec{r})$ and which are defined in Theorem 6.1.

Proof

Observe that $\vec{v}_s(\vec{r})$ and $\vec{v}_p(\vec{r})$ are obtained by passing $\vec{v}(\vec{r})$ through the linear space-invariant filters $\mathbf{H}_s(\vec{\lambda})$ and $\mathbf{H}_p(\vec{\lambda})$ respectively, where

$$\begin{aligned}\mathbf{H}^s(\vec{\lambda}) &= [\delta_{i,j} - \frac{\lambda_i\lambda_j}{\lambda^2}] \\ &= \begin{bmatrix} \frac{\lambda_2^2}{\lambda^2} & -\frac{\lambda_1\lambda_2}{\lambda^2} \\ -\frac{\lambda_1\lambda_2}{\lambda^2} & \frac{\lambda_1^2}{\lambda^2} \end{bmatrix} \qquad (6.1.37) \\ \mathbf{H}^p(\vec{\lambda}) &= [\frac{\lambda_i\lambda_j}{\lambda^2}] \\ &= \begin{bmatrix} \frac{\lambda_1^2}{\lambda^2} & \frac{\lambda_1\lambda_2}{\lambda^2} \\ \frac{\lambda_1\lambda_2}{\lambda^2} & \frac{\lambda_2^2}{\lambda^2} \end{bmatrix}. \qquad (6.1.38)\end{aligned}$$

Hence, by direct substitution we have

$$\begin{aligned}\mathbf{S}^s(\vec{\lambda}) &= \mathbf{H}_s(\vec{\lambda})\mathbf{S}(\vec{\lambda})\mathbf{H}_s^T(\vec{\lambda}) \\ &= \begin{bmatrix} \frac{\lambda_2^2}{\lambda^2}S_s(\lambda) & -\frac{\lambda_1\lambda_2}{\lambda^2}S_s(\lambda) \\ -\frac{\lambda_1\lambda_2}{\lambda^2}S_s(\lambda) & \frac{\lambda_1^2}{\lambda^2}S_s(\lambda) \end{bmatrix} \qquad (6.1.39) \\ \mathbf{S}^p(\vec{\lambda}) &= \mathbf{H}_p(\vec{\lambda})\mathbf{S}(\vec{\lambda})\mathbf{H}_p^T(\vec{\lambda}) \\ &= \begin{bmatrix} \frac{\lambda_1^2}{\lambda^2}S_p(\lambda) & \frac{\lambda_1\lambda_2}{\lambda^2}S_p(\lambda) \\ \frac{\lambda_1\lambda_2}{\lambda^2}S_p(\lambda) & \frac{\lambda_2^2}{\lambda^2}S_p(\lambda) \end{bmatrix}. \qquad (6.1.40)\end{aligned}$$

Similarly, the cross-power spectral density matrix $\mathbf{S}^{sp}(\vec{\lambda})$ of $z_s(\vec{r})$ and $z_p(\vec{r})$ is given by

$$\begin{aligned}\mathbf{S}^{sp}(\vec{\lambda}) &= \mathbf{H}_s(\vec{\lambda})S(\vec{\lambda})\mathbf{H}_p^T(\vec{\lambda}) \\ &= 0. \qquad (6.1.41)\end{aligned}$$

□□

Equations (6.1.35) - (6.1.36) explain the terminology solenoidal and potential power spectra introduced earlier; the power spectral densities $S_s(\lambda)$ and $S_p(\lambda)$ respectively characterize the power spectral density matrices of the solenoidal and potential parts of the isotropic random vector field $\vec{v}(\vec{r})$.

We conclude this section by showing that the isotropy property is preserved under all elementary vector differential operators such as the gradient, the divergence and the curl operators. Consider a scalar isotropic random field $z(\vec{r})$ and let $\vec{v}(\vec{r})$ be the gradient of $z(\vec{r})$, i.e. let

$$\vec{v}(\vec{r}) = \nabla z(\vec{r}). \tag{6.1.42}$$

Equation (6.1.42) implies that $\vec{v}(\cdot)$ is an isotropic random potential vector field with the power spectral density matrix

$$\mathbf{S}_v(\vec{\lambda}) = \begin{bmatrix} \lambda_1^2 S_z(\lambda) & \lambda_1\lambda_2 S_z(\lambda) \\ \lambda_1\lambda_2 S_z(\lambda) & \lambda_2^2 S_z(\lambda) \end{bmatrix}, \tag{6.1.43}$$

where $S_z(\lambda)$ is the power spectrum of $z(\cdot)$. Comparing (6.1.43) and (6.1.36), we find that

$$S_p(\lambda) = \lambda^2 S_z(\lambda). \tag{6.1.44}$$

Similarly, the vector field

$$\begin{aligned} \vec{w}(\vec{r}) &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \nabla z(\vec{r}) \\ &= \begin{bmatrix} -\frac{\partial}{\partial r_2} z(\vec{r}) \\ \frac{\partial}{\partial r_1} z(\vec{r}) \end{bmatrix} \end{aligned} \tag{6.1.45}$$

is an isotropic random solenoidal vector field with the power spectral density matrix

$$\mathbf{S}_w(\vec{\lambda}) = \begin{bmatrix} \lambda_2^2 S_z(\lambda) & -\lambda_1\lambda_2 S_z(\lambda) \\ -\lambda_1\lambda_2 S_z(\lambda) & \lambda_1^2 S_z(\lambda) \end{bmatrix}. \tag{6.1.46}$$

Again, comparing (6.1.46) and (6.1.35) we find that

$$S_s(\lambda) = \lambda^2 S_z(\lambda). \tag{6.1.47}$$

Observe that (6.1.42) and (6.1.45) imply that we can express any isotropic potential vector field as the gradient of a properly chosen isotropic scalar field and any isotropic solenoidal vector field as the rotated gradient of a properly chosen isotropic

scalar field. Specifically, given the isotropic potential vector field $\vec{v}(\vec{r})$ characterized by the potential power spectrum $S_p(\lambda)$ we can write

$$\vec{v}(\vec{r}) = \nabla z_p(\vec{r}) \tag{6.1.48}$$

where $z_p(\vec{r})$ is an isotropic random scalar field with the power spectral density $S_p(\lambda)/\lambda^2$. Similarly, given the isotropic solenoidal vector field $\vec{w}(\vec{r})$ characterized by the solenoidal power spectrum $S_s(\lambda)$ we can write

$$\vec{w}(\vec{r}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \nabla z_s(\vec{r}) \tag{6.1.49}$$

where $z_s(\vec{r})$ is an isotropic random scalar field with the power spectral density $S_s(\lambda)/\lambda^2$. Hence, any isotropic random vector field $\vec{v}(\vec{r})$ can be written as

$$\vec{v}(\vec{r}) = \nabla z_p(\vec{r}) + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \nabla z_s(\vec{r}) \tag{6.1.50}$$

where $z_p(\vec{r})$ and $z_s(\vec{r})$ are two uncorrelated isotropic random scalar fields with the power spectral densities $S_p(\lambda)/\lambda^2$ and $S_s(\lambda)/\lambda^2$ respectively. Observe that since the isotropy property of random scalar fields is preserved under scaling and addition then equation (6.1.50) implies that the isotropy property of the random vector $\vec{v}(\vec{r})$ is preserved under scaling, addition and rotation by a multiple of $\pi/2$. Rotation by $\pi/2$ is equivalent to multiplying the vector $v(\vec{r})$ by the matrix R, where R is given by

$$R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \tag{6.1.51}$$

and has for effect to interchange the potential and solenoidal parts of the vector field. Finally, note that the divergence and the non-zero component of the curl of a mean-square differentiable 2-D isotropic random vector field $\vec{v}(\vec{r})$ are isotropic random scalar fields. Specifically, if

$$\begin{aligned} z(\vec{r}) &= \nabla \cdot \vec{v}(\vec{r}) \\ &= \frac{\partial}{\partial r_1} v_1(\vec{r}) + \frac{\partial}{\partial r_2} v_2(\vec{r}) \end{aligned} \tag{6.1.52}$$

and if

$$\begin{aligned} y(\vec{r}) &= \nabla \times \vec{v}(\vec{r}) \\ &= \frac{\partial}{\partial r_1} v_2(\vec{r}) - \frac{\partial}{\partial r_2} v_1(\vec{r}), \end{aligned} \tag{6.1.53}$$

then by direct substitution we find that $S_z(\lambda)$ and $S_y(\lambda)$ are given respectively by

$$S_z(\lambda) = \lambda^2 S_p(\lambda) \tag{6.1.54}$$

$$S_y(\lambda) = \lambda^2 S_s(\lambda). \tag{6.1.55}$$

6.2 SPECTRAL REPRESENTATIONS

In this section we use the theory presented in the last section to develop a spectral representation for isotropic random vector fields similar to the spectral representation (2.1.21) that is known for isotropic random scalar fields. This spectral representation also leads to a useful series expansion for isotropic random vector fields similar to the Fourier series expansions for isotropic random scalar fields. We begin by presenting the spectral representation for isotropic random vector fields.

Theorem 6.4 *Any isotropic random vector field $\vec{v}(\vec{r})$ has a spectral representation of the form*

$$\vec{v}(\vec{r}) = \sum_{n=-\infty}^{\infty} \mathbf{U}(\theta) e^{jn\theta} \int_0^{\infty} \mathbf{J}_n(\lambda r)\, \vec{V}_n(d\lambda), \tag{6.2.1}$$

where $\{\vec{V}_n(d\lambda)\}$ are zero-mean uncorrelated orthogonal vector random measures with

$$E[\vec{V}_n(d\lambda_1)\vec{V}_m^H(d\lambda_2)] = \begin{bmatrix} S_p(\lambda_1)/\lambda_1 & 0 \\ 0 & S_s(\lambda_1)/\lambda_1 \end{bmatrix} \delta_{n,m}\delta(\lambda_1 - \lambda_2)\, d\lambda_1\, d\lambda_2, \tag{6.2.2}$$

where $\vec{V}_m^H(\cdot)$ denotes the complex conjugate transpose of $\vec{V}_m(\cdot)$. In (6.2.1) the matrices $\mathbf{U}(\theta)$ and $\mathbf{J}_n(x)$ are given by

$$\mathbf{U}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{6.2.3}$$

$$\mathbf{J}_n(x) = \begin{bmatrix} \frac{d}{dx}J_n(x) & \frac{-jn}{x}J_n(x) \\ \frac{jn}{x}J_n(x) & \frac{d}{dx}J_n(x) \end{bmatrix}. \tag{6.2.4}$$

Proof

Recall that any isotropic random vector field $\vec{v}(\vec{r})$ can be written as

$$\vec{v}(\vec{r}) = \nabla z_p(\vec{r}) + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \nabla z_s(\vec{r}) \tag{6.2.5}$$

where $z_p(\vec{r})$ and $z_s(\vec{r})$ are two uncorrelated isotropic random scalar fields with the power spectral densities $S_p(\lambda)/\lambda^2$ and $S_s(\lambda)/\lambda^2$ respectively. Furthermore, recall that according to equation (2.1.21) $z_p(\vec{r})$ and $z_s(\vec{r})$ have the spectral representations

$$z_p(\vec{r}) = \sum_{n=-\infty}^{\infty} \int_0^{\infty} J_n(\lambda r) Z_n^p(d\lambda)\, e^{jn\theta} \tag{6.2.6}$$

and

$$z_s(\vec{r}) = \sum_{n=-\infty}^{\infty} \int_0^{\infty} J_n(\lambda r) Z_n^s(d\lambda)\, e^{jn\theta} \tag{6.2.7}$$

where $Z_n^p(\lambda)$ and $Z_n^s(\lambda)$ are uncorrelated[1] zero-mean orthogonal random measures with

$$E\left[\begin{bmatrix} Z_n^p(d\lambda_1) \\ Z_n^s(d\lambda_1) \end{bmatrix} \begin{bmatrix} Z_m^{pH}(d\lambda_2) & Z_m^{sH}(d\lambda_2) \end{bmatrix}\right] = \begin{bmatrix} S_p(\lambda_1)/\lambda_1 & 0 \\ 0 & S_s(\lambda_1)/\lambda_1 \end{bmatrix} \delta_{n,m}\delta(\lambda_1-\lambda_2)\, d\lambda_1\, d\lambda_2. \tag{6.2.8}$$

Note also that in a Cartesian coordinate system we have

$$\nabla J_n(\lambda r)\, e^{jn\theta} = \lambda \dot{J}_n(\lambda r) e^{jn\theta} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} + \frac{jn}{r} J_n(\lambda r) e^{jn\theta} \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}, \tag{6.2.9}$$

where $\dot{J}_n(x) = \frac{d}{dx} J_n(x)$. Hence, if we use the notation

$$\mathbf{U}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{6.2.10}$$

$$\mathbf{J}_n(x) = \begin{bmatrix} \frac{d}{dx} J_n(x) & \frac{-jn}{x} J_n(x) \\ \frac{jn}{x} J_n(x) & \frac{d}{dx} J_n(x) \end{bmatrix} \tag{6.2.11}$$

$$\vec{V}_n(d\lambda) = \begin{bmatrix} Z_n^p(d\lambda_1) \\ Z_n^s(d\lambda_1) \end{bmatrix}, \tag{6.2.12}$$

[1]The random measures $\{Z_n^p(\lambda)\}$ and $\{Z_n^s(\lambda)\}$ are uncorrelated as a consequence of the fact that $z_p(\cdot)$ and $z_s(\cdot)$ are uncorrelated.

we obtain the following spectral representation for isotropic random vector fields by combining (6.2.5)-(6.2.12),

$$\vec{v}(\vec{r}) = \sum_{n=-\infty}^{\infty} \mathbf{U}(\theta) e^{jn\theta} \int_0^{\infty} \mathbf{J}_n(\lambda r)\, \vec{V}_n(d\lambda). \tag{6.2.13}$$

□□

In the sequel, we shall call the matrix functions $\mathbf{J}_n(x)$ *matrix Bessel functions* of order n. These functions have a number of interesting properties. First, note that they are Hermitian. Furthermore, by using the identities [7]

$$\dot{J}_n(x) + \frac{n}{x} J_n(x) = J_{n-1}(x) \tag{6.2.14}$$

$$\dot{J}_n(x) - \frac{n}{x} J_n(x) = -J_{n+1}(x), \tag{6.2.15}$$

we find that the matrix $\mathbf{J}_n(x)$ has the representation

$$\mathbf{J}_n(x) = \mathbf{P}\ \mathrm{diag}(J_{n-1}(x)\ ,\ -J_{n+1}(x))\mathbf{P}^H, \tag{6.2.16}$$

where the matrix $\mathbf{P}$ is given by

$$\mathbf{P} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & j \\ j & 1 \end{bmatrix}. \tag{6.2.17}$$

Note also that this implies that

$$\begin{aligned} \int_0^{\infty} \mathbf{J}_n(\lambda_1 r)\mathbf{J}_n(\lambda_2 r)\, r\, dr &= \int_0^{\infty} \mathbf{P} \begin{bmatrix} J_{n-1}(\lambda_1 r)J_{n-1}(\lambda_2 r) & 0 \\ 0 & J_{n+1}(\lambda_1 r)J_{n+1}(\lambda_2 r) \end{bmatrix} \mathbf{P}^H\, r\, dr \\ &= \frac{\delta(\lambda_1 - \lambda_2)}{\lambda_1} I_2, \end{aligned} \tag{6.2.18}$$

where I_2 is the 2×2 identity matrix. Note that (6.2.18) is a generalization of identity [55]

$$\int_0^{\infty} J_n(\lambda_1 r) J_n(\lambda_2 r)\, r\, dr = \frac{\delta(\lambda_1 - \lambda_2)}{\lambda_2} \tag{6.2.19}$$

which holds for scalar Bessel functions. Finally, observe that the matrix Bessel function of order 1 is related to the matrix function $\overline{\mathbf{J}}_1(x)$ of Theorem 6.1 as

$$\overline{\mathbf{J}}_1(x) = \begin{bmatrix} 1 & 0 \\ 0 & j \end{bmatrix} \mathbf{J}_1(x) \begin{bmatrix} 1 & 0 \\ 0 & -j \end{bmatrix}. \tag{6.2.20}$$

Equation (6.2.1) can also be used to derive a useful series representation for isotropic random vector fields. Specifically, if we denote by

$$\vec{v}_n(r) = \int_0^\infty \mathbf{J}_n(\lambda r)\vec{V}_n(d\lambda) \tag{6.2.21}$$

we obtain the following theorem.

Theorem 6.5 *Any isotropic random vector field $\vec{v}(\vec{r})$ can be expanded in a series representation of the form*

$$\vec{v}(\vec{r}) = \sum_{n=-\infty}^{\infty} \mathbf{U}(\theta)\, e^{jn\theta}\vec{v}_n(r), \tag{6.2.22}$$

where the vectors $\{\vec{v}_n(r)\}$ are zero-mean uncorrelated 1-D random vector processes with

$$\begin{aligned} E[\vec{v}_n(r)\vec{v}_m^H(s)] &= \mathbf{k}_n(r,s)\delta_{n,m} \\ &= \int_0^\infty \mathbf{J}_n(\lambda r) \begin{bmatrix} S_p(\lambda)/\lambda^2 & 0 \\ 0 & S_s(\lambda)/\lambda^2 \end{bmatrix} \mathbf{J}_n^H(\lambda s)\, \lambda\, d\lambda. \end{aligned} \tag{6.2.23}$$

The proof of the above theorem follows directly from Theorem 6.4 and (6.2.21). Note that the coefficients $\vec{v}_n(r)$ can be easily obtained from $\vec{v}(\vec{r})$ by using the fact that

$$\frac{1}{2\pi}\int_0^{2\pi} \mathbf{U}^T(\theta)\mathbf{U}(\theta)e^{j(n-m)\theta}d\theta = I_2\delta_{n,m}, \tag{6.2.24}$$

where $\mathbf{U}^T(\theta)$ denotes the transpose of $\mathbf{U}(\theta)$. Hence, the coefficients $\vec{v}_n(r)$ can be computed from $\vec{v}(\vec{r})$ as

$$\vec{v}_n(r) = \frac{1}{2\pi}\int_0^{2\pi} \mathbf{U}^T(\theta)e^{-jn\theta}\,\vec{v}(\vec{r})\,d\theta. \tag{6.2.25}$$

Theorem 6.5 together with equation 6.2.25 is a generalization to the vector case of equations (2.1.9) and (2.1.15)-(2.1.17) which characterize the Fourier series expansions of isotropic random scalar fields. Finally, note that a decomposition of $\vec{v}(\vec{r})$ into its longitudinal and transverse coordinates can easily be obtained from (6.2.22). In particular, if we denote by $v_n^{(1)}(r)$ and $v_n^{(2)}(r)$ the components of $\vec{v}_n(r)$

along the r_1 and r_2 axes respectively in a Cartesian coordinate representation of the underlying plane, and if we let

$$v_l(\vec{r}) = \sum_{n=-\infty}^{\infty} v_n^{(1)}(r)e^{jn\theta} \tag{6.2.26}$$

and

$$v_t(\vec{r}) = \sum_{n=-\infty}^{\infty} v_n^{(2)}(r)e^{jn\theta} \tag{6.2.27}$$

we obtain from (6.2.22)

$$\vec{v}(\vec{r}) = v_l(\vec{r}) \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} + v_t(\vec{r}) \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}. \tag{6.2.28}$$

In (6.2.28) $v_l(\cdot)$ and $v_t(\cdot)$ are respectively the longitudinal and transverse components of $\vec{v}(\cdot)$. Note however that the representation (6.2.28) is not very useful since $v_l(\vec{r})$ and $v_t(\vec{s})$ are *not* uncorrelated unless $\vec{r}$ and $\vec{s}$ are colinear in which case we have

$$E\left[\begin{bmatrix} v_l(\vec{r}) \\ v_t(\vec{r}) \end{bmatrix} [v_l(\vec{s}) \quad v_t(\vec{s})]\right] = \begin{bmatrix} K_l(|\vec{r}-\vec{s}|) & 0 \\ 0 & K_t(|\vec{r}-\vec{s}|) \end{bmatrix}, \tag{6.2.29}$$

as was observed in Theorem 6.2.

6.3 CONCLUSION

We conclude this chapter by briefly discussing some estimation problems for isotropic random vector fields. We begin by looking at spectral estimation problems for isotropic random vector fields.

Note that according to Theorem 6.1, an isotropic random vector field is completely characterized by its potential and solenoidal power spectra. Hence, one may be interested in estimating these power spectra given measurements of the isotropic random vector field over some region of the plane. Equations (6.1.51) and (6.1.52) provide a means for estimating these spectra. Specifically, by taking the divergence of the isotropic random vector field we obtain an isotropic random scalar field with a power spectrum $\lambda^2 S_p(\lambda)$ which can then be estimated using the techniques of

Chapters 4 and 5. Similarly, by taking the curl of the isotropic random vector field and looking at the component of the resulting vector field which is normal to the plane, we obtain an isotropic random scalar field with a power spectrum $\lambda^2 S_s(\lambda)$ which again can be estimated by using the techniques of Chapters 4 and 5. However, such an approach involves differential operators and is thus numerically unstable. A challenging problem in this context is to come up with a numerically stable *local* operation that would yield an isotropic random scalar field with a power spectrum that is proportional to either $S_s(\lambda)$ or $S_p(\lambda)$. One can easily come up with non-local operators which can accomplish this. For example, if we are interested in estimating the solenoidal power spectrum of an isotropic random vector field $\vec{v}(\cdot)$, we can use equations (6.1.28) and (6.1.30) to compute the solenoidal part $\vec{v}_s(\cdot)$ of the vector field $\vec{v}(\cdot)$. However, $\vec{v}_s(\cdot)$ is a vector field, and therefore the spectral estimation techniques for scalar fields described in Chapters 4 and 5 cannot be applied directly to $\vec{v}_s(\cdot)$ to estimate $S_s(\lambda)$. To obtain a scalar field, we can apply the operator

$$D' = [\int_{-\infty}^{r_2} dr_2' \cdot -\pi \int_{-\infty}^{\infty} dr_2' \cdot \quad - \int_{-\infty}^{r_1} dr_1' \cdot +\pi \int_{-\infty}^{\infty} dr_1' \cdot] \tag{6.3.1}$$

to $\vec{v}_s(\cdot)$, which gives an isotropic random scalar field $z(\cdot)$ with a power spectrum $4S_s(\lambda)/\lambda^2$. However, both (6.1.30) and the use of D' involve global operations since one needs the value of $\vec{v}(\cdot)$ over the whole plane to compute $\vec{v}_s(\cdot)$ and $z(\cdot)$. In the case where $\vec{v}(\cdot)$ is not given over the whole plane, but only over a subregion $\mathcal{A}$ of the plane, the procedure (6.1.28)-(6.1.30) to separate $\vec{v}(\cdot)$ in its solenoidal and potential components needs to be adapted by replacing the Green's function $-\ln(|\vec{r}-\vec{s}|)/2\pi$ of the Laplacian operator over the whole plane by an equivalent Green's function $G(\vec{r},\vec{r'})$ defined over the subregion $\mathcal{A}$. Then if

$$\vec{w'}(\vec{r}) = \frac{1}{2\pi} \int_{\mathbf{R}^2} G(\vec{r},\vec{r'}) \vec{v}(\vec{r'})\, d\vec{r'} \tag{6.3.2}$$

we have

$$\begin{aligned}
\vec{v}_s(\vec{r}) &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \nabla \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \nabla \cdot \vec{w}'(\vec{r}) \\
&= \begin{bmatrix} \frac{\partial^2}{\partial r_2^2} & -\frac{\partial^2}{\partial r_1 \partial r_2} \\ -\frac{\partial^2}{\partial r_1 \partial r_2} & \frac{\partial^2}{\partial r_1^2} \end{bmatrix} \vec{w}'(\vec{r}) \qquad (6.3.3) \\
\vec{v}_p(\vec{r}) &= \nabla \nabla \cdot \vec{w}'(\vec{r}) \\
&= \begin{bmatrix} \frac{\partial^2}{\partial r_1^2} & \frac{\partial^2}{\partial r_1 \partial r_2} \\ \frac{\partial^2}{\partial r_1 \partial r_2} & \frac{\partial^2}{\partial r_2^2} \end{bmatrix} \vec{w}'(\vec{r}). \qquad (6.3.4)
\end{aligned}$$

The problem is then to estimate $S_s(\lambda)$ and $S_p(\lambda)$ given the values of $\vec{v}_s(\cdot)$ and $\vec{v}_p(\cdot)$ over $\mathcal{A}$. Note however that again, the procedure described above to separate $\vec{v}(\cdot)$ in its components $\vec{v}_s(\cdot)$ and $\vec{v}_p(\cdot)$ is not local since it involves the global operation (6.3.2) over the domain $\mathcal{A}$.

In some applications on the other hand, one might be interested in estimating the value of an isotropic random vector field rather than its power spectral density matrix. For example, one may be interested in estimating the value of an isotropic random vector field $\vec{v}(\cdot)$ given some noisy observations $\vec{y}(\cdot)$ of $\vec{v}(\cdot)$ where

$$\vec{y}(\vec{r}) = \vec{v}(\vec{r}) + \vec{n}(\vec{r}). \qquad (6.3.5)$$

In (6.3.5) $\vec{n}(\cdot)$ is a white noise vector field, i.e. a field which can be integrated along curves and whose covariance measures curve lengths. Specifically, if C is a smooth not necessarily closed curve of $\mathbf{R}^2$ then the random variable

$$a = \int_C \vec{n}(\vec{r}) \cdot d\vec{l} \qquad (6.3.6)$$

has zero-mean and a variance $E[a^2] = l(C)$, where $l(C)$ is the length of the curve C. Note that since $E[a^2] = l(C)$ then $\vec{n}(\cdot)$ cannot be a pure potential field, otherwise $E[a^2]$ would be independent of the shape of the curve C. The exact structure of white noise vector fields is not known at present and is the subject of current investigation (see for example [89]). The construction of such fields is of great importance in the development of a theory of Markovian random fields. Apart from the problem of

determining the structure of the white noise vector field $\vec{n}(\cdot)$, a challenging problem in this estimation context is to develop efficient recursive input/output estimation schemes for the isotropic random vector field $\vec{v}(\cdot)$ given the observations $\vec{y}(\cdot)$, similar to those developed in [41] for isotropic random scalar fields.

In some other cases, the isotropic random vector field $\vec{v}(\cdot)$ may obey a vector partial differential equation. For example, one may be interested in estimating the value of the electromagnetic vector potential $\vec{A}(\vec{r})$ where

$$\nabla^2 \vec{A}(\vec{r}) + \omega^2 \mu\epsilon \vec{A}(\vec{r}) = -\mu \vec{n}(\vec{r}) \tag{6.3.7}$$

$$\vec{H}(\vec{r}) = -\frac{1}{\mu} \nabla \times \vec{A}(\vec{r}) \tag{6.3.8}$$

$$\vec{E}(\vec{r}) = \frac{\nabla \nabla \cdot \vec{A}(\vec{r})}{j\omega\mu\epsilon} - j\omega \vec{A}(\vec{r}) \tag{6.3.9}$$

$$\vec{y}(\vec{r}) = \vec{E}(\vec{r}) + \vec{m}(\vec{r}) \tag{6.3.10}$$

In (6.3.8)-(6.3.10) $\vec{E}(\vec{r})$ and $\vec{H}(\vec{r})$ denote respectively the electric and magnetic fields and ω, ϵ and μ are respectively the steady state frequency of the electromagnetic field and the permittivity and permeability of the medium in which the electromagnetic field is propagating. The vector fields $\vec{n}(\cdot)$ and $\vec{m}(\cdot)$ are independent source and observation white noise vector fields respectively. The problem here is to develop a recursive estimation technique for $\vec{v}(\cdot)$, similar to the one developed in Chapter 3 for isotropic random scalar fields with an internal realization of the form (3.1.1)-(3.1.2).

We believe that the series representation (6.2.22) of isotropic random vector fields may provide a means for extending our approach and that of [41] to the above estimation problems involving isotropic random vector fields. Such spectral representations may also be important in deriving computationally efficient solutions to spectral estimation problems for isotropic random vector fields.

Chapter 7

CONCLUSION

In this thesis we have used Fourier series expansions of 2-D isotropic random fields to solve a number of recursive estimation and spectral estimation problems for such fields. In particular, we have shown in Chapter 2 that a wave-number limited isotropic random field, with a power spectrum that is zero outside of a disk of radius B centered at the origin of the wave-number plane, can be reconstructed in a mean-square sense from its observation on the countable set of circles of radii $r_i = a_{i,m}/B$, $i \in \mathbf{N}$, where $a_{i,m}$ is the ith zero of the mth order Bessel function $J_m(x)$, or of radii $r_i = i\pi/B$, $i \in N$. This result is a direct consequence of the sampling theorems that we derived for the Fourier coefficient processes associated with the given isotropic random field. Observe the parallel between our sampling procedures and the corresponding one-dimensional results. In one dimension, a stationary process can be reconstructed in the mean-square sense from its observation on a countable number of spheres in a space of dimension one, i.e. at a countably infinite number of points. In the general m-D case, an isotropic random field can be reconstructed in the mean-square sense from its observation on a countably infinite number of spheres in the m-D space.

We have also obtained in Chapter 3 efficient recursive estimation techniques for isotropic random fields described by non-causal internal differential realizations. By exploiting the properties of isotropic random fields, we showed that the problem of estimating an isotropic random field given noisy observations over a finite disk is equivalent to a countably infinite set of decoupled one-dimensional two-point

boundary value system (TPBV) estimation problems for the Fourier coefficient processes of the random field. The 1-D TPBV estimation problems are then solved using either the method of Adams et al. [2] or by using a Markovianization approach followed by standard 1-D smoothing techniques. Our smoothing scheme results in a processing structure that is recursive with respect to the radius r in a polar coordinate representation of the field.

In Chapter 4, we have presented a high resolution spectral estimation method for isotropic random fields with a covariance function equal to a weighted sum of cylindrical harmonics. Such fields are often used to model some types of background noise in geophysics and in ocean acoustics. The algorithm that we have obtained takes maximal advantage of the symmetries implied by the special structure of covariance functions which are equal to a weighted sum of cylindrical harmonics. Our approach is similar in spirit to the 1-D spectral estimation methods based on harmonic retrieval from an eigenanalysis of the covariance matrix. In the 2-D isotropic context, the spectral estimate is determined by performing an eigenanalysis of the covariance matrix of samples of the zeroth-order Fourier coefficient process corresponding to the given noisy observations of the underlying field. We have also discussed the estimation of this covariance matrix and presented examples to illustrate the high resolution and robustness properties of our procedure.

In Chapter 5, we have presented a new linear MEM algorithm for 2-D isotropic random fields. Our procedure differs from previous 2-D MEM algorithms by the fact that we take maximal advantage of the symmetries implied by isotropy which is the natural generalization to several dimensions of the 1-D notion of stationarity. Unlike general 2-D covariances, isotropic covariance functions which are positive definite on a disk are known to be extendible. Here, we have developed a computationally efficient procedure for computing the MEM isotropic spectral estimate corresponding to an isotropic covariance function which is given over a finite disk of radius $2R$. We have shown that the isotropic MEM problem has a linear solution which can be obtained by constructing the optimal linear filter for estimating the underlying isotropic field at a point on the boundary of a disk of radius R given noisy measurements of the field inside the disk. Our procedure is based on Fourier

series expansions in both the space and wave-number domains of the inverse of the MEM spectral estimate. Furthermore, our method is guaranteed to yield a valid isotropic spectral estimate and it is computationally efficient since it requires only $O(BRL^2)$ operations where L is the number of points used to discretize the interval $[0, R]$, and where B is the bandwidth of the spectrum that we want to estimate. We have also presented examples to illustrate the behavior of our algorithm and its high resolution property.

Finally, note that, as mentioned earlier, the results that we have developed in this thesis can easily be extended to isotropic random fields in higher dimensions, provided that the random fields are expanded in spherical harmonics instead of in Fourier series.

7.1 OPEN TECHNICAL QUESTIONS

In this section we discuss some interesting areas in which future research is needed and may be fruitful. Some of the topics that we discuss involve a generalization of the work that we have presented in this thesis and some may require taking a different approach. We begin by discussing some issues that are suggested by the recursive estimation part of this thesis. (See Chapter 3.) We then present some possible directions in which the work that we have done on spectral estimation can be extended. We do not discuss in this section open problems involving random vector fields. The reader should refer to Chapter 6 for a discussion of these problems.

1. **Recursive estimation with non-white noise sources:**
   Note that most the results presented in this thesis are based on the Fourier series expansions of isotropic random fields. Note also that, as mentioned in Chapter 2, such Fourier series expansions are really Karhunen-Loeve expansions of isotropic random fields in terms of the coordinate angle θ in a polar coordinate representation of the underlying plane. In a number of applications of practical interest, the source term appearing on the right hand side of (3.1.1) is not spatially white but has some directionality associated to it and it turns out that a Karhunen-Loeve expansion approach does not help

in such a case. When the source term $u(\cdot)$ is not isotropic, its covariance function $K_u(\vec{r},\vec{s})$ is *not* symmetric in the r and s variables nor in the θ and ϕ variables, where $\vec{r} = (r,\theta)$ and $\vec{s} = (s,\phi)$. Hence, the covariance function $K_u(\cdot,\cdot)$ does not possess a symmetric Mercer's expansion in terms of orthogonal eigenfunctions that depend solely on r or θ, and there is no obvious way in which the expansion approach that we have used in this thesis can be generalized. Therefore, alternative procedures need to be derived to tackle the problem of a non-isotropic source term.

2. **Angular Processing of isotropic random fields:**
An interesting extension of the results of Chapter 3 is the development of alternate "recursive " structures. In particular, one can imagine developing algorithms that process data recursively in the angular direction as shown in Figs. 7.1 a and 7.1 b, rather than radially (see [66] for such an algorithm in a different setting). For example, one may try to expand the isotropic random fields in terms of complete orthogonal sets of functions of the polar coordinate variable r. However, such expansions do not seem to lead to uncorrelated 1-D estimation problems for the coefficients appearing in the Karhunen-Loeve expansions of the underlying fields. Note that these coefficients are now random functions of the coordinate variable θ. Developing alternative angular processing algorithms for isotropic random fields may prove as difficult as the above problem of developing efficient recursive estimation techniques for the signal field when the noise term appearing on the right hand side of (3.1.1) is spatially non-white.

3. **Eigenstructure spectral estimation procedures for more general background noise models:**
The covariance functions that we have considered in Chapter 4 are a special case of the model [36]

$$K(r) = \sum_{p=1}^{P} A_p H_0^{(1)}(\lambda_p r) + A_p^* H_0^{(1)*}(\lambda_p r) \qquad (7.1.1)$$

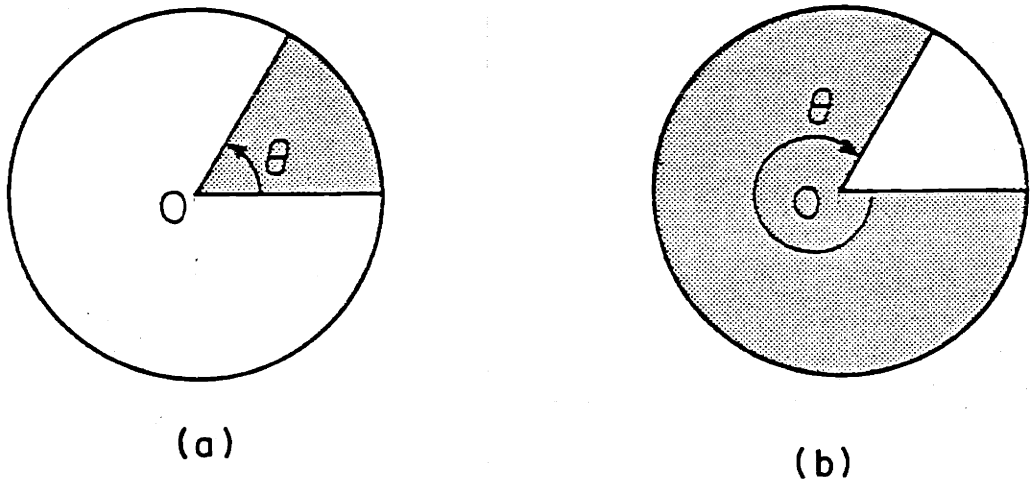

Figure 7.1: Counterclockwise and clockwise angular recursions.

where A_p^* denotes the complex conjugate of A_p. In (7.1.1) A_p and λ_p are complex numbers and $H_0^{(1)}(\cdot)$ denotes a Hankel function of the first type and of order zero. The imaginary parts of the complex frequencies $\{\lambda_p\}$ are assumed to be positive. This assumption guarantees that $K(r)$ in (7.1.1) is a valid covariance function. When all of the A_p and λ_p are real, model (7.1.1) reduces to the model (4.1.3) that we studied in Chapter 4. Since covariance functions of the form (7.1.1) are often used to model some types of background noise in ocean acoustics, the development of efficient procedures for estimating P, $\{A_p\}$ and $\{\lambda_p\}$ for $1 \leq p \leq P$ given some noisy measurements of the background fields is important. However, even though the covariance function of the zeroth order Fourier coefficient process corresponding to a background noise field with a covariance function of the form (7.1.1) has the nice structure

$$k_0(r,s) = \sum_{p=1}^{P} A_p J_0(\lambda_p r_<) H_0^{(1)}(\lambda_p r_>) + A_p^* J_0^*(\lambda r_<) H_0^{(1)*}(\lambda_p r_>) \qquad (7.1.2)$$

where $r_< = \min(r,s)$ and $r_> = \max(r,s)$, the methods of Chapter 4 do not carry over to this more general case. The approach of Chapter 4 is predicated on the fact that the matrix S of normalized sample values of the covariance function of the zeroth order Fourier coefficient process (see (4.2.5)) has rank P. Unfortunately, this is not true for covariance functions of the form (7.1.1). For example, the 2×2 S matrix corresponding to a background covariance function of the form

$$K(r) = H_0^{(1)}(\lambda_0 r) + H_0^{(1)*}(\lambda_0 r) \qquad (7.1.3)$$

has the structure

$$S = [\sqrt{r_i r_j}(J_0(\lambda_0 r_i) H_0^{(1)}(\lambda_0 r_j) + J_0^*(\lambda_0 r_i) H_0^{(1)*}(\lambda_0 r_j))]. \qquad (7.1.4)$$

Clearly, S has rank 2 in general. We speculate that the special structure of covariance functions of the form (7.1.1) together with the properties of Hankel and Bessel functions may lead to efficient estimation algorithms for the parameters of covariance functions of the form (7.1.1).

4. **Modeling data by a sum of zeroth-order Hankel functions of the first type:**

   Another interesting issue related to the one that we have discussed above is the issue of modeling data measurements by a sum of zeroth-order Hankel functions of the first type with an additive 2-D white Gaussian noise component. Specifically, the measurements are assumed to come from a model of the form

$$y(\vec{r}) = \sum_{p=1}^{P} A_p H_0^{(1)}(\lambda_p r) + n(\vec{r}) \tag{7.1.5}$$

   where the amplitudes $\{A_p\}$ and the frequencies $\{\lambda_p\}$ are complex numbers and where the imaginary parts of the frequencies $\{\lambda_p\}$ are assumed to be positive. In (7.1.5) $n(\cdot)$ is a 2-D white Gaussian noise process. The challenge here is to determine P, the amplitudes $\{A_p\}$ and the frequencies $\{\lambda_p\}$. The problem that we have described arises in electromagnetic theory where the field $y(\cdot)$ corresponds to noisy measurements of 2-D cylindrical waves propagating in a lossy medium. Note that this problem is similar to the 1-D Prony problem of fitting noisy data with a sum of complex exponentials [33]. In the 1-D case, existing solutions to Prony's method are based on the observation that a sum of exponentials is the homogeneous solution of a constant coefficients difference equation. In the 2-D case, a sum of zeroth-order Hankel functions of the first type is the homogeneous solution of a constant coefficient differential equation involving powers of the Laplacian operator. In theory, it is possible to use this fact to determine all the parameters of model (7.1.5). However, in practice any method involving the use of powers of the Laplacian operator is numerically unstable, and it is desirable to develop alternative techniques for estimating P, the amplitudes $\{A_p\}$ and the frequencies $\{\lambda_p\}$ in (7.1.5) based on the properties of the Hankel functions [1], [7].

5. **Maximum entropy spectral estimation for more general homogeneous fields:**

   There are several directions in which one can try to generalize the work presented in Chapter 5. For example, 2-D covariance functions which are

constant along ellipses rather than along circles, arise in some cases of practical interest. Such covariance functions become radially symmetric under an appropriate scaling and rotation of the underlying coordinate axes, and the techniques of Chapter 5 can then be used to estimate a warped version of the power spectrum of the underlying random field. A challenging problem here is to develop an algorithm for finding the correct scaling and rotation operations to be performed given limited measurements of the random field. More generally, another interesting problem is to extend some of the ideas that appear throughout Chapter 5 and in particular the spectral factorization result, to homogeneous, but not necessarily isotropic, covariance functions which are defined continuously over the plane. This will require a study of filtering problems for homogeneous fields aimed at determining whether the homogeneity property can be exploited in higher dimensional spaces to develop fast filtering algorithms.

6. **Lattice structures for isotropic random fields:**
As mentioned in Chapter 1, the Levinson recursions of 1-D prediction give rise to lattice realizations of the underlying stationary process [56]. These lattice realizations have the feature that their stability can be checked by inspection, and have also the property of having a very low sensitivity to roundoff errors. The existence of fast Levinson-like equations for the construction of the linear least squares estimate of a 2-D isotropic random field given noisy observations of the field over a finite disk [41], and the fact that these equations were used in this thesis to develop a fast MEM spectral estimation procedure leads us to speculate that lattice realizations for the Fourier coefficient processes corresponding to a 2-D isotropic random field may indeed exist. The search for such realizations should be an integral part of any future research effort focusing on 2-D isotropic random fields.

Bibliography

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications Inc., New York, N.Y. 1972. Vol. 67, No. 6, pp. 1988-1996, June 1980.

[2] M. B. Adams, A. S. Willsky and B. C. Levy, "Linear Estimation of Boundary value Processes," *IEEE Trans. Automat. Control*, Vol. AC-29, No. 9, pp. 803-821, Sept. 1984.

[3] W. F. Ames, *Numerical Methods for Partial Differential Equations*, Academic Press, New York, N. Y., 1977.

[4] S. Attasi, "Modelling and Recursive Estimation for Double Indexed Sequences," in *Systems Identification: Advances and Case Studies*, R. K. Mehra and D. G. Lainiotis, eds., Academic Press, New York, N.Y. 1976.

[5] A. B. Baggeroer, *Space/Time Random Processes and Optimum Array Processing*, Naval Undersea report NUC TP 506, San Diego, CA 1976.

[6] A. B. Baggeroer, "Sonar Signal Processing," in *Applications of Digital Signal Processing*, A. V. Oppenheim, Ed., Prentice-Hall, Englewood Cliffs, N.J. 1978.

[7] H. Bateman, *Higher Transcendal Functions*, Volume II, McGraw-Hill, New-York, N.Y. 1953.

[8] R. E. Bellman, *Matrix Analysis*, 2nd ed., McGraw-Hill Inc., New York, N.Y. 1968.

[9] G. Bienvenu and L. Kopp, "Adaptivity to Background Noise Spatial Coherence for High Resolution Passive Methods," in *Proc. ICASSP 80*,

[10] A. Blanc-Lapierre and P. Faure, " Stationary and Isotropic Random Functions," in *Bernoulli, Bayes, Laplace Anniversary Volume*, Springer-Verlag, Berlin 1965.

[11] N. S. Burdic, *Underwater Acoustic System Analysis*, Prentice Hall Inc., Englewood Cliffs, N. J., 1984.

[12] S. M. Candel, "Simultaneous Calculation of Fourier-Bessel Transforms up to Order N," *J. Comput. Phys.*, vol. 23, pp. 343-353, 1981.

[13] J. Capon, "Maximum-Likelihood Spectral Estimation," in *Nonlinear Methods of Spectral Analysis*, S. Haykin, Ed., Springer-Verlag, New York, N.Y. 1979.

[14] G. F. Carrier, M. Krook and C. E. Pearson, *Functions of a Complex Variable; Theory and Technique*, McGraw-Hill, New York, NY, 1966.

[15] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill Inc., New York, N.Y. 1955.

[16] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, Orlando, FL. 1984.

[17] J. J. Dongarra et al., *Linpack User's Guide*, SIAM, Philadelphia, PA., 1979.

[18] H. Dym and I. Gohberg, "On an Extension Problem, Generalized Fourier Analysis, and an Entropy Formula," *Integral Equations and Operator Theory*, vol. 3, pp. 143-215, 1980.

[19] M. P. Ekstrom and J. W. Woods, "2-D Spectral Factorization with Applications in Recursive Digital Filtering," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-24, no. 2, pp. 115-178, Apr. 1976.

[20] A. Erdelyi, *Table of Integral Transforms*, Volume I, McGraw-Hill, New-York, N.Y. 1954.

[21] D. C. Fraser, *A New Technique for Optimal Smoothing of Data*, Sc.D. Dissertation, M.I.T., Cambridge, MA, 1967.

[22] I. C. Gohberg and M. G. Krein, *Theory and Applications of Volterra Operators in Hilbert Space*, Trans. Math. Monographs, vol. 24, Amer. Math. Soc., Providence, R.I., 1970.

[23] J. L. Griffith, "Hankel Transforms of Functions Zero Outside a Finite Interval," *Proc. Roy. Soc. New South Wales*, vol. 89, pp. 109-115, 1955.

[24] A. Habibi, "Two Dimensional Bayesian Estimates of Images," *Proc. IEEE*, Vol. 60, No. 7, pp. 878-883, July 1972.

[25] O. S. Halpeny and D. G. Childers, "Composite Wavefront Decomposition via Multidimensional Digital Filtering of Array Data," *IEEE Trans. Circuits and Syst.*, vol. CAS-22, pp. 552-562, June 1975.

[26] E. W. Hansen, "Fast Hankel Transform Algorithm," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-33, no. 3, pp. 666-671, June 1985.

[27] K. Ito, "Isotropic Random Current," in *Proc. 3rd Berkeley Symposium on Math. Stat. and Prob.*, pp. 125-132, 1956.

[28] A. K. Jain and E. Angel, "Image Restoration Modelling and Reduction of Dimensionality," *IEEE Trans. Computers*, Vol. C-23, No. 5, pp. 470-476, May 1974.

[29] A. K. Jain and S. Raganath, "Two-Dimensional Spectral Estimation," in *Proc. RADC Spectrum Estimation Workshop*, Rome, N. Y. May 1978, pp. 217-225.

[30] A. J. Jerri, "The Various Extensions and Applications of Shannon's Sampling Theorem: A Tutorial Review," *Proc. IEEE*, Vol. 65, No. 11, pp. 1565-1596, Nov. 1977.

[31] P. R. Julian and A. Cline, "The Direct Estimation of Spatial Wavenumber Spectra of Atmospheric Variables," *Jour. Atmospheric Sci.*, vol. 31, pp. 1526-1539, 1976.

[32] T. Kailath, *Lectures on Wiener and Kalman Filtering*, Springer-Verlag, New York, N. Y., 1981.

[33] S. M. Kay and S. L. Marple, "Spectrum Analysis - A Modern Perspective," *Proc. IEEE*, vol. 69, no. 11, pp. 1380-1419, Nov. 1981.

[34] E. A. Kraut, *Fundamentals of Mathematical Physics*, McGraw-Hill Inc., New York, N. Y., 1967.

[35] A. J. Krener, "Acausal Realization Theory, Part I: Linear Deterministic Systems," to appear in *SIAM J. Control and Opt.*.

[36] W. A. Kuperman and F. Ingenito, "Spatial Correlation of Surface Generated Noise in a Stratified Ocean," *J. Acoust. Soc. Am*,

[37] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, John Wiley and Sons Inc., New York, N.Y. 1972.

[38] S. W. Lang and J. H. McClellan, "The Extension Of Pisarenko's Method to Multiple Dimensions," in *Proc. ICASSP 82*, Paris, France, May 1982.

[39] S. W. Lang and J. H. McClellan, "Multidimensional MEM Spectral Estimation," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-30, no. 6, pp. 880-887, Dec. 1982.

[40] L. Lapidus and G. F. Pinder, *Numerical Solution of Partial Differential Equations in Science and Engineering*, John Wiley, New York, N. Y., 1982.

[41] B. C. Levy and J. N. Tsitsiklis, "A fast Algorithm for Linear Estimation of Two-Dimensional Isotropic Random Fields," *IEEE Trans. Inform. Theory*, vol. IT-31, no. 5, pp 635-644, Sept. 1985.

[42] J. S. Lim and N. A. Malek, "A New Algorithm for Two-Dimensional Maximum Entropy Power Spectrum Estimation," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-29, pp. 401-413, June 1981.

[43] J. Makhoul, "Spectral Analysis of Speech Estimation," *IEEE Trans. Audio Electroacoust*, vol. AU-21, pp. 140-148, June 1973.

[44] T. L. Marzetta, "Two-Dimensional Linear Prediction: Autocorrelation Arrays, Minimum-Phase Prediction Error Filters, and Reflection Coefficient Arrays," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-28, pp. 726-733, Dec. 1980.

[45] D. Q. Mayne, "A Solution of the Smoothing Problem for Linear Dynamic Systems," *Automatica*,

[46] J. H. McClellan and R. J. Purdy, "Applications of Digital Signal Processing to Radar," in *Applications of Digital Signal Processing*, A. V. Oppenheim, Ed., Prentice-Hall, Englewood Cliffs, N.J., 1978.

[47] J. H. McClellan, "Multidimensional Spectral Estimation," *Proc. IEEE*, vol. 70, pp. 1029-1039, Sept. 1982. Vol. 4, pp. 73-92, 1966.

[48] R. N. McDonough, "Applications of the Maximum-Likelihood Method and the Maximum-Entropy Method to Array Processing," in *Nonlinear Methods of Spectral Analysis*, S. Haykin, Ed., Springer-Verlag, New York, N.Y. 1979.

[49] C. L. Mehta and E. Wolf, "Coherence Properties of Black Body Radiation," Parts I and II, *Physical Review*, Vol. 134, No. 4, pp. A 1143-A 1153, 1964.

[50] A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics: Mechanics of Turbulence*, Volume 2, MIT Press, Cambridge, MA., 1975.

[51] D. R. Mook, "An Algorithm for the Numerical Evaluation of the Hankel and Abel Transforms," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-31, pp. 979-985, August 1983.

[52] J. J. Murray, "Spectral Factorization and Quarter-plane Digital Filters," *IEEE Trans. Circuits and Systems*, vol. CS-25, no. 8, pp. 586-592, Aug. 1978.

[53] R. G. Ogier and E. Wong, "Recursive Linear Smoothing for Two Dimensional Random Fields," *IEEE Trans. Inform. Theory*, Vol. IT-27, No. 1, pp. 72-82, Jan. 1981.

[54] H. Ogura, "Spectral Representation of Vector Random Fields," *Jour. Phys. Soc. of Japan*, Vol. 24, pp. 1370-1380, June 1968.

[55] A. Papoulis, *Systems and Transforms with Applications in Optics*, MacGraw-Hill, New York, N. Y., 1968.

[56] A. Papoulis, "Maximum Entropy and Spectral Estimation: A Review," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-29, no. 6, pp. 1176-1186, Dec. 1981.

[57] D. P. Petersen and D. Middleton, "Sampling and Reconstruction of Wave-number Limited Functions in N-dimensional Euclidean Spaces," *Inform. Contr.*, Vol. 5, pp. 279-323, 1962.

[58] V. F. Pisarenko, "The Retrieval of Harmonics from a Covariance Function," *Geophys. J. Roy. Astron. Soc.*, Vol. 33, pp. 347-366, 1973.

[59] L. D. Pitt, "A Markov Property for Gaussian Processes with a Multidimensional Parameter," *Arch. Rat. Mech. Anal.*, Vol. 43, pp. 367-391, 1971.

[60] M. Plancherel and G. Polya, "Fonctions Entieres et Integrales de Fourier Multiples," *Comment. Math. Helv.*, vol. 9, pp 224-248, 1937.

[61] H. E. Rauch, F. Tung and C. T. Striebel, "Maximum-Likelihood Estimates of Linear Dynamic Systems," *AIAA J.*, Vol. 3, pp. 1445-1450, 1965.

[62] R. D. Reasenberg, "The Moment of Inertia and Isostacy of Mars," *Jour. Geophys. Research*, Vol. 82, pp. 369-375, Jan. 1977.

[63] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Ungar, New York, N.Y. 1955.

[64] H. P. Robertson, "The Invariant Theory of Isotropic Turbulence," *Proc. Cambridge Phil. Soc.*, Vol. 36, pp. 209-223, 1940.

[65] E. A. Robinson and S. Treitel, "Maximum Entropy and the Relationship of the Partial Autocorrelation to the Reflection Coefficients of a Layered System," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-28, no. 2, pp. 224-235, April 1980.

[66] G. Rodriguez, "A Function Space Approach to State and Model Error Estimation for Elliptic Systems," in *Proc. 22nd IEEE Conf. on Decision and Control*, San Antonio, Texas, pp. 256-267, Dec. 1983.

[67] I. Rodriguez-Iturbe and J. M. Mejia, "The Design of Rainfall Networks in Space and Time," *Water Resour. Res.*, vol. 10, no. 4, Aug. 1974.

[68] L. I. Ronkin, *Introduction to the Theory of Entire Functions of Several Variables*, Trans. Math. Monographs, vol. 44, Amer. Math. Soc., Providence, R.I., 1974.

[69] S. E. Roucos and D. G. Childers, "A Two Dimensional Maximum Entropy Spectral Estimator," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 554-560, Sept. 1980.

[70] W. Rudin, "The Extension Problem for Positive-Definite Functions," *Ill. J. Math.*, vol. 7, pp. 532-539, 1963.

[71] W. Rudin, "An Extension Theorem for Positive-Definite Functions," *Duke Math. J.*, vol. 37, pp. 49-53, 1970.

[72] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, N.Y. 1974.

[73] R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," in *Proc. RADC Spectrum Estimation Workshop*, Griffths AFB, N.Y., pp. 243-258, Oct. 1979.

[74] J. P. Schott and J. H. McClellan, "Maximum Entropy Power Spectral Estimation with Uncertainty in Correlation Measurements," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-32, pp. 410-418, April 1984.

[75] G. Sharma and R. Chellappa, "A Model Based Approach for Estimation of Two-Dimensional Maximum Entropy Power Spectra," *IEEE Trans. Inform. Theory*, vol. IT-31, no. 1, pp 90-99, Jan. 1985.

[76] M. Shinozuka and C.-M. Jan, "Digital Simulation of Random Processes and its Applications," *J. Sound Vib.*, vol. 25, no. 1, pp. 111-128, 1972.

[77] D. R. Smith, *Variational Methods in Optimization*, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1974.

[78] G. Strang, *Linear Algebra and its Applications*, Academic Press, New York, N.Y., 1980.

[79] E. C. Titchmarsh, *The Theory of Functions*, Second Edition, Oxford University Press, London, UK 1939.

[80] F. G. Tricomi, *Integral Equations*, Interscience Publishers, New York, N.Y. 1957.

[81] H. L. Van Trees, *Detection, Estimation and Modulation Theory*, vol. 1, John Wiley, New York, NY, 1968.

[82] G. Verghese and T. Kailath, "A Further Note on Backwards Markovian Models," *IEEE Trans. Inform. Theory*, Vol. IT-25, No. 1, pp. 121-124, Jan. 1979.

[83] M. Wax and T. Kailath, "Determining the Number of Signals by Information Theoretic Criteria," in *Proc. ICASSP 84*, San Diego, California, March 1984.

[84] M. Wax and T. Kailath, "Covariance Eigenstructure Approach to 2-D Harmonic Retrieval," in *Proc. ICASSP 83*, Boston, MA, April 1983.

[85] S. J. Wernecke and L. R. D'Addario, "Maximum Entropy Image Reconstruction," *IEEE Trans. Comput.*, vol. C-26, pp. 351-364, April 1977.

[86] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, U.K., 1965.

[87] E. Wong, "Recursive Causal Linear Filtering for Two Dimensional Random Fields," *IEEE Trans. Inform. Theory*, Vol. IT-24, No. 1, pp. 50-59, Jan. 1978.

[88] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems*, Springer-Verlag, New York, N.Y. 1985.

[89] E. Wong, "In Search of Multiparameter Markov Processes," in *Communications and Networks: A Survey of Recent Advances*, I. F. Blake and H. V. Poor editors, Springer-Verlag, New York, N.Y. 1986.

[90] J. W. Woods, "Two Dimensional Markov Spectral Estimation," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 232-240, March 1972.

[91] J. W. Woods and C. H. Radewan, "Kalman Filtering in Two Dimensions," *IEEE Trans. Inform. Theory*, Vol. IT-23, No. 4, pp. 473-482, July 1977.

[92] M. I. Yadrenko, *Spectral Theory of Random Fields*, Optimization Software Inc., New York, N.Y. 1986.

[93] A. M. Yaglom, "Some Classes of Random Fields in N-Dimensional Space Related to Stationary Random Processes," *Theory of Prob. and Appl.*, vol. II, pp. 273-319, 1957.

[94] A. M. Yaglom, "Second Order Homogeneous Random Fields," *Proc. 4th Berkeley Symp. on Math. Stat. and Prob.*, vol. II, pp. 593-622, Univ. of California Press, Berkeley, CA 1961.

[95] A. M. Yaglom, "Some Mathematical Models Generalizing the Model of Homogeneous and Isotropic Turbulence," *Jour. Geophys. Research*, vol. 67, no. 8, pp. 3081-3087, July 1962.