ELECTRICAL AND COMPUTER ENGINEERING DEPARTMENT

CLEMSON UNIVERSITY

CLEMSON, SC 29634-0915

TR-011795-3570P

p- 111

ON THE JOINT SPECTRAL DENSITY OF BIVARIATE RANDOM SEQUENCES

by

David D. Aalfs

Radar Systems Laboratory
Technical Report No. 21

On The Joint Spectral Density Of Bivariate Random Sequences

by

David D. Aalfs

Technical Report #21
January 17, 1995

Radar Systems Laboratory
Electrical and Computer Engineering Department
Clemson University
Clemson, SC 29634-0915

Windshear Detection Radar Signal Processing Studies
Grant NGT- 50414 and NAG-928
National Aeronautics and Space Administration
Langley Research Center
Hampton, VA 23665

ABSTRACT

For univariate random sequences, the power spectral density acts like a probability density function of the frequencies present in the sequence. This dissertation extends that concept to bivariate random sequences. For this purpose, a function called the joint spectral density is defined that represents a joint probability weighting of the frequency content of pairs of random sequences. Given a pair of random sequences, the joint spectral density is not uniquely determined in the absence of any constraints. Two approaches to constraining the sequences are suggested: (1) assume the sequences are the margins of some stationary random field, (2) assume the sequences conform to a particular model that is linked to the joint spectral density. For both approaches, the properties of the resulting sequences are investigated in some detail, and simulation is used to corroborate theoretical results. It is concluded that under either of these two constraints, the joint spectral density can be computed from the non-stationary cross-correlation.

DEDICATION

To Laurie

ACKNOWLEDGMENTS

The author would like to express his gratitude to Dr. E. G. Baxa for his valuable guidance and encouragement. He also thanks Dr. J. J. Komo, Dr. R. J. Schalkoff, and Dr. P. Kiessler for their support and helpful criticism as members of the committee. Dr. Kiessler is especially acknowledged for making the time for what the author found to be very fruitful discussions.

The National Aeronautics and Space Administration is to be thanked for its financial support under the Graduate Student Researchers Program. Special thanks are extended to the Antenna and Microwave Research Branch at Langley Research Center for their technical support and Mohamed Moosa for his friendship and assistance.

Finally, he expresses his appreciation to his wife and family for their patience and support throughout his graduate studies.

TABLE OF CONTENTS

Page

TITLE PAGE i

ABSTRACT ii

DEDICATION iii

ACKNOWLEDGMENTS iv

LIST OF FIGURES vii

CHAPTER

1. INTRODUCTION. 1

   Problem Statement 1
   The Radar Problem 3
   A Note on Terminology 4
   Contribution to the Field 6
   Organization 7

2. THEORY OF STATIONARY RANDOM SEQUENCES. 8

   Distribution Functions 8
   Density Functions 10
   Characteristic Functions 11
   Stieltjes Integral 12
   Moments 12
   Stationarity 13
   Properties of Autocorrelation Functions 14
   Spectral Representation of Correlation Functions 15
   Spectral Representation of Stationary Random Sequences 16
   Orthogonal Increment Process 18
   Brownian Motion 19

3. JOINT SPECTRAL DENSITY. 21

   Bivariate Correlation and Spectral Properties 22
   Stationary Random Fields 25
   Properties of $Z(\lambda_1, \lambda_2)$ 28
   Joint Spectral Density 29

Table of Contents (Continued)

	Page
4. MARGINAL SEQUENCES.	38
Introduction	38
Basic Concepts	38
Properties in the Spectral Domain	41
Jointly Stationary Marginal Sequences	44
Some Examples	46
Uniform Margins	47
Gaussian Margins	48
Jointly Harmonizable Marginal Sequences	50
Some Examples	54
Generalized Brownian Motion	54
Bivariate Gaussian	57
Simulating Marginal Spectra with Particular JSDs	60
Jointly Stationary with Gaussian Marginal Spectra	61
Jointly Harmonizable with Bivariate Gaussian JSD	64
Summary	70
5. AN EXPONENTIAL MODEL.	71
Correlation and Spectral Properties	72
Simulations	75
Single Exponential ($K = 1$)	76
Multiple Exponentials ($K > 1$)	78
Summary	79
6. CONCLUSIONS.	81
Motivation	81
Properties of the Joint Spectral Density	81
Marginal Sequences	83
Exponential Model	83
Future Work	84
Final Assessments	85
APPENDICES	86
A. Families of Bivariate Uniform Distributions.	87
B. Some Properties of the Gaussian Distribution.	93
REFERENCES	100

LIST OF FIGURES

Figure		Page
1.1	Scan region	5
1.2	Range cell volume	5
3.1	Covariance of a two-dimensional orthogonal increment process	30
3.2	Example of a two-dimensional distribution function, $F(\lambda_1, \lambda_2)$	32
3.3	A slice of the two-dimensional correlation function and the corresponding projection of the spectral density	36
3.4	Schematic diagram of the relationships among various time-domain and spectral representations	37
4.1	Marginal process of a two-dimensional random field	39
4.2	Domain of a two-dimensional orthogonal increment process	44
4.3	Contour plot of $F(\lambda_1, \lambda_2)$ for jointly stationary marginal sequences	47
4.4	Plot of $F(\lambda_1, \lambda_2)$ for jointly stationary marginal sequences with uniform marginal spectra	49
4.5	Plot of $F(\lambda_1, \lambda_2)$ for jointly stationary marginal sequences with Gaussian marginal spectra	51
4.6	Plot of $R(\tau, \upsilon)$ for jointly stationary marginal sequences with Gaussian marginal spectra	52
4.7	Plot of equation (4.16) for continuous τ	56
4.8	Plot of equation (4.17) for discrete τ	56
4.9	Plot of $R(\tau, \upsilon)$ for continuous τ and υ	[illegible]
4.10	Plot of $P(\lambda_1, \lambda_2)$	[illegible]
4.11	Plot of the marginal spectra $f_{11}(\lambda)$ and $f_{22}(\lambda)$	[illegible]
4.12	Plot of the JSD, $\hat{f}(\lambda_1, \lambda_2)$	[illegible]
4.13	Plot of $P(\lambda_1, \lambda_2)$	[illegible]
4.14	Contour plot of $P(\lambda_1, \lambda_2)$	[illegible]

List of Figures (Continued)

Page

4.15 Plot of $\hat{f}(\lambda_1, \lambda_2)$. 68

4.16 Contour plot of $\hat{f}(\lambda_1, \lambda_2)$. 69

5.1 Power spectral density of $X(t)$ for K=256. 76

5.2 Contour plot of $\hat{f}_{\lambda\nu}(\lambda, \nu)$ $(K = 1, \rho = 0)$..... 77

5.3 Contour plot of $\hat{f}_{\lambda\nu}(\lambda, \nu)$ $(K = 1, \rho = .5)$. 77

5.4 Contour plot of $\hat{f}_{\lambda\nu}(\lambda, \nu)$ $(K = 1, \rho = .9)$. 78

5.5 Contour plot of $\hat{f}_{\lambda\nu}(\lambda, \nu)$ $(N = 1000, K = 2, \rho = .9)$. 80

A-1. Morgenstern's uniform density for $\kappa = 0$. 88

A-2. Morgenstern's uniform density for $\kappa = 0.5$..... 88

A-3. Morgenstern's uniform density for $\kappa = 1$. 89

A-4. Morgenstern's uniform density for $\kappa = -0.5$. 89

A-5. Morgenstern's uniform density for $\kappa = -1$..... 90

A-6. Plackett's uniform density for $\psi = 0$. 90

A-7. Plackett's uniform density for $\psi = 0.5$..... 91

A-8. Plackett's uniform density for $\psi = 0.2$..... 91

A-9. Plackett's uniform density for $\psi = 2$. 92

A-10.Plackett's uniform density for $\psi = 5$. 92

B-1. 2-dimensional Gaussian with $\rho = 0$ 95

B-2. 2-dimensional Gaussian with $\rho > 0$ 96

B-3. 2-dimensional Gaussian with $\rho < 0$ 97

B-4. Family of ellipses for $r = 0$..... 98

B-5. Family of ellipses for $r \neq 0$..... 99

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

The power spectral density (PSD) of a random sequence can be loosely interpreted as a probability density function (PDF) of the frequency content of that sequence. Stated more precisely, if the PSD of a random sequence $x(t)$ is denoted by $f(\lambda)$, then $f(\lambda)\,d\lambda$ represents the expected value of the portion of the total power in $x(t)$ due to components with frequency in the interval $(\lambda, \lambda + d\lambda]$ [1]. Normalizing $f(\lambda)$ by the total power gives a probability weighting for each frequency analogous to a PDF of the frequency. As such, the power spectral density has all the properties of a probability density function. This is one of many parallels between the spectral theory of random sequences and the probability theory of random variables. The integral of the PSD, sometimes called the integrated spectrum or the spectral distribution function, is analogous to a probability (cumulative) distribution function. Also, in the time domain, the auto-covariance is closely related to the class of characteristic functions. Much of the terminology echoes the similarity between concepts in these two bodies of theory. In this context it makes sense to talk about the statistics of the spectrum where frequency is treated as a random variable. For example, computing the spectral moments is often of considerable interest in applications such as radar signal processing [2]. Covariance based approaches to estimating the spectral mean and width can be understood in terms of the moment theorem from probability theory applied to covariances and spectral densities instead of characteristic functions and probability densities [3].

This one-to-one correspondence begins to break down when the joint statistics of more than one random sequence are under consideration. Consider the bivariate case of two random sequences. What we would like to have is a single function of

two frequency arguments that acts like a joint probability density function for the frequency content in the two sequences. Conventional spectral analysis of bivariate random sequences involves four functions of a single frequency argument that are typically written as a two-by-two matrix valued function called the spectral density matrix. The PSDs of the two sequences, sometimes called their auto-spectra, are placed along the main diagonal and their cross-spectra are placed along the cross diagonal. In essence, the auto-spectra correspond to the marginal probability densities of the frequency content for each sequence, and the cross-spectra contribute information about the relationships between the two sequences. It is well known from probability theory that the joint probability density function cannot be uniquely determined based on knowledge of the marginals [4]. Thus the ability to form this desired function via conventional bivariate analysis depends on whether the cross-spectra contain the extra information that is necessary to uniquely determine the joint probability density function.

This function that acts like a joint probability density function of the frequency content in the two sequences will be called the joint spectral density (JSD). The terminology, joint spectral density, requires some clarification because the individual words (joint, spectral, and density) are rather generic and have been used in various combinations to denote things that are not intended here. To help reduce the risk of confusion, first consider what is NOT meant by the joint spectral density as the term is defined in this dissertation. It is not the spectral density matrix of a bivariate random function. It is not the coherency, or the magnitude squared coherency. It is not the bispectrum or any other higher-order spectrum. What is meant by the joint spectral density is a joint probability weighting of the frequency content in a set of two or more random sequences that is completely analogous to a joint probability density function.

Why bother about this joint spectral density function when there exists a well developed theory for multivariate random functions based on the covariance matrix in the time domain and the spectral density matrix in the frequency domain? The short answer to this question is that having the joint spectral density will make all the techniques for manipulating PDFs available to the spectral analysis of bivariate random sequences. In particular, random variable transformations, or the so-called algebra of random variables, are useful because they can be used to determine PDFs of functions of the frequency in two random sequences. If x and y are any two random variables, which may represent the frequency content in two random sequences, and $g(\cdot,\cdot)$ is a function, then the random variable

$$h = g(x, y),$$

can be expressed in terms of $g(\cdot,\cdot)$ and the joint probability density function of x and y [4].

The focus of this dissertation is to form a precise definition of the joint spectral density function in terms of the spectral representation of random sequences, explore its properties, and determine under what conditions (if any) it can be estimated from realizations of a bivariate random sequence.

1.2 The Radar Problem

An application where the joint spectral density could be used to great advantage is in the processing of radar backscatter return from distributed targets such as weather. A pulsed Doppler weather radar can be used to construct a map of the radial windspeed[1] over a scanned air volume (see Figure 1.1) [2]. Each of the boxed regions in Figure 1.1 represent a range cell volume from which samples of the backscatter return are collected over time (see Figure 1.2). These samples form a random sequence

[1]The component of the vector wind on a radial from the radar platform will be referred to as the radial windspeed.

in which the frequency content is related via the Doppler shift to the velocities of wind-blown particles within the range cell volume.

An important part of the signal processing task involves estimating the zeroth, first, and second moments of the Doppler spectrum in each range cell [5]. These three moments of the Doppler spectrum correspond to the signal power, mean velocity, and spectrum width respectively. When the goal of signal processing is the detection of certain types of weather events (e.g., gust fronts, tornadoes, windshears), the radial windspeed gradient, $\frac{\partial v}{\partial r}$, is often of particular interest. Knowledge of the joint spectral density of two adjacent range cells would make it possible to compute a probability density of windspeed gradient at each range increment through the use of random variable transformations.

An example of a function of the windspeed gradient that is of interest in the windshear detection problem, is a hazard index called the "F"–factor given by

$$F = \frac{V_g}{g}\frac{\partial v}{\partial r}\left[1 + \frac{2hg}{V_g V_a}\right],$$

where V_g is the aircraft groundspeed, V_a is the aircraft airspeed, h is the altitude, and g is the acceleration due to gravity [6]. The radial windspeed gradient can be estimated by

$$\frac{\partial v}{\partial r} = \frac{1}{\Delta r}(v_2 - v_1),$$

where v_1 and v_2 are the windspeeds in two range cells and Δr is the distance between them. If the joint probability density of v_1 and v_2 is known, then it is a simple matter to derive a probability density of hazard factor as the density of the difference of two random variables scaled by a constant [4]. The JSD provides a way to get that joint probability density of the windspeeds in two range cells.

1.3 A Note on Terminology

Much of the terminology associated with the spectral analysis of random sequences is used inconsistently in the various texts and papers available on the

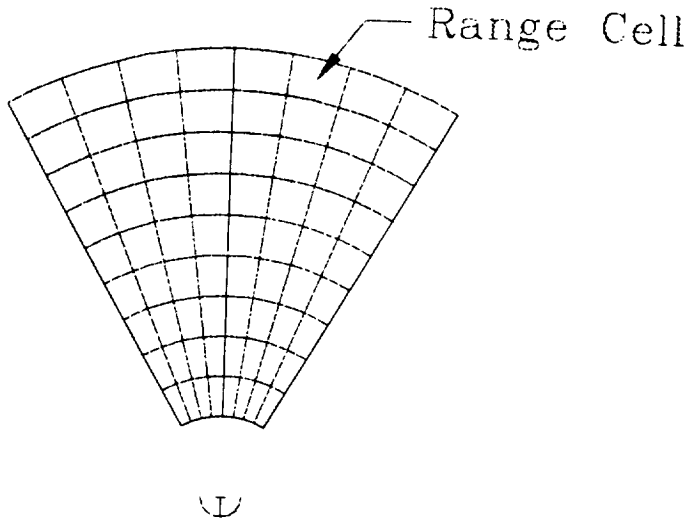

Figure 1.1. Scan region.

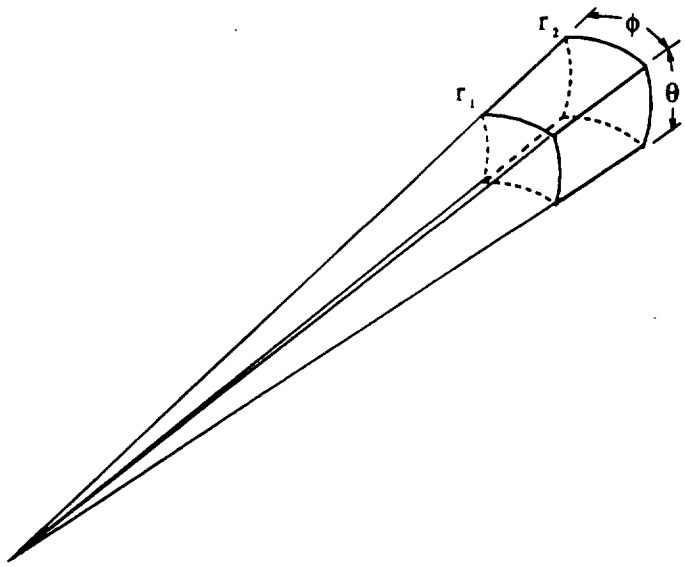

Figure 1.2. Range cell volume.

ject. In this section, ground rules are established for how various terms are used in this dissertation. First, there is no consistent way of distinguishing between random functions of a continuous parameter (sometimes time, but not restricted to be so) and random functions of a discrete parameter. In this regard, a convention following that of Yaglom is used so that random functions of a continuous parameter are called random processes and random functions of a discrete parameter are called random sequences [7]. When the argument may be either continuous or discrete it is called a random function. Throughout most of what follows, random sequences are treated exclusively, however, in most cases the results are equally applicable to random processes with only minor adjustments (usually involving a change of the limits of integration from $(-\pi, \pi]$ to $(-\infty, \infty]$).

When more that one random function is observed, they are collectively referred to as a multivariate random function. Single random functions of more than one argument are sometimes called multidimensional random functions, but here the term random field is preferred. The term random field makes no distinction between continuous and discrete arguments, but unless otherwise specified, the arguments are assumed to be discrete.

The exact meaning of margin as it is used in this dissertation depends on the type of function to which it is applied. For random fields and their correlation functions, the margin refers to the function that results from setting one or more of the arguments equal to zero. The margins of spectral distribution functions result from setting one or more of the arguments equal to π (or ∞ in the continuous case). Finally, for spectral density functions the margin is formed by integrating out one or more of the arguments.

1.4 Contribution to the Field

The contributions of this dissertation are twofold. First, a definition of the joint spectral density is given in terms of the spectral representation of random fields.

setting the problem on firm mathematical ground. According to this definition, the joint spectral density is not uniquely determined by the two sequences. Second, two classes of sequences are identified that enable the joint spectral density to be uniquely determined. The class of marginal sequences is considered first. Several theorems are offered on the properties of marginal sequences, and examples are provided to illustrate those properties. The second class of sequences is based on an exponential model that explicitly incorporates the joint spectral density in its definition. Simulations are used in both cases to corroborate the theory.

1.5 Organization

In the second chapter, the spectral and correlation theory of univariate random sequences is reviewed to establish basic concepts that are extended to bivariate random sequences and random fields in Chapter 3. Also in Chapter 3 a definition for the joint power spectral density is proposed based on establishing a connection between the spectral representation of a bivariate random sequence and the spectral representation of a random field. In Chapter 4 this connection is made by assuming a particular relationship between the bivariate random sequence and a stationary random field. Another approach to constraining the problem is explored in Chapter 5 where the bivariate random sequence is assumed to conform to a special model that allows the joint power spectral density to be computed under some special circumstances. Finally, in Chapter 6 some conclusions are drawn from this research and suggestions are made for future work.

CHAPTER 2

THEORY OF STATIONARY RANDOM SEQUENCES

In this chapter basic results in the correlation theory of stationary random sequences are reviewed. Sections 2.1 thru 2.3 provide some necessary background material from probability theory concerning functions that characterize random variables and, by extension, sequences of random variables. Section 2.4 is a mathematical aside defining the Stieltjes integral which occurs frequently in this and the following chapters. The remaining sections deal strictly with second order theory of stationary random sequences. The spectral representation is of particular interest because extensions of this theory are used in the next chapter to define the joint spectral density of a pair of random sequences. Many excellent references are available that cover this material, so well known theorems are stated without proof except where a brief sketch will provide insight to the theory without diverging too far from the flow of the discussion [8, 9, 10, 11, 1, 7, 12].

2.1 Distribution Functions

Consider a complex valued random sequence $\{X(t)\}, t = 0, \pm 1, \pm 2, \cdots \pm (n-1)$. Since $\{X(t)\}$ is complex valued, each element of the sequence can be written as

$$X(t) = U(t) + jV(t),$$

where $\{U(t)\}$ and $\{V(t)\}$ are real valued random variables. For finite n, $\{X(t)\}$ is completely characterized by its $2n$-dimensional probability distribution function,

$$F_{U(t_1),\ldots,U(t_n),V(t_1),\ldots,V(t_n)}(u_1,\ldots,u_n,v_1,\ldots,v_n) =$$

$$P(U(t_1) \le u_1,\ldots,U(t_n) \le u_n, V(t_1) \le v_1,\ldots,V(t_n) \le v_n). \tag{2.1}$$

where $(x_1, \ldots, x_n) \in \mathbb{C}^n$ and $x_i = u_i + jv_i$ [13]. For notational convenience, the distribution function for a complex valued $\{X(t)\}$ will be written as

$$F_{X(t_1),\ldots,X(t_n)}(x_1, \ldots, x_n), \tag{2.2}$$

by which is meant (2.1). When it is not necessary to distinguish whether a set of n random variables corresponds to n-length sequence of real valued random variables or a $\frac{n}{2}$-length sequence of complex valued random variables, the notation $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ will be used.

A function F is defined as a distribution function if

1. F is monotonically increasing in each of its arguments,
2. F is right continuous in each of its arguments,
3. $\lim_{x_t \to -\infty} F(x_1, \ldots, x_n) = 0 \quad (1 \leq t \leq n)$,
4. $F(+\infty, \ldots, +\infty) < \infty$,
5. $F(b_1, \ldots, b_n) - \sum_{i=1}^{n} p_i + \sum_{i<j} p_{ij} \mp \cdots + (-1)^n F(a_1, \ldots, a_n) \geq 0$,

for any a_i and b_i $(i = 1, \ldots, n)$, where $p_{ij\ldots k}$ is $F(c_1, \ldots, c_n)$ with $c_i = a_i, c_j = a_j, \ldots, c_k = a_k$ and $c_l = b_l$ for all $l \neq i, j, \ldots, k$ [9]. The left hand side of the condition in item (5) above is equal to the probability that $(\mathcal{X}_1, \ldots, \mathcal{X}_n)$ falls within the parallelepiped

$$a_i \leq \mathcal{X}_i < b_i,$$

that is,

$$P(a_1 \leq \mathcal{X}_1 < b_1, \ldots, a_n \leq \mathcal{X}_n < b_n).$$

If $F(+\infty, \ldots, +\infty) = 1$ then F is a probability distribution function.

The existence of a stochastic process corresponding to a particular *a priori* set of finite dimensional distribution functions,

$$F_{\mathcal{X}_1}(x_1), F_{\mathcal{X}_1,\mathcal{X}_2}(x_1, x_2), \ldots, F_{\mathcal{X}_1,\ldots,\mathcal{X}_n}(x_1, \ldots, x_n),$$

is assured by Kolmogorov's theorem if two conditions are met [14]. That is, a set of probability distribution functions are the distribution functions of some stochastic process if and only if the *symmetry* (or *permutation*) condition,

$$F_{X_{i_1},\ldots,X_{i_n}}(x_{i_1},\ldots,x_{i_n}) = F_{X_1,\ldots,X_n}(x_1,\ldots,x_n),$$

holds for all permutations $(i_1,\ldots,i_n)$ of $(1,\ldots,n)$, and the *compatibility* (or *consistency*) condition

$$F_{X_1,\ldots,X_m}(x_1,\ldots,x_m) = F_{X_1,\ldots,X_n}(x_1,\ldots,x_m,+\infty,\ldots,+\infty),$$

holds for $m < n$. The *compatibility* condition ensures that the lower dimensional distributions coincide with the margins of the n-dimensional distribution.

2.2 Density Functions

The derivative of the distribution function, if it exists, gives the density function as

$$f(x_1,\ldots,x_n) = \frac{dF(x_1,\ldots,x_n)}{dx_1\ldots dx_n}.$$

Explicit reference to the random variables $\{X_i\}$ has been dropped from the notation for simplicity when the meaning is clear from the context. A density function has the following properties [9]:

1. $f(x_1,\ldots,x_n) \geq 0$.
2. $P(a_1 \leq X_1 < b_1,\ldots,a_n \leq X_n < b_n) = \int_{a_n}^{b_n}\cdots\int_{a_1}^{b_1} f(x_1,\ldots,x_n)\,dx_1\ldots dx_n$.
3. $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1,\ldots,x_n)\,dx_1\ldots dx_n = 1$.

As a consequence of the compatibility condition

$$f(x_1,\ldots,x_m) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1,\ldots,x_n)\,dx_{m+1}\ldots dx_n,$$

so that the marginal density of $\{X_1,\ldots,X_m\}$ is obtained by integrating out the random variables $\{X_{m+1},\ldots,X_n\}$.

2.3 Characteristic Functions

The characteristic function of a random variable $\mathcal{X}$ is defined as

$$\Phi(\omega) = E[e^{j\omega x}] = \int_{-\infty}^{\infty} e^{j\omega x} dF(x), \tag{2.3}$$

where $F(x)$ is the distribution function of $\mathcal{X}$.

Theorem 2.1 (Bochner–Khinchin) *A function $\Phi(\omega_1, \ldots, \omega_n)$ with $\Phi(0, \ldots, 0) = 1$ is a characteristic function if and only if it is non-negative definite.*

In other words, the Bochner–Khinchin theorem says that the class of characteristic functions coincides with the class of non-negative definite functions, which by definition means that

$$\sum_{j=1}^{n} \sum_{k=1}^{n} \Phi(\omega_j - \omega_k) c_j c_k^* \geq 0, \tag{2.4}$$

for any integer n, where $\omega_1, \ldots, \omega_n$ are vectors in R^n and $c_1, \ldots, c_n$ are arbitrary complex numbers. Any function that satisfies (2.4) is a valid characteristic function for some random variable. Using (2.4) it is easy to show that non-negative definite functions have the following properties:

1. $\Phi(0, \ldots, 0) \geq 0$,
2. $\Phi(-\omega_1, \ldots, -\omega_n) = \Phi^*(\omega_1, \ldots, \omega_n)$,
3. $\Phi(0, \ldots, 0) \geq |\Phi(\omega_1, \ldots, \omega_n)|$.

A unique inverse relationship exists between a characteristic function $\Phi(\omega_1, \ldots, \omega_n)$ and its corresponding distribution function $F(x_1, \ldots, x_n)$ given by [9],

$$P(a_1 \leq \mathcal{X}_1 < b_1, \ldots, a_n \leq \mathcal{X}_n < b_n) =$$

$$\lim_{T \to \infty} \frac{1}{(2\pi)^n} \int_{-T}^{T} \cdots \int_{-T}^{T} \prod_{k=1}^{n} \frac{e^{-j\omega_k a_k} - e^{-j\omega_k b_k}}{j\omega_k} \Phi(\omega_1, \ldots, \omega_n)\, d\omega_1 \ldots d\omega_n.$$

2.4 Stieltjes Integral

The integral on the right-hand side of (2.3) is the so-called Stieltjes integral which will be used extensively in what follows [9]. The Stieltjes integral of a function $g(x)$ with respect to the distribution function $F(x)$ is defined as

$$\int_a^b g(x)\,dF(x) = \lim_{\max|x_k - x_{k-1}|\to 0} \sum_{k=1}^{N} g(\check{x}_k)[F(x_k) - F(x_{k-1})]$$

where

$$a = x_0 < x_1 < \cdots < x_N = b, \text{ and } x_{k-1} \le \check{x}_k \le x_k.$$

This generalization of the integral remains well defined even when $F(x)$ is not everywhere differentiable. The improper Stieltjes integral where the interval of integration goes from $-\infty$ to $+\infty$ is defined in the usual way as

$$\lim_{\substack{a\to-\infty \\ b\to+\infty}} \int_a^b g(x)\,dF(x).$$

If $F(x)$ is differentiable with respect to x, then the Stieltjes integral reduces to the ordinary integral

$$\int_a^b g(x)f(x)\,dx$$

where $f(x) = \frac{dF(x)}{dx}$. The Stieltjes integral can be applied to either Riemann or Lebesgue integration in which case it is called the Riemann-Stieltjes or Lebesgue-Stieltjes integral respectively.

2.5 Moments

It is often more convenient to work with a partial characterization of a random sequence in terms of a finite set of statistical moments. The first moment is simply the mean and is defined as

$$\mu(t) = E[X(t)] = \int_{-\infty}^{\infty} x\,dF_{X(t)}(x).$$

The second moment is defined as

$$R(t_1, t_2) = E[X^*(t_1)X(t_2)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 \, dF_{X(t_1),X(t_2)}(x_1, x_2), \qquad (2.5)$$

and is called the autocorrelation function. Alternatively, the second central moment can be formed by taking the autocorrelation of $\{X(t) - \mu(t)\}$ in which case it is called the autocovariance,

$$C(t_1, t_2) = R(t_1, t_2) - E[\mu(t_1)]E[\mu(t_2)].$$

The convention of applying the term autocorrelation to (2.5) is standard among the engineering community, however; mathematicians use the term autocorrelation to signify the normalized autocovariance. Henceforth, use of the term autocorrelation always refers to (2.5), and the mean of $\{X(t)\}$ is assumed to be zero so that the autocorrelation and autocovariance are identical. This assumption results in no loss of generality.

2.6 Stationarity

A random sequence, $\{X(t)\}$, is strict-sense stationary if its distribution functions are independent of shifts in the index t

$$F_{X(t_1),\ldots,X(t_m)}(x_1, \ldots, x_m) = F_{X(t_{1+i}),\ldots,X(t_{m+i})}(x_1, \ldots, x_m),$$

for any integer i [4]. If the joint distribution of two sequences $\{X(t), Y(t)\}$ is stationary then the sequences are said to be jointly (or mutually) stationary. Clearly from (2.1) a complex sequence is stationary if its real and imaginary parts are jointly stationary.

A less restrictive form of stationarity, called wide-sense stationarity, only requires that the mean be constant

$$\mu(t) = E[X(t)] = \mu,$$

and the correlation function depend only on the difference of the indexes

$$R(t_1, t_2) = R(\tau) = E[X^*(t_1)X(t_1 + \tau)],$$

where $\tau = t_2 - t_1$. Consequently, the autocorrelation is independent of the absolute starting point, t_1. Henceforth, when a random sequence is called stationary it means stationary in the wide-sense. Since Gaussian random sequences are fully described in terms of their first two moments, in this special case wide-sense stationarity implies strict-sense stationarity.

2.7 Properties of Autocorrelation Functions

Theorem 2.2 *A function $R(\tau)$ defined on integers is the autocorrelation function of a zero-mean, stationary random sequence if and only if it is non-negative definite. That is,*

$$\sum_{j=1}^{n}\sum_{k=1}^{n} R(j-k)c_j c_k^* \geq 0,$$

for any positive integer n and arbitrary complex numbers $c_1, \ldots c_n$ [12].

As a consequence of being a non-negative definite function, the autocorrelation of a complex valued random sequence has the following properties:

1. $R(0) \geq 0$,
2. $|R(\tau)| \leq R(0)$,
3. $R(\tau) = R^*(-\tau)$.

From the first and second property, the autocorrelation must have its maximum value at zero lag and be non-negative at that point. The third property indicates that $R(\cdot)$ is a Hermitian function.

Comparing the properties of autocorrelation functions with the properties of characteristic functions from Section 2.3, it is apparent that an close relationship exists between the two. A variation of Herglotz' Theorem (Theorem 2.3 in the next

section) states that a function, $R(\tau)$, defined on integers is non-negative definite, and thus an autocorrelation function, if and only if it coincides on the set of integers with a characteristic function, $\Phi(\omega)$, defined by

$$\Phi(\omega) = \int_{-\pi}^{\pi} e^{j\omega x} dF(x),$$

where ω is continuous on R^1 [15].

2.8 Spectral Representation of Correlation Functions

Theorem 2.3 (Herglotz) *A complex valued function $R(\tau)$ is non-negative definite if and only if it has a representation*

$$R(\tau) = \int_{-\pi}^{\pi} e^{j\lambda m} dF(\lambda), \tag{2.6}$$

where $F(\lambda)$ has the properties of a distribution function on $(-\pi, \pi]$.

$F(\lambda)$ is called the spectral distribution function and as stated in the theorem it has all the usual properties of a distribution function except that it is defined on $(-\pi, \pi]$, such that

$$F(\lambda) = 0, \qquad \lambda \leq -\pi,$$

$$F(\lambda) = F(\pi), \qquad \lambda \geq \pi.$$

If a correlation function is absolutely summable,

$$\sum_{\tau=-\infty}^{\infty} |R(\tau)|^2 < \infty,$$

then $R(\tau)$ can be represented in terms of it Fourier coefficients

$$R(\tau) = \int_{-\pi}^{\pi} e^{j\lambda\tau} f(\lambda) d\lambda,$$

where

$$f(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} e^{-j\lambda\tau} R(\tau),$$

is the power spectral density which has all the properties of a density function defined on $(-\pi, \pi]$.

2.9 Spectral Representation of Stationary Random Sequences

Theorem 2.4 (Cramér) *Any zero mean stationary random sequence can be represented in the form*

$$X(t) = \int_{-\pi}^{\pi} e^{j\lambda t} dZ(\lambda), \tag{2.7}$$

where $\{Z(\lambda)\}$ *is a complex valued random process with orthogonal increments defined on* $(-\pi, \pi]$.

A proof of this theorem due to Cramér is given below in a non-rigorous form to help illustrate the essential properties of the spectral representation [8, 10, 1, 16, 17]. Thorough treatments of the Hilbert space theory necessary to understand what follows are available in the literature [12, 17].

Let H be the Hilbert space formed by the collection of all complex valued random variables with zero mean and finite variance. Then for each value of t, the random variable $X(t)$ belongs to H. Let H_x denote the closed linear subspace of H that is spanned by the random sequence $\{X(t)\}$. The inner product of two elements of H_x is defined as

$$\langle x_1, x_2 \rangle = E[x_1^* x_2], \tag{2.8}$$

and the distance between two elements is defined as

$$d(x_1, x_2) = \|x_1 - x_2\| = \sqrt{E\left[|x_1 - x_2|^2\right]}. \tag{2.9}$$

Denote by $L_2(F)$ the set of all complex valued functions, $\Theta(\lambda)$, on the interval $(-\pi, \pi]$ for which the Lebesgue-Stieltjes integral

$$\int_{-\pi}^{\pi} |\Theta(\lambda)|^2 dF(\lambda),$$

exists and is finite. Here F is the spectral distribution function of $X(t)$. Then $L_2(F)$ forms a Hilbert space H_θ with the inner product of two elements θ_1 and θ_2 defined as

$$\langle \theta_1, \theta_2 \rangle = \int_{-\pi}^{\pi} \theta_1^*(\lambda)\theta_2(\lambda) dF(\lambda),$$

and the distance between θ_1 and θ_2 defined as

$$d(\theta_1, \theta_2) = \|\theta_1 - \theta_2\| = \left[\int_{-\pi}^{\pi} |\theta_1(\lambda) - \theta_2(\lambda)|^2 \, dF(\lambda)\right]^{\frac{1}{2}}.$$

A linear, one-to-one mapping between two spaces that preserves inner products is known as a congruence. A congruence between H_x and H_θ can be established as follows. Let $x(t) \in H_x$ and $\theta_t(\lambda) = e^{j\lambda t} \in H_\theta$ be corresponding elements in the two spaces. By Herglotz' Theorem,

$$R_{xx}(t, u) = E[x^*(t)x(u)] = \int_{-\pi}^{\pi} e^{-j\lambda t} e^{-j\lambda u} dF(\lambda).$$

Comparing the spectral representation of the autocorrelation with the definition of the inner product in H_x and H_θ it is evident that the inner product is preserved for a mapping M such that

$$M\left[e^{j\lambda t}\right] = x(t),$$

and for linear combinations of $\{\theta_t(\lambda)\}$

$$M\left[\sum_i c_i \theta_{t_i}(\lambda)\right] = \sum_i c_i M[\theta_{t_i}(\lambda)].$$

Now let $\mathbf{1}_{(\lambda_1,\lambda_2]}(\lambda)$ be the indicator function defined such that

$$\mathbf{1}_{(\lambda_1,\lambda_2]}(\lambda) = \begin{cases} 1 & \lambda_1 < \lambda \leq \lambda_2 \\ 0 & \text{otherwise.} \end{cases}$$

Simple functions in H_θ can be written in terms of the indicator function in the following manner:

$$\theta_t(\lambda) = \lim_{n\to\infty} \sum_{i=1}^{n} \theta_t(\check{\lambda}) \mathbf{1}_{(\lambda_{i-1},\lambda_i]}(\lambda), \tag{2.10}$$

where $\lambda_{i-1} \leq \check{\lambda} \leq \lambda_i$ and $-\pi = \lambda_0 < \lambda_1 < \cdots < \lambda_n = \pi$. The indicator function itself is an element of H_θ. Define $Z(\lambda)$ as the random process in H_x corresponding to the indicator function in H_θ through the mapping M as follows:

$$Z(\lambda_2) - Z(\lambda_1) = M[\mathbf{1}_{(\lambda_1,\lambda_2]}(\lambda)].$$

Clearly, $Z(\lambda)$ is orthogonal over disjoint increments since for $[\lambda_1, \lambda_2) \cap [\lambda_3, \lambda_4) = 0$,

$$E\left[\{Z(\lambda_4) - Z(\lambda_3)\}^*\{Z(\lambda_2) - Z(\lambda_1)\}\right] = \langle \mathbf{1}_{(\lambda_3,\lambda_4]}(\lambda), \mathbf{1}_{(\lambda_1,\lambda_2]}(\lambda)\rangle$$
$$= 0.$$

If inner product is preserved by the mapping M, that implies that distance is also preserved so that,

$$E\left[|Z(\lambda_2) - Z(\lambda_1)|^2\right] = \|\mathbf{1}_{(\lambda_2,\lambda_1]}(\lambda)\|^2$$
$$= \int_{\lambda_1}^{\lambda_2} dF(\lambda)$$
$$= F(\lambda_2) - F(\lambda_1).$$

Applying the mapping M to (2.10) gives

$$M[\theta_t(\lambda)] = \lim_{n\to\infty} \sum_{i=0}^{n} \theta_t(\check{\lambda}) M[\mathbf{1}_{(\lambda_{i-1},\lambda_i]}(\lambda)] = \lim_{n\to\infty} \sum_{i=0}^{n} \theta_t(\check{\lambda})\left[Z(\lambda_{i-1}) - Z(\lambda_i)\right]$$

which converges to the Stieltjes integral

$$x(t) = \int_{-\pi}^{\pi} \theta_t(\lambda) dZ(\lambda).$$

Substituting $\theta_t(\lambda) = e^{j\lambda t}$ gives the spectral representation in (2.7). The integral on the right-hand side of (2.7) is a stochastic Stieltjes integral with respect to a random measure, in this case $Z(\lambda)$.

2.10 Orthogonal Increment Process

An orthogonal increment process, $Z(\lambda)$, has the following properties:

1. $E[Z(\lambda)] = 0$,
2. $Var[Z(\lambda)] \leq \infty$,
3. $Cov[Z(\lambda_4) - Z(\lambda_3), Z(\lambda_2) - Z(\lambda_1)] = 0$, (2.11)

where $(\lambda_1, \lambda_2] \cap (\lambda_3, \lambda_4] = 0$ and $Cov[X, Y]$ and $Var[X]$ are defined as

$$Cov[X, Y] = E[X^*Y],$$
$$Var[X] = Cov[X, X].$$

The orthogonal increment process, $Z(\lambda)$, is related to the spectral distribution function in (2.6), $F(\lambda)$, by

$$F(\lambda_2) - F(\lambda_1) = Var[Z(\lambda_2) - Z(\lambda_1)]. \tag{2.12}$$

Recall that $F(-\pi) = 0$, so letting $\lambda_1 = -\pi$ causes (2.12) to reduce to

$$F(\lambda) = Var[Z(\lambda) - Z(-\pi)],$$

where $Z(-\pi)$ is the starting value of the orthogonal increment process which can also be set to zero without affecting the values of the increments [12]. Considering (2.11) and (2.12), the relationship between $Z(\lambda)$ and $F(\lambda)$ can be expressed as

$$E[dZ^*(\lambda)dZ(\check{\lambda})] = \delta(\lambda - \check{\lambda})dF(\lambda).$$

The inversion formula for $Z(\lambda)$ is given by

$$Z(\lambda) = \lim_{n\to\infty} \frac{1}{2\pi} \sum_{t=-n}^{n} \frac{e^{-j\lambda t} - (1)^t}{-jt} X(t) + Z(-\pi).$$

It can be shown that $X(t)$ is Gaussian if and only if $Z(\lambda) - Z(-\pi)$ is Gaussian [8].

2.11 Brownian Motion

One dimensional Brownian motion [12], $\{B(\lambda),\ -\pi \leq \lambda \leq \pi\}$, is a simple example of an orthogonal increment process which is Gaussian distributed with

$$E[B(\lambda)] = 0$$

$$Var[B(\lambda)] = \frac{\sigma^2(\lambda + \pi)}{2\pi}.$$

The spectral distribution function corresponding to a Brownian motion process specified as above is given by

$$F(\lambda) = \begin{cases} 0 & \lambda \leq -\pi, \\ \frac{\sigma^2(\lambda+\pi)}{2\pi} & -\pi \leq \lambda \leq \pi, \\ \sigma^2 & \lambda \geq \pi. \end{cases}$$

Substituting $F(\lambda)$ into (2.6) gives the autocorrelation as

$$\begin{aligned} R(\tau) &= \int_{-\pi}^{\pi} e^{j\lambda\tau}\, dF(\lambda), \\ &= \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} e^{j\lambda\tau}\, d\lambda, \\ &= \sigma^2 \delta(h). \end{aligned}$$

As expected, the autocorrelation of a random sequence with a Brownian orthogonal increments process is an impulse at zero lag because the spectral density is uniform on $(-\pi, \pi]$.

CHAPTER 3

JOINT SPECTRAL DENSITY

In Chapter 2, a univariate random sequence was shown to have a spectral representation in terms of an orthogonal increments process. Furthermore, taking the variance of this orthogonal increments process gives the spectral distribution function which acts like a probability distribution function of the frequency content of the random sequence. At times it is more convenient to work with the distribution function rather than the density function. This should cause no difficulty since the density function can be found by taking the derivative of the distribution function (assuming that it is differentiable).

The purpose of this chapter is to extend the idea of a PDF of frequency content to the case of two random sequences. Clearly if the joint spectral density is to be completely analogous to a joint PDF, it must be a real valued function of multiple arguments (in this case 2). In Section 3.1 classical spectral analysis of bivariate random sequences is discussed in terms of the correlation matrix and the spectral density matrix. Since these are matrix valued functions of a single argument, they cannot be considered candidates for the JSD. And yet, they contain all the available information about the joint probability structure of the two sequences. In Section 3.2 the spectral representation of stationary random fields is discussed and it is noted that the spectral density function of a stationary random field has the same form and properties that are desired for the JSD. The properties of two-dimensional orthogonal increments processes are explored in Section 3.3. The joint spectral density is formally introduced in Section 3.4 by treating the autospectra of the bivariate random sequence as the marginal spectra of a stationary random field.

3.1 Bivariate Correlation and Spectral Properties

The correlation theory of bivariate random sequences is well documented in the literature [1, 12, 7, 19]. Consider two jointly stationary random sequences $\{X_1(t)\}, \{X_2(t)\}$ that have the following spectral representations:

$$X_1(t) = \int_{-\pi}^{\pi} e^{j\lambda t} dZ_1(\lambda), \tag{3.1}$$

$$X_2(t) = \int_{-\pi}^{\pi} e^{j\lambda t} dZ_2(\lambda), \tag{3.2}$$

where $\{Z_1(\lambda)\}, \{Z_2(\lambda)\}$ are orthogonal increment processes. Then assuming that both process have zero mean (i.e., $E[X_1(t)] = E[X_2(t)] = 0$), the autocorrelation functions can be written as follows:

$$R_{11}(\tau) = E[X_1^*(t)X_1(t+\tau)] = \int_{-\pi}^{\pi} e^{j\lambda\tau} dF_{11}(\lambda), \tag{3.3}$$

$$R_{22}(\tau) = E[X_2^*(t)X_2(t+\tau)] = \int_{-\pi}^{\pi} e^{j\lambda\tau} dF_{22}(\lambda), \tag{3.4}$$

where $F_{11}(\lambda)$ and $F_{22}(\lambda)$ are the spectral distributions of the two processes. Recall from Section 2.9 that

$$Var[dZ_1(\lambda)] = dF_{11}(\lambda) \tag{3.5}$$

$$Var[dZ_2(\lambda)] = dF_{22}(\lambda). \tag{3.6}$$

A facet of the analysis of multivariate processes that is absent when considering the univariate case is the interaction between pairs of processes. This information is captured by the cross-correlation. The cross-correlation is given by

$$R_{12}(t,u) = E[X_1^*(t)X_2(u)]. \tag{3.7}$$

Substituting (3.1) and (3.2) into (3.7) gives

$$R_{12}(t,u) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{-j\lambda t} e^{j\check{\lambda}u} E[dZ_1^*(\lambda)dZ_2(\check{\lambda})]. \tag{3.8}$$

If $\{X_1(t)\}, \{X_2(t)\}$ are jointly 2^{nd}-order stationary then their cross-correlation will be a function only of the shift index $\tau = u - t$ and will be independent of the starting index t. For this to be the case in (3.8), the following condition must hold:

$$E[dZ_1^*(\lambda)dZ_2(\check{\lambda})] = 0 \quad \lambda \neq \check{\lambda}, \tag{3.9}$$

and therefore,

$$R_{12}(\tau) = \int_{-\pi}^{\pi} e^{j\lambda\tau} E[dZ_1^*(\lambda)dZ_2(\lambda)] = \int_{-\pi}^{\pi} e^{j\lambda\tau} dF_{12}(\lambda), \tag{3.10}$$

where $F_{12}(\lambda)$ is the cross-spectral distribution function. The cross-correlation function is not necessarily non-negative definite, so it need not share the same properties as the autocorrelation. One significant difference is that the cross-correlation may have its maximum value away from zero lag. As a result, $F_{12}(\lambda)$ is in general complex valued and even if it is real valued is not necessarily non-decreasing. It is however a function of bounded variation because its real and imaginary parts are the difference of two non-decreasing functions respectively [7].

For bivariate random sequences, the auto- and cross-correlations are often assembled into a correlation matrix given by

$$\mathbf{R}(\tau) = \begin{bmatrix} R_{11}(\tau) & R_{12}(\tau) \\ R_{21}(\tau) & R_{22}(\tau) \end{bmatrix}. \tag{3.11}$$

The two cross-correlation elements of the correlation matrix yield redundant information since

$$R_{12}(\tau) = R_{21}^*(-\tau). \tag{3.12}$$

If the elements of $\mathbf{R}(\tau)$ are absolutely summable,

$$\sum_{\tau=-\infty}^{\infty} |R_{ij}(\tau)|^2 < \infty, \quad i,j = 1,2, \tag{3.13}$$

then spectral density matrix is given by

$$\mathbf{f}(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} e^{-j\lambda\tau} \mathbf{R}(\tau) \tag{3.14}$$

where

$$\mathbf{f}(\lambda) = \begin{bmatrix} f_{11}(\lambda) & f_{12}(\lambda) \\ f_{21}(\lambda) & f_{22}(\lambda) \end{bmatrix} \tag{3.15}$$

and

$$f_{ij}(\lambda) = \frac{dF_{ij}(\lambda)}{d\lambda}. \tag{3.16}$$

Properties of the elements of the spectral density matrix include

$$f_{11}(\lambda) \geq 0,$$

$$f_{22}(\lambda) \geq 0,$$

$$|f_{12}|^2 \leq f_{11}(\lambda) f_{22}(\lambda),$$

where the third relation is a consequence of Schwartz' inequality. The cross-spectral density, f_{12}, is also complex valued in general, so it can be written in terms of pairs of real valued functions in a couple of different ways. Each of these ways of writing f_{12} offers a particular interpretive viewpoint. Breaking f_{12} down into its real and imaginary parts gives the co-spectrum (real part) and quadrature spectrum (imaginary part) as

$$f_{12}(\lambda) = c_{12}(\lambda) - jq_{12}(\lambda).$$

Writing f_{12} in polar form gives the cross-amplitude spectrum and phase spectrum as

$$f_{12}(\lambda) = \alpha_{12}(\lambda) e^{j\phi_{12}(\lambda)},$$

where

$$\alpha_{12}(\lambda) = |f_{12}(\lambda)| = \sqrt{c_{12}^2(\lambda) + q_{12}^2(\lambda)},$$

$$\phi_{12}(\lambda) = \arg(f_{12}(\lambda)) = \arctan\left(\frac{-q_{12}(\lambda)}{c_{12}(\lambda)}\right).$$

The properties and interpretation of the co-spectrum, quadrature spectrum, cross-amplitude spectrum, and the phase spectrum are treated in depth in the literature [1, 7, 19].

The complex coherence is defined as

$$\gamma_{12}(\lambda) = \frac{f_{12}(\lambda)}{\sqrt{f_{11}(\lambda) f_{22}(\lambda)}}.$$

The complex coherence is sometimes referred to as the correlation coefficient in the frequency domain, but it should not be confused with a correlation coefficient between the frequency content in the two random sequences. It is, rather, "the correlation coefficient between the random coefficients of the components" in the two random sequences at each frequency [1].

The correlation matrix and the spectral density matrix describe the joint probability structure of the bivariate random sequence, however; if the auto-spectra are interpreted as the probability densities of the frequency in each random sequence, then it might be possible to reorganize the information available in these matrix valued functions to form a joint probability density function of the frequency content in the two sequences. In the following two sections this problem is set within the framework of the spectral representation of stationary random fields and their marginal spectra.

3.2 Stationary Random Fields

It is well known that the spectral representation theorem for random sequences can also be extended to stationary random fields [20, 7, 1]. A random field is called stationary if the mean is constant

$$\mu(t,u) = E[X(t,u)] = \mu, \tag{3.17}$$

and the correlation function is function only of shifts in its arguments

$$R(t,u,v,w) = R(\tau,\upsilon) = E[X^*(t,u)X(v,w)], \tag{3.18}$$

where $\tau = v - t$ and $v = w - u$. This two-dimensional autocorrelation function has a spectral representation

$$R(\tau, v) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j(\lambda_1 \tau + \lambda_2 v)} dF(\lambda_1, \lambda_2), \tag{3.19}$$

where $F(\lambda_1, \lambda_2)$ has the properties of a distribution function on $(-\pi, \pi] \times (-\pi, \pi]$. If $F(\lambda_1, \lambda_2)$ is differentiable, then the random field has a spectral density, $f(\lambda_1, \lambda_2)$, given by

$$f(\lambda_1, \lambda_2) = \frac{dF(\lambda_1, \lambda_2)}{d\lambda_1 d\lambda_2},$$

where $f(\lambda_1, \lambda_2)$ has the properties of a density function on $(-\pi, \pi] \times (-\pi, \pi]$.

Theorem 3.1 *Any zero mean stationary random field can be represented in the form*

$$X(t, u) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j(\lambda_1 t + \lambda_2 u)} dZ(\lambda_1, \lambda_2), \tag{3.20}$$

where $\{Z(\lambda_1, \lambda_2)\}$ *is an orthogonal increment process on* $(-\pi, \pi] \times (-\pi, \pi]$ *[20, 7].*

The proof of this theorem given below is the author's extension of Cramér's proof of Theorem 2.7 to random fields.

Let H_x be the closed linear subspace of H that is spanned by the random field $\{X(t, u)\}$ (see Section 2.9). The inner product and distance for H_x are defined as (2.8) and (2.9) respectively.

Denote by $L_2(G)$ the set of all complex valued functions, $\Psi(\lambda_1, \lambda_2)$, on $(-\pi, \pi] \times (-\pi, \pi]$ for which the two-dimensional Lebesgue-Stieltjes integral

$$\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |\Psi(\lambda_1, \lambda_2)|^2 dG(\lambda_1, \lambda_2), \tag{3.21}$$

exists and is finite. Here G is the two-dimensional spectral distribution function of $X(t, u)$. Then $L_2(G)$ forms a Hilbert space H_ψ with the inner product of elements ψ_1 and ψ_2 defined as

$$\langle \psi_1, \psi_2 \rangle = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \psi_1^*(\lambda_1, \lambda_2) \psi_2(\lambda_1, \lambda_2) dG(\lambda_1, \lambda_2). \tag{3.22}$$

To construct a congruence between H_x and H_ψ, let $X(t,u) \in H_x$ and $\psi_{tu}(\lambda_1,\lambda_2) = e^{j(\lambda_1 t+\lambda_2 u)} \in H_\psi$ be corresponding elements in the two Hilbert spaces. By Herglotz theorem in two dimensions,

$$\begin{aligned} &R_{xx}(t,u,v,w) = \\ &E\left[x^*(t,u)x(v,w)\right] = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{-j(\lambda_1 t+\lambda_2 u)} e^{j(\lambda_1 v+\lambda_2 w)} dG(\lambda_1,\lambda_2) \\ &\qquad = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \psi^*_{tu}(\lambda_1,\lambda_2)\psi_{vw}(\lambda_1,\lambda_2) dG(\lambda_1,\lambda_2) \\ &\langle x(t,u), x^*(v,w)\rangle = \langle \psi_{tu}, \psi_{vw}\rangle. \end{aligned}$$

Then the inner product is preserved for the mapping T such that

$$T\left[e^{j(\lambda_1 t+\lambda_2 u)}\right] = X(t,u).$$

Let $\mathbf{1}_{(\mu_i,\mu_j]\times(\nu_i,\nu_j]}(\lambda_1,\lambda_2)$ be the indicator function defined on rectangles such that

$$\mathbf{1}_{(\mu_i,\mu_j]\times(\nu_i,\nu_j]}(\lambda_1,\lambda_2) = \begin{cases} 1 & \mu_i < \lambda_1 \le \mu_j, \quad \nu_i < \lambda_2 \le \nu_j \\ 0 & \text{otherwise.} \end{cases}$$

Simple functions can be constructed from this two-dimensional version of the indicator function as follows:

$$\psi_{st}(\lambda_1,\lambda_2) = \lim_{n\to\infty} \sum_{i,j=1}^{n} \psi_{st}(\check{\lambda}_1,\check{\lambda}_2)\mathbf{1}_{(\mu_{i-1},\mu_i]\times(\nu_{j-1},\nu_j]}(\lambda_1,\lambda_2)$$

where

$$\begin{aligned} \mu_{i-1} < &\check{\lambda}_1 \le \mu_i, \\ \nu_{i-1} < &\check{\lambda}_2 \le \nu_i, \\ -\pi = \lambda_0 < &\cdots < \lambda_n = \pi, \\ -\pi = \nu_0 < &\cdots < \nu_n = \pi. \end{aligned}$$

The indicator function is also an element of H_ψ, so we can define its corresponding random process, $Z(\lambda_1,\lambda_2)$, in H_x as

$$Z(\mu_2,\nu_2) - Z(\mu_2,\nu_1) - Z(\mu_1,\nu_2) + Z(\mu_1,\nu_1) = T\left[\mathbf{1}_{(\mu_1,\mu_2]\times(\nu_1,\nu_2]}(\lambda_1,\lambda_2)\right]. \quad (3.2$$

It can be shown that $Z(\lambda_1, \lambda_2)$ has orthogonal increments since for non-intersecting rectangles, the inner product of the indicator function is equal to zero. Since distance is preserved by the mapping T,

$$E\left[\left|Z(\mu_2,\nu_2) - Z(\mu_2,\nu_1) - Z(\mu_1,\nu_2) + Z(\mu_1,\nu_1)\right|^2\right]$$

$$= \|\mathbf{1}_{(\mu_1,\mu_2]\times(\nu_1,\nu_2]}(\lambda_1,\lambda_2)\|^2$$

$$= \int_{\nu_1}^{\nu_2}\int_{\mu_1}^{\mu_2} dG(\lambda_1,\lambda_2)$$

$$= G(\mu_2,\nu_2) - G(\mu_2,\nu_1) - G(\mu_1,\nu_2) + G(\mu_1,\nu_1).$$

Applying the mapping T to (3.23) gives

$$T\left[\psi_{st}(\lambda_1,\lambda_2)\right] = \lim_{n\to\infty}\sum_{i,j=1}^{n}\psi_{st}(\check{\lambda}_1,\check{\lambda}_2)T\left[\mathbf{1}_{(\mu_{i-1},\mu_i]\times(\nu_{j-1},\nu_j]}(\lambda_1,\lambda_2)\right],$$

$$= \lim_{n\to\infty}\sum_{i,j=1}^{n}\psi_{st}(\check{\lambda}_1,\check{\lambda}_2)\left[Z(\mu_i,\nu_j) - Z(\mu_i,\nu_{j-1}) - Z(\mu_{i-1},\nu_j) + Z(\mu_{i-1},\nu_{j-1})\right],$$

which converges to the two-dimensional Stieltjes integral

$$X(t,u) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\psi_{tu}(\lambda_1,\lambda_2)dZ(\lambda_1,\lambda_2).$$

Letting $\psi_{tu}(\lambda_1,\lambda_2) = e^{j(\lambda_1 t+\lambda_2 u)}$ gives the spectral representation of a two-dimensional random field

$$X(t,u) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}e^{j(\lambda_1 t+\lambda_2 u)}dZ(\lambda_1,\lambda_2). \tag{3.24}$$

3.3 Properties of $Z(\lambda_1, \lambda_2)$

An orthogonal increment process in two dimensions has properties similar to the one dimensional case.

1. $E[Z(\lambda_1,\lambda_2)] = 0$,
2. $Var[Z(\lambda_1,\lambda_2)] \leq \infty$,
3. $Cov\left[\Delta Z\left\{(\mu_3,\mu_4]\times(\nu_3,\nu_4]\right\}, \Delta Z\left\{(\mu_1,\mu_2]\times(\nu_1,\nu_2]\right\}\right] = 0$,

where $(\mu_1, \mu_2] \times (\nu_1, \nu_2] \cap (\mu_3, \mu_4] \times (\nu_3, \nu_4] = 0$. The increment operator Δ in two dimensions is defined as follows:

$$\Delta Z\{(\mu_1, \mu_2] \times (\nu_1, \nu_2]\} = Z(\mu_2, \nu_2) - Z(\mu_1, \nu_2) - Z(\mu_2, \nu_1) + Z(\mu_1, \nu_1).$$

The random measure $\Delta Z\{(\mu_1, \mu_2] \times (\nu_1, \nu_2]\}$ is related to the spectral measure by

$$\Delta F\{(\mu_1, \mu_2] \times (\nu_1, \nu_2]\} = Var[\Delta Z\{(\mu_1, \mu_2] \times (\nu_1, \nu_2]\}].$$

To see this relationship more clearly, consider computing

$$Cov\left[\Delta Z\left\{(\mu_3, \mu_4] \times (\nu_3, \nu_4]\right\}, \Delta Z\left\{(\mu_1, \mu_2] \times (\nu_1, \nu_2]\right\}\right],$$

where $(\mu_1, \mu_2] \times (\nu_1, \nu_2] \cap (\mu_3, \mu_4] \times (\nu_3, \nu_4] \neq 0$. This situation is depicted in Figure 3.1. The only non-zero contribution to the covariance is from the region where the rectangular increments overlap (the shaded region in the figure). Therefore, the covariance reduces to

$$Cov\left[\Delta Z\left\{(\mu_3, \mu_4] \times (\nu_3, \nu_4]\right\}, \Delta Z\left\{(\mu_1, \mu_2] \times (\nu_1, \nu_2]\right\}\right] = Var\left[\Delta Z\left\{(\mu_3, \mu_2] \times (\nu_1, \nu_4]\right\}\right].$$

As the rectangular increments are shrunk down to infinitesimals, the rectangles will only overlap if their vertices coincide so that we are left with the short-hand notation,

$$Cov[dZ(\lambda_1, \lambda_2), dZ(\tilde{\lambda_1}, \tilde{\lambda_2})] = \delta(\lambda_1 - \tilde{\lambda_1})\delta(\lambda_2 - \tilde{\lambda_2})dF(\lambda_1, \lambda_2).$$

3.4 Joint Spectral Density

Definition 3.1 *The joint spectral distribution, $F(\lambda_1, \lambda_2)$, of two random sequences, $X_1(t)$ and $X_2(t)$ with corresponding orthogonal increment process $Z_1(\lambda)$ and $Z_2(\lambda)$ is equal to the spectral distribution function of some stationary two-dimensional random field such that*

$$F(\lambda, \pi) = Var[Z_1(\lambda)],$$

$$F(\pi, \lambda) = Var[Z_2(\lambda)],$$

and is consistent with the $Cov[Z_1(\lambda_1), Z_2(\lambda_2)]$ away from the margins.

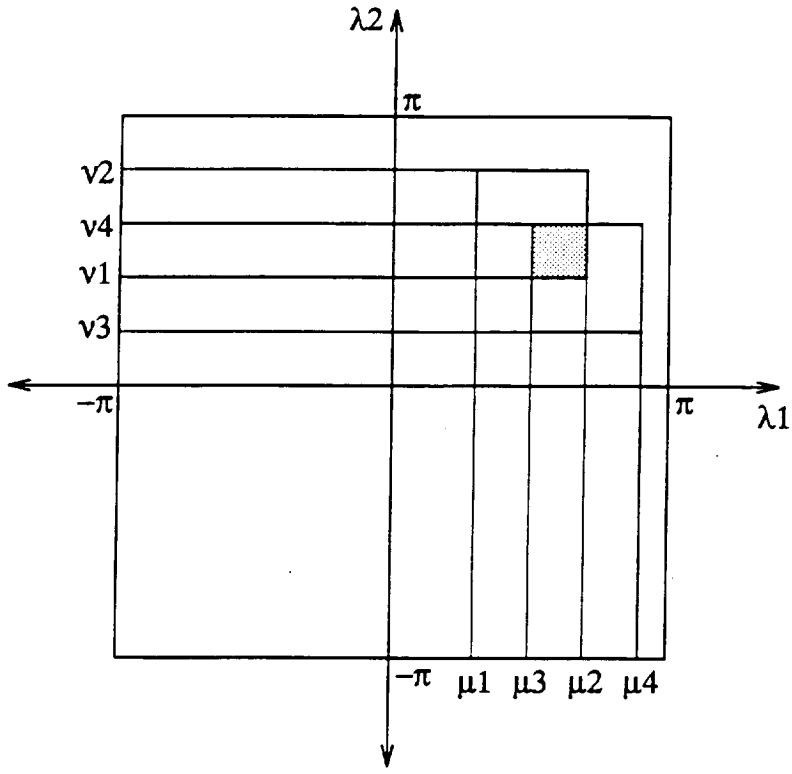

Figure 3.1. Covariance of a two-dimensional orthogonal increment process.

The last line of the definition deserves some comment. In general, the joint spectral distribution (or density) is underdetermined by $X_1(t)$ and $X_2(t)$. However, under certain conditions it is possible to establish a relationship between the cross spectral distribution, $F_{12}(\lambda)$, and $F(\lambda_1, \lambda_2)$ at points away from the margins. In these cases, the joint spectral distribution can be uniquely determined for a given pair of random sequences. Initially it might seem that this idea violates the well known maxim that the joint probability density can not be uniquely determined from the marginal densities. However, there is more information available than what is contained in the marginal spectra. Recall that the power spectral density is not a sufficient statistic for a random sequence since the phase information is discarded by taking the magnitude squared of the orthogonal increment process, $Z(\lambda)$. Thus conditions may exist under which the cross-correlation or equivalently the cross-spectrum contains the additional information necessary to form the joint density.

To construct the joint spectral density of a bivariate random sequence, it is necessary to take the information in the spectral density matrix, (3.15), and somehow form it into a two-dimensional spectral density function like $f(\lambda_1, \lambda_2)$. Consider the spectral representation of the autocorrelation of some random field, $X(t, u)$, that was given in (3.19)

$$R(\tau, \upsilon) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j(\lambda_1 \tau + \lambda_2 \upsilon)} dF(\lambda_1, \lambda_2).$$

Notice that setting either τ or υ equal to zero is equivalent to integrating out λ_1 or λ_2 respectively,

$$R_{11}(\tau) = R(\tau, 0) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j\lambda_1 \tau} dF(\lambda_1, \lambda_2) = \int_{-\pi}^{\pi} e^{j\lambda_1 \tau} dF_{11}(\lambda_1), \qquad (3.25)$$

$$R_{22}(\tau) = R(0, \tau) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j\lambda_2 \tau} dF(\lambda_1, \lambda_2) = \int_{-\pi}^{\pi} e^{j\lambda_2 \tau} dF_{22}(\lambda_2), \qquad (3.26)$$

where $R_{11}(\tau)$ and $R_{22}(\tau)$ are the marginal correlation functions and $F_{11}(\lambda_1)$ and $F_{22}(\lambda_2)$ are the marginal spectral distributions. This is necessary for $F(\lambda_1, \lambda_2)$ to

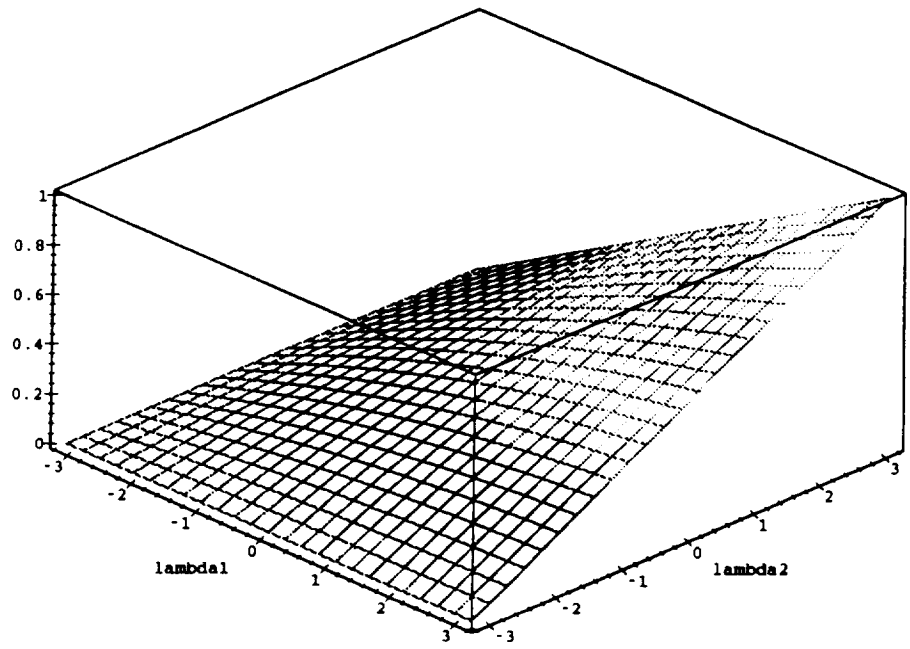

Figure 3.2. Example of a two-dimensional distribution function, $F(\lambda_1, \lambda_2)$.

satisfy the compatibility condition. Then $F_{11}(\lambda_1)$ and $F_{22}(\lambda_2)$ are related to $F(\lambda_1, \lambda_2)$ by

$$\begin{aligned} F_{11}(\lambda) &= Var[Z_1(\lambda) - Z_1(-\pi)] \\ &= Var[Z(\lambda, \pi) - Z(\lambda, -\pi) - Z(-\pi, \pi) + Z(-\pi, -\pi)] \\ &= F(\lambda, \pi), \end{aligned} \tag{3.27}$$

$$\begin{aligned} F_{22}(\lambda) &= Var[Z_2(\lambda) - Z_2(-\pi)] \\ &= Var[Z(\pi, \lambda) - Z(\pi, -\pi) - Z(-\pi, \lambda) + Z(-\pi, -\pi)] \\ &= F(\pi, \lambda). \end{aligned} \tag{3.28}$$

If the marginal spectral distributions of a stationary random field are set equal to the auto-spectral distributions of a pair of random sequences, then $F(\lambda_1, \lambda_2)$ will be consistent with the joint spectral distribution of those sequences along the boundaries $\lambda_1 = \pi$ and $\lambda_2 = \pi$. Figure 3.2 shows a plot of a two-dimensional distribution function on $(-\pi, \pi] \times (-\pi, \pi]$. Notice that on the boundaries along $\lambda_1 = -\pi$ and $\lambda_2 = -\pi$ the joint distribution function is equal to zero and that it is monotonically increasing in each of its arguments until they reach π where the joint distribution function is equal to the marginal distribution function. To complete the specification of the joint spectral distribution, it remains to determine $F(\lambda_1, \lambda_2)$ in the interior region away from the margins in terms of the joint statistics of $X_1(t)$ and $X_2(t)$.

Under certain symmetry conditions, knowing the diagonal slice, $F(\lambda, \lambda)$, in addition to the margins is sufficient to completely specify $F(\lambda_1, \lambda_2)$ over $(-\pi, \pi] \times (-\pi, \pi]$. The Gaussian distribution is an example since its density function exhibits elliptical symmetry (see Appendix B). The projection-slice theorem is an interesting property that relates slices through a two-dimensional correlation function to projections in its Fourier transform as depicted in Figure 3.3 [21]. If $R(\tau, \upsilon)$ is evaluated along the slice formed by setting $\upsilon = \tau$, then

$$\tilde{R}(\tau) = R(\tau, \tau) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j(\lambda_1 + \lambda_2)\tau} dF(\lambda_1, \lambda_2), \tag{3.29}$$

results. Taking the one-dimensional Fourier transform of $\tilde{R}(\tau)$ yields a projection of the joint power spectral density along the 45° line formed by setting $\lambda_2 = \lambda_1$.

$$\tilde{f}(\lambda) = \frac{1}{2\pi}\sum_{-\infty}^{\infty}\tilde{R}(\tau)e^{-j\lambda\tau} \tag{3.30}$$

$$= \frac{1}{2\pi}\sum_{-\infty}^{\infty}e^{-j\lambda\tau}\left[\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}e^{j(\lambda_1+\lambda_2)\tau}dF(\lambda_1,\lambda_2)\right]. \tag{3.31}$$

Assuming $F(\lambda_1, \lambda_2)$ is differentiable, i.e.,

$$dF(\lambda_1,\lambda_2) = f(\lambda_1,\lambda_2)\,d\lambda_1\,d\lambda_2, \tag{3.32}$$

then

$$\begin{aligned}\tilde{f}(\lambda) &= \frac{1}{2\pi}\sum_{-\infty}^{\infty}e^{-j\lambda\tau}\left[\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}e^{j(\lambda_1+\lambda_2)\tau}f(\lambda_1,\lambda_2)\,d\lambda_1\,d\lambda_2\right]\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left[\sum_{-\infty}^{\infty}e^{j(\lambda_1+\lambda_2-\lambda)\tau}\right]f(\lambda_1,\lambda_2)\,d\lambda_1\,d\lambda_2\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}f(\lambda_1,\lambda-\lambda_1)\,d\lambda_1. \end{aligned}\tag{3.33}$$

If $f(\lambda_1, \lambda_2)$ is separable in its arguments such that

$$f(\lambda_1,\lambda_2) = f_1(\lambda_1)f_2(\lambda_2),$$

then (3.33) reduces to the convolution of $f_1(\lambda_1)$ and $f_2(\lambda_2)$. Notice that in this case, $\tilde{f}(\lambda)$ corresponds to the probability density function of the sum of λ_1 and λ_2 where they are statistically independent random variables.

The author knows of no existing method for computing arbitrary slices $R(\tau, \upsilon)$ that are not on the margins. At this point we reach an impasse and need to find a way to further constrain the problem. The interrelationships among the various time and spectral domain representations are summarized in Figure 3.4. Lines that have arrows on both ends signify Fourier type operations which are bidirectional. Lines with arrows on just one end signify one-way operations that are either of the covariance type, or involve extracting the margins. Starting from $X_1(t)$ and $X_2(t)$

at the bottom notice that there is no direct path to the goal of a joint spectral distribution at the top. The premise is that given certain restrictions on $X_1(t)$ and $X_2(t)$, the combined information in the auto- and cross-spectra (or correlations) can form a link to $F(\lambda_1, \lambda_2)$ (or $R(\tau, \upsilon)$).

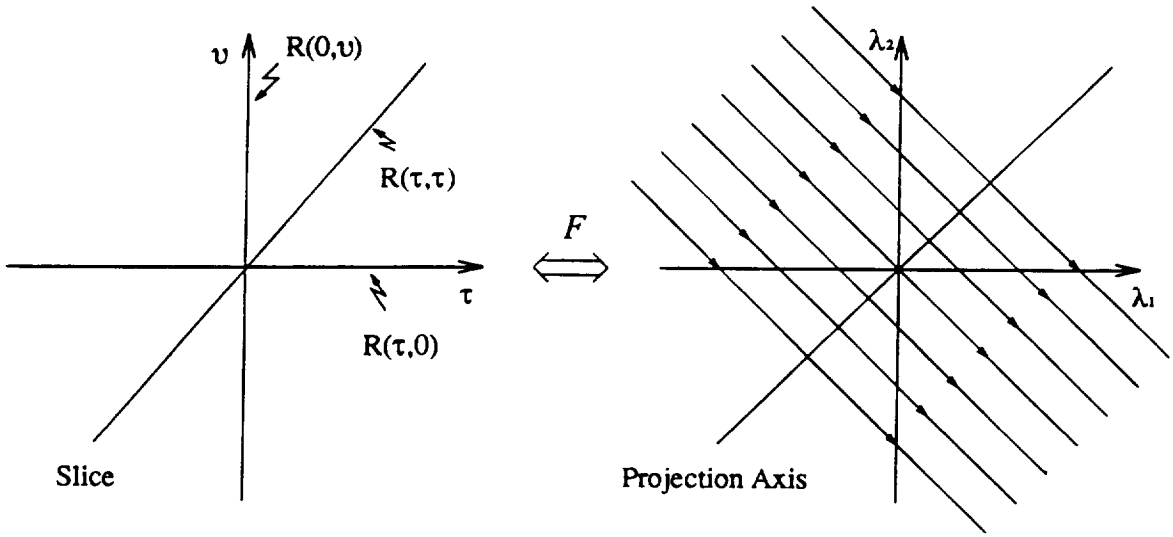

Figure 3.3. A slice of the two-dimensional correlation function and the corresponding projection of the spectral density.

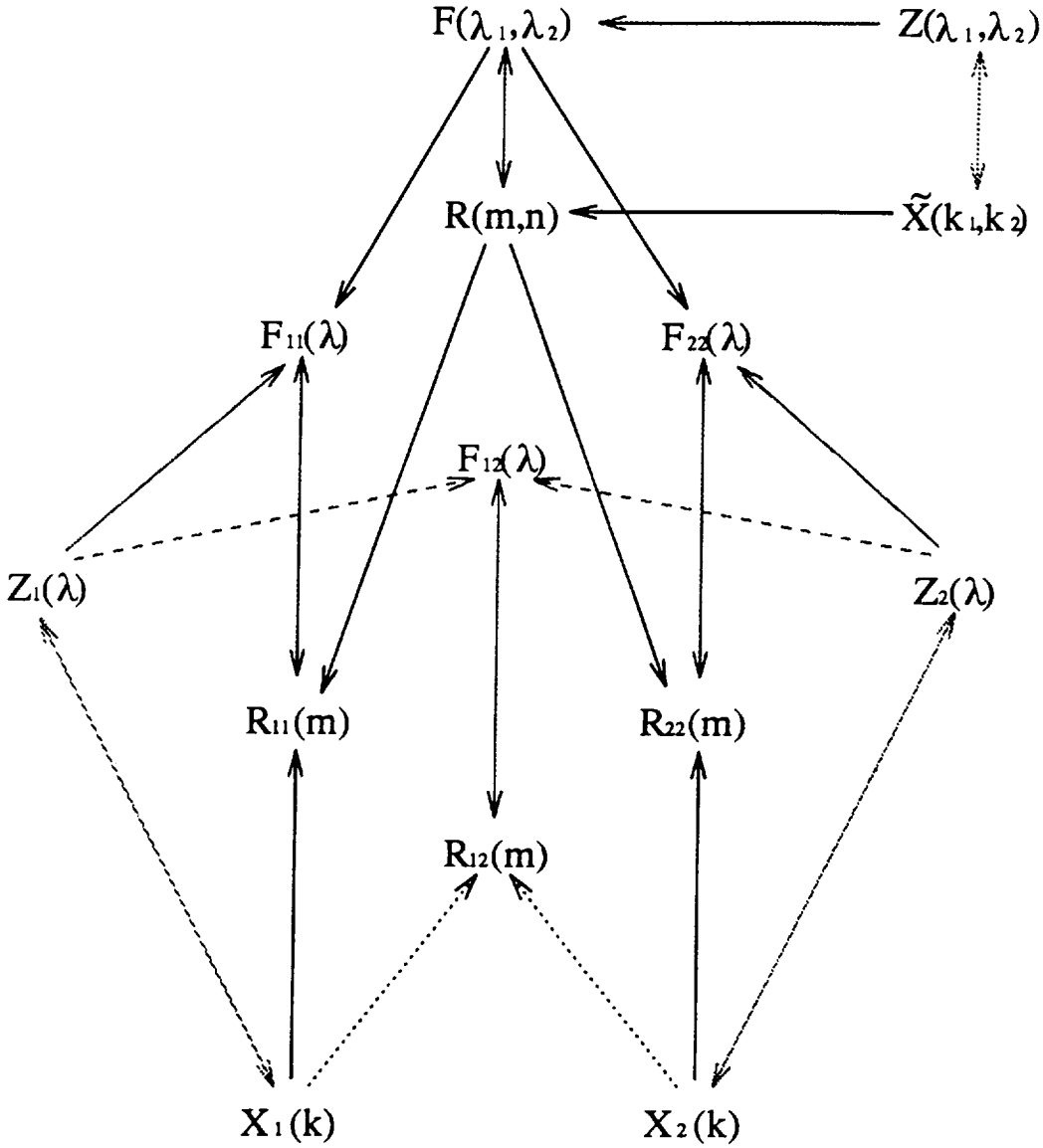

Figure 3.4. Schematic diagram of the relationships among various time-domain and spectral representations.

CHAPTER 4

MARGINAL SEQUENCES

4.1 Introduction

One way to constrain a pair of random sequences so that they uniquely specify the joint spectra density is to assume that they are the marginal sequences of some stationary random field. This restriction implies that the spectra of the two sequences correspond to the marginal spectra of the random field, but in general the converse is not true. That is, if the spectra of two sequences are the marginal spectra of a stationary random field, it does not necessarily follow that the sequences are the marginal sequences of that or any other stationary random field. This is a consequence of the "many-to-one" relationship between sequences and their power spectral densities. It is well known that many sequences may share the same power spectral density, but that any particular sequence has one and only one power spectral density [22].

4.2 Basic Concepts

The marginal sequences of a two-dimensional stationary random field, $\{X(t,u)\}$, are defined as the pair of random sequences that result from setting $t = 0$ and $u = 0$ respectively,

$$X_1(t) = X(t,0)$$

$$X_2(u) = X(0,u).$$

Figure 4.1 shows how the marginal process are related to the random field as a whole. One of the distinguishing properties of marginal processes that is clearly illustrated in Figure 4.1 is that the two sequences must start at the same value.

$$X_1(0) = X_2(0) = X(0,0).$$

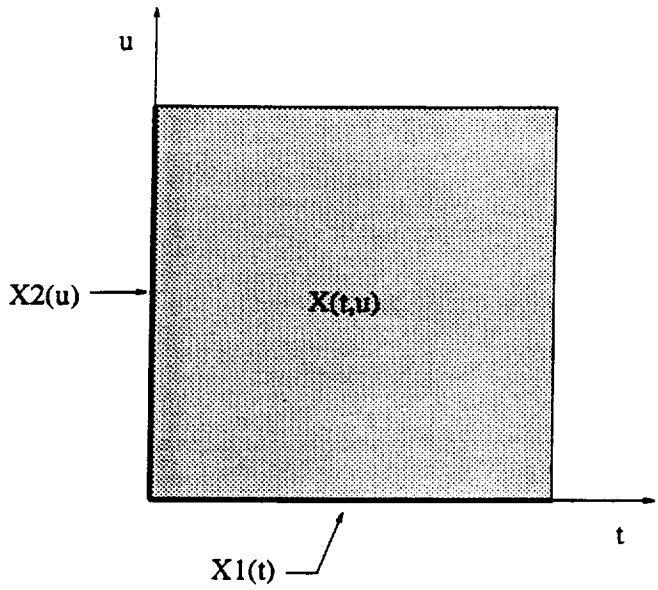

Figure 4.1. Marginal process of a two-dimensional random field.

Further restrictions result directly from the assumption that the random field is stationary.

Theorem 4.1 *The marginal sequences of a stationary random field are individually stationary.*

This theorem follows directly from the definitions of a stationary random field and its marginal processes. A random field, $\{X(t,u)\}$ is stationary iff

$$R_{xx}(\tau,v) = E[X^*(t,u)X(t+\tau,u+v)]. \tag{4.1}$$

Letting $u = 0$ and $v = 0$, gives

$$\begin{aligned} R_{xx}(\tau,0) &= E[X^*(t,0)X(t+\tau,0)], \\ &= E[X_1^*(t)X_1(t+\tau)], \\ &= R_{11}(\tau). \end{aligned} \tag{4.2}$$

Similarly it can be shown that $X_2(t)$ is also stationary. From (4.2) it is clear that the autocorrelations of the marginal processes correspond to the margins of the autocorrelation of the stationary random field. That is,

$$R_{11}(\tau) = R_{xx}(\tau, 0)$$

$$R_{22}(\tau) = R_{xx}(0, \tau).$$

Theorem 4.2 *The marginal sequences of a stationary random field, $X(t,u)$, are jointly stationary iff the autocorrelation of $X(t,u)$ can be written in the form*

$$R_{xx}(\tau, v) = R_{xx}(v - \tau).$$

Again, the proof follows directly from the definitions of a stationary random field and its marginal processes. In (4.1) let $u = 0$ and $t + \tau = 0$, then

$$\begin{aligned} R_{xx}(\tau, v) &= E[X^*(-\tau, 0)X(0, v)], \\ &= E[X_1^*(-\tau)X_2(v)], \\ &= R_{12}(-\tau, v). \end{aligned} \tag{4.3}$$

Only in the special case where $R_{xx}(\tau, v) = R_{xx}(v - \tau)$ will the marginal sequences be jointly stationary. Returning to (4.1), if we let $t = 0$ and $u + v = 0$, then

$$\begin{aligned} R_{xx}(\tau, v) &= E[X^*(0, -v)X(\tau, 0)], \\ &= E[X_2^*(-v)X_1(\tau)], \\ &= R_{21}(-v, \tau). \end{aligned} \tag{4.4}$$

Comparing (4.3) and (4.4), gives

$$R_{12}(-\tau, v) = R_{21}(-v, \tau).$$

Since $R_{xx}(\tau, v) = R_{xx}^*(-\tau, -v)$, we also have

$$R_{12}(-\tau, v) = R_{12}^*(\tau, -v)$$

$$R_{21}(-v, \tau) = R_{21}^*(v, -\tau).$$

Therefore R_{12} and R_{21} are hermitian functions and are non-negative definite. As a result, F_{12} and F_{21} are real-valued and non-decreasing.

4.3 Properties in the Spectral Domain

It might be expected that the correlation relationships for marginal processes from the previous section have a significant effect on the spectral domain characterizations. In this section that is shown to indeed be that case.

Theorem 4.3 *If $X_i(t), \quad i = 1, 2$ are the marginal sequences of a stationary random field, $X(t, u)$, then*

$$Z_1(\lambda) = Z(\lambda, \pi) - Z(\lambda, -\pi), \tag{4.5}$$

$$Z_2(\lambda) = Z(\pi, \lambda) - Z(-\pi, \lambda), \tag{4.6}$$

where $Z_i(\lambda)$ are the orthogonal increment processes corresponding to the marginal sequences and $Z(\lambda_1, \lambda_2)$ is the orthogonal increment process for $X(t, u)$.

To prove this theorem, consider the spectral representation of a stationary random field,

$$X(t, u) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j(\lambda_1 t + \lambda_2 u)} dZ(\lambda_1, \lambda_2).$$

Setting $u = 0$ gives an expression for the marginal sequence in terms of the two-dimensional orthogonal increment process, $Z(\lambda_1, \lambda_2)$,

$$\begin{aligned} X_1(t) = X(t, 0) &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j\lambda_1 t} dZ(\lambda_1, \lambda_2), \\ &= \int_{-\pi}^{\pi} e^{j\lambda_1 t} \int_{-\pi}^{\pi} dZ(\lambda_1, \lambda_2), \end{aligned} \tag{4.7}$$

where the inner integral is with respect to λ_2 and outer integral is with respect to λ_1. Recalling that the spectral representation of $X_1(t)$ is given by

$$X_1(t) = \int_{-\pi}^{\pi} e^{j\lambda t} dZ_1(\lambda),$$

and comparing this with (4.7) gives a relationship between the marginal orthogonal increment process, $Z_1(\lambda)$ and the two-dimensional orthogonal increment process, $Z(\lambda_1, \lambda_2)$,

$$dZ_1(\lambda_1) = \int_{-\pi}^{\pi} dZ(\lambda_1, \lambda_2),$$

where the integral on the right hand side is with respect to λ_2. This expression can be evaluated by integrating both sides with respect to λ_1 over the interval $(a, b]$ and applying the definition Stieltjes integral.

$$\int_a^b dZ_1(\lambda_1) = \int_a^b \int_{-\pi}^{\pi} dZ(\lambda_1, \lambda_2)$$

$$Z_1(b) - Z_1(a) = \lim_{n\to\infty} \sum_{i,j=1}^{n} Z(\mu_i, \nu_j) - Z(\mu_i, \nu_{j-1}) - Z(\mu_{i-1}, \nu_j) + Z(\mu_{i-1}, \nu_{j-1})$$

where $a = \mu_0 < \cdots < \mu_n = b$ and $-\pi = \nu_0 < \cdots < \nu_n = \pi$. By the noting the cancelation of terms in the summation, it is easy to verify that

$$Z_1(b) - Z_1(a) = Z(b, \pi) - Z(b, -\pi) - Z(a, \pi) + Z(a, -\pi).$$

Equating terms on both sides gives

$$Z_1(b) = Z(b, \pi) - Z(b, -\pi)$$

$$Z_1(a) = Z(a, \pi) - Z(a, -\pi).$$

In a similar manner (4.6) can be proved.

Corollary 4.4 *If $Z_1(\lambda)$ and $Z_2(\lambda)$ are orthogonal increment processes corresponding to the random sequences $X_1(t)$ and $X_2(t)$ respectively, and are related to the two dimensional orthogonal increment process, $Z(\lambda_1, \lambda_2)$, by*

$$Z_1(\lambda) = Z(\lambda, \pi) - Z(\lambda, -\pi),$$

$$Z_2(\lambda) = Z(\pi, \lambda) - Z(-\pi, \lambda),$$

then

$$X_1(t) = X(t, 0),$$

$$X_2(t) = X(0, t),$$

where $X(t, u)$ *is the stationary random field corresponding to* $Z(\lambda_1, \lambda_2)$.

The proof is a simple variation on the proof of Theorem 4.3 and will not be given here.

Setting the starting point of the orthogonal increment processes to zero which results in no loss of generality, we get the following expressions:

$$Z_1(\lambda) = Z(\lambda, \pi)$$

$$Z_2(\lambda) = Z(\pi, \lambda).$$

Substituting $Z(\lambda, \pi)$ and $Z(\pi, \lambda)$ into the expression for the cross-spectral distribution gives

$$\begin{aligned} F_{12}(\lambda) &= E[Z_1^*(\lambda) Z_2(\lambda)] \\ &= Cov[Z(\lambda, \pi), Z(\pi, \lambda)]. \end{aligned}$$

Recall that $Z(\lambda_1, \lambda_2)$ is a two-dimensional orthogonal increment process so the covariance of $Z(\lambda_1, \lambda_2)$ over two intervals is only non-zero where the two intervals overlap. If the two intervals are disjoint then the covariance is identically zero. Therefore, referring to Figure 4.2, it is easy to show that

$$Cov[Z(\lambda_1, \pi), Z(\pi, \lambda_2)] = F(\lambda_1, \lambda_2). \tag{4.8}$$

Hence for this special case, the cross spectral distribution is equal to the joint spectral distribution along the diagonal $\lambda_1 = \lambda_2$.

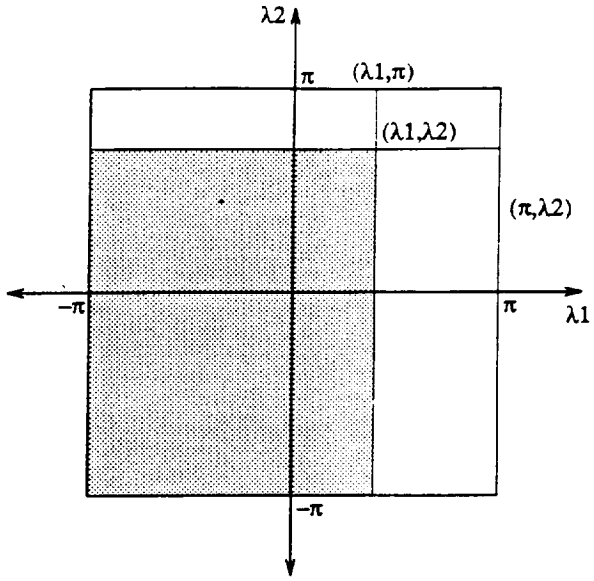

Figure 4.2. Domain of a two-dimensional orthogonal increment process.

4.4 Jointly Stationary Marginal Sequences

To evaluate the effect that assuming the marginal sequences are jointly stationary has on the JSD, start with the expression for the correlation of a stationary random field

$$R(\tau, \upsilon) = E\left[X^*(t,u)X(t+\tau, u+\upsilon)\right].$$

Suppose $R(\tau, \upsilon)$ is evaluated along $\upsilon = 0$, then

$$\begin{aligned} R(\tau,0) &= E\left[X^*(t,u)X(t+\tau,u)\right], \\ &= E\left[X^*(0,0)X(\tau,0)\right], \\ &= \underbrace{E\left[X_1^*(0)X_1(\tau)\right]}_{R_{11}(\tau)} = \underbrace{E\left[X_2^*(0)X_1(\tau)\right]}_{R_{21}(\tau)}. \end{aligned} \tag{4.9}$$

Similarly, if $R(\tau, v)$ is evaluated along $\tau = 0$, then

$$\begin{aligned} R(0,v) &= E\left[X^*(t,u)X(t,u+v)\right], \\ &= E\left[X^*(0,0)X(0,v)\right], \\ &= \underbrace{E\left[X_1^*(0)X_2(v)\right]}_{R_{12}(v)} = \underbrace{E\left[X_2^*(0)X_2(v)\right]}_{R_{22}(v)}. \end{aligned} \tag{4.10}$$

Also since both $X_1(0)$ and $X_2(0)$ are equal to $X(0,0)$,

$$R_{11}(\tau) = E\left[X_1^*(0)X_1(\tau)\right] = E[\,X_2^*(0)X_1(\tau)] = R_{22}(\tau). \tag{4.11}$$

Comparing (4.11) to (4.9) and (4.10) it is recognized that all of the auto- and cross-correlations functions are identical,

$$R_{11}(\tau) = R_{12}(\tau) = R_{21}(\tau) = R_{22}(\tau).$$

In the spectral domain this implies that

$$F_{11}(\lambda) = F_{12}(\lambda) = F_{21}(\lambda) = F_{22}(\lambda).$$

Recalling that

$$F_{12}(\lambda) = F(\lambda, \lambda),$$

we have

$$F(\lambda, \pi) = F(\lambda, \lambda) = F(\pi, \lambda).$$

Unfortunately this is a trivial result from the standpoint of estimating a joint spectral density given an arbitrary pair of random sequences because the two sequences are required to have the same PSD. Even so, there is some pedagogical benefit to carrying this a bit further with a couple of simple examples.

4.5 Some Examples

A class of two-dimensional distribution functions defined on $(-\pi, \pi] \times (-\pi, \pi]$ satisfying the conditions,

$$F(\lambda, \pi) = F(\lambda, \lambda) = F(\pi, \lambda),$$

can be constructed as follows:

1. Choose any valid one-dimensional distribution function $G(\lambda)$,
2. Set $F(\lambda_1, \lambda_2) = G(\min(\lambda_1, \lambda_2))$.

Since $F(\lambda_1, \lambda_2)$ must be non-decreasing in each of its arguments, it will have contours of constant value as shown in Figure 4.3. The joint spectral density is given by

$$\begin{aligned} f(\lambda_1, \lambda_2) &= \frac{F(\lambda_1, \lambda_2)}{d\lambda_1\, d\lambda_2} \\ &= \frac{G(\min(\lambda_1, \lambda_2))}{d\lambda_1\, d\lambda_2}. \end{aligned}$$

Note that the $\min(\lambda_1, \lambda_2)$ can be written as

$$\min(\lambda_1, \lambda_2) = \frac{1}{2}(\lambda_1 + \lambda_2 - |\lambda_1 - \lambda_2|).$$

By applying the theory of generalized functions (also known as the theory of distributions) we can evaluate the derivative of $\min(\lambda_1, \lambda_2)$ with respect to λ_1 and λ_2 as follows:

$$\begin{aligned} \frac{d}{d\lambda_1 d\lambda_2} \min(\lambda_1, \lambda_2) &= \frac{1}{2} \frac{d}{d\lambda_1 d\lambda_2} (\lambda_1 + \lambda_2 - |\lambda_1 - \lambda_2|), \\ &= \frac{1}{2} \frac{d}{d\lambda_1} (1 + \operatorname{sgn}(\lambda_1 - \lambda_2)), \\ &= \frac{1}{2} \frac{d}{d\lambda_1} (1 + (2u(\lambda_1 - \lambda_2) - 1)), \\ &= \frac{d}{d\lambda_1} u(\lambda_1 - \lambda_2), \\ &= \delta(\lambda_1 - \lambda_2), \end{aligned}$$

where $\operatorname{sgn}(x)$ is the signum function and $u(x)$ is the unit step function [23, 24, 25].

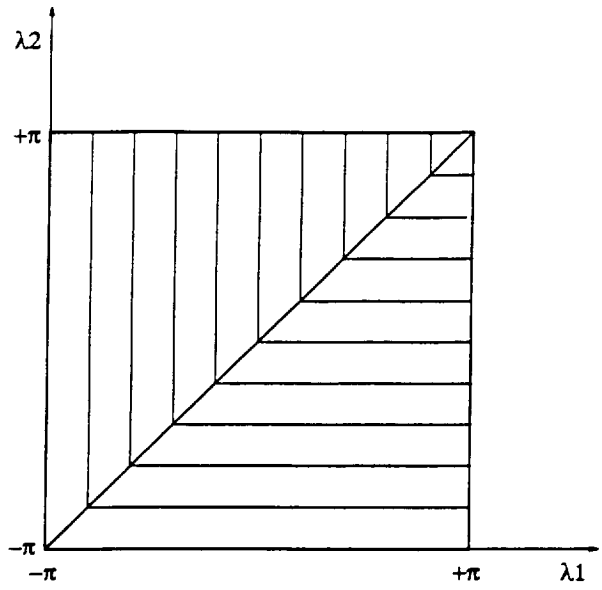

Figure 4.3. Contour plot of $F(\lambda_1, \lambda_2)$ for jointly stationary marginal sequences.

Applying the chain rule of differentiation, it can be shown that

$$f(\lambda_1, \lambda_2) = g(\min(\lambda_1, \lambda_2))\delta(\lambda_1 - \lambda_2),$$

where $g(\lambda)$ is the density function corresponding to $G(\lambda)$. The joint spectral density is concentrated on the line $\lambda_1 = \lambda_2$ and is zero everywhere else. The following two examples illustrate this class of JSD.

4.5.1 Uniform Margins

The distribution function, $F(\lambda_1, \lambda_2)$, when the marginal spectra are uniform is given by

$$F(\lambda_1, \lambda_2) = \begin{cases} 0 & \lambda_1 \text{ or } \lambda_2 \leq -\pi, \\ \sigma^2\left[\frac{1}{2}\left(\lambda_1 + \lambda_2 - |\lambda_1 - \lambda_2|\right) + \pi\right] & -\pi < \lambda_1, \lambda_2 \leq \pi, \\ \frac{\sigma^2(\lambda_1+\pi)}{2\pi} & -\pi < \lambda_1 \leq \pi,\ \lambda_2 \geq \pi, \\ \frac{\sigma^2(\lambda_2+\pi)}{2\pi} & \lambda_1 \geq \pi,\ -\pi < \lambda \leq \pi, \\ \sigma^2 & \lambda_1 \text{ and } \lambda_2 \geq \pi. \end{cases}$$

This function is plotted in Figure 4.4. The marginal spectral distributions $F_{11}(\lambda_1)$ and $F_{22}(\lambda_2)$ are equal to each other and correspond to a standard uniform distribution on $(-\pi, \pi]$. The joint spectral distribution, however, is not strictly uniform in the sense that the joint spectral density is not flat over $(-\pi, \pi] \times (-\pi, \pi]$.

In order to find the autocorrelation function $R(\tau, \upsilon)$, first evaluate

$$\begin{aligned} f(\lambda_1, \lambda_2) &= \frac{dF(\lambda_1, \lambda_2)}{d\lambda_1 d\lambda_2} \\ &= \frac{\sigma^2 \delta(\lambda_1 - \lambda_2)}{2\pi}. \end{aligned}$$

The spectral density corresponds to a uniform density concentrated along the diagonal $\lambda_1 = \lambda_2$.

Substituting $f(\lambda_1, \lambda_2)$ into (3.19) gives

$$\begin{aligned} R(\tau, \upsilon) &= \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j(\lambda_1 \tau + \lambda_2 \upsilon)} \delta(\lambda_1 - \lambda_2)\, d\lambda_1 d\lambda_2, \\ &= \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} e^{j(\tau+\upsilon)\lambda_2}\, d\lambda_2. \end{aligned}$$

$$\int_{-\pi}^{\pi} e^{j(\tau+\upsilon)\lambda_2}\, d\lambda_2 = \begin{cases} 2\pi & \tau = -\upsilon \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$R(\tau, \upsilon) = \sigma^2 \delta(\tau + \upsilon).$$

The autocorrelation is thus a train of impulses running along the diagonal $\tau = -\upsilon$.

4.5.2 Gaussian Margins

The distribution function, $F(\lambda_1, \lambda_2)$, for the case of Gaussian marginal spectra with zero mean and unit variance is given by

$$F(\lambda_1, \lambda_2) = \begin{cases} 0 & \lambda_1 \text{ or } \lambda_2 \leq -\pi, \\ Q\left[\frac{\frac{1}{2}(\lambda_1 + \lambda_2 - |\lambda_1 - \lambda_2|) - \mu}{\sigma}\right] & -\pi < \lambda_1, \lambda_2 \leq \pi, \\ Q(\frac{\lambda_1 - \mu}{\sigma}) & -\pi < \lambda_1 \leq \pi,\ \lambda_2 \geq \pi, \\ Q(\frac{\lambda_2 - \mu}{\sigma}) & \lambda_1 \geq \pi,\ -\pi < \lambda \leq \pi, \\ 1 & \lambda_1 \text{ and } \lambda_2 \geq \pi. \end{cases}$$

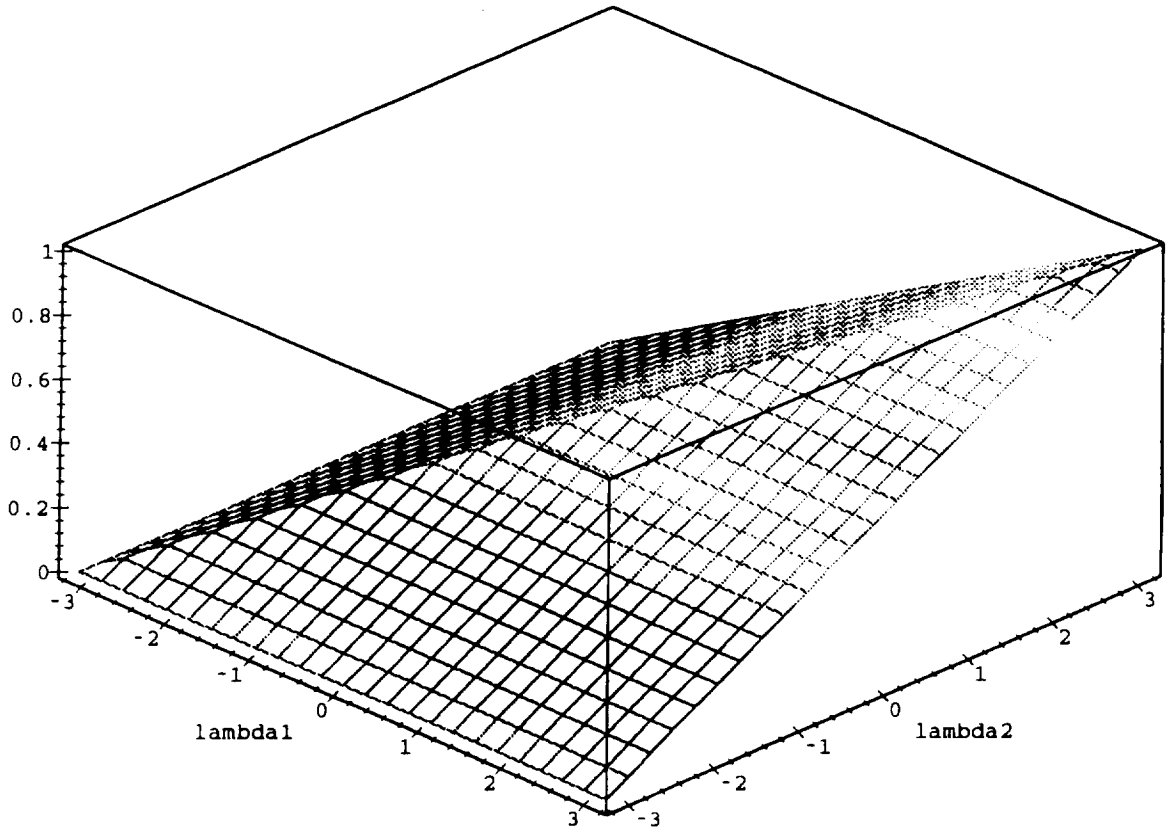

Figure 4.4. Plot of $F(\lambda_1, \lambda_2)$ for jointly stationary marginal sequences with uniform marginal spectra.

See Appendix B for the definition of the Q-function.

The marginal spectral distributions $F_{11}(\lambda_1)$ and $F_{22}(\lambda_2)$ are equal to each other and correspond to a univariate Gaussian distribution. Figure 4.5 shows a plot of this spectral distribution function.

The joint spectral density is given by

$$\begin{aligned} f(\lambda_1, \lambda_2) &= \frac{dF(\lambda_1, \lambda_2)}{d\lambda_1 d\lambda_2} \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\lambda_1-\mu)^2}{2\sigma^2}} \delta(\lambda_1 - \lambda_2). \end{aligned}$$

Substituting into (3.19) gives

$$\begin{aligned} R(\tau, \upsilon) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j(\lambda_1\tau + \lambda_2\upsilon)} e^{-\frac{(\lambda_1-\mu)^2}{2\sigma^2}} \delta(\lambda_1 - \lambda_2)\, d\lambda_1 d\lambda_2, \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\pi}^{\pi} e^{j(\tau+\upsilon)\lambda_2} e^{-\frac{(\lambda_2-\mu)^2}{2\sigma^2}}\, d\lambda_2, \\ &= e^{-\frac{1}{2}\sigma^2(\tau+\upsilon)^2} e^{j\mu(\tau+\upsilon)}. \end{aligned}$$

The envelope of this autocorrelation function is shown in Figure 4.6.

4.6 Jointly Harmonizable Marginal Sequences

Assuming that the marginal sequences are jointly stationary places an unacceptable constraint on the joint spectral density. This constraint can be relaxed however by considering a generalization of stationarity called harmonizability [15]. A harmonizable random sequence can be represented by

$$X(t) = \int_{-\pi}^{\pi} e^{j\lambda t} dZ(\lambda),$$

but whereas in the stationary case, $Z(\lambda)$ has orthogonal increments, for harmonizable sequences, the increments are not necessarily orthogonal. The autocorrelation is given by

$$\begin{aligned} R(t, u) &= E[X^*(t)X(u)] \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{-j\lambda t} e^{j\check{\lambda} u} E[dZ^*(\lambda) dZ(\check{\lambda})] \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j(\check{\lambda}u - \lambda t)} dF(\lambda, \check{\lambda}), \end{aligned} \tag{4.12}$$

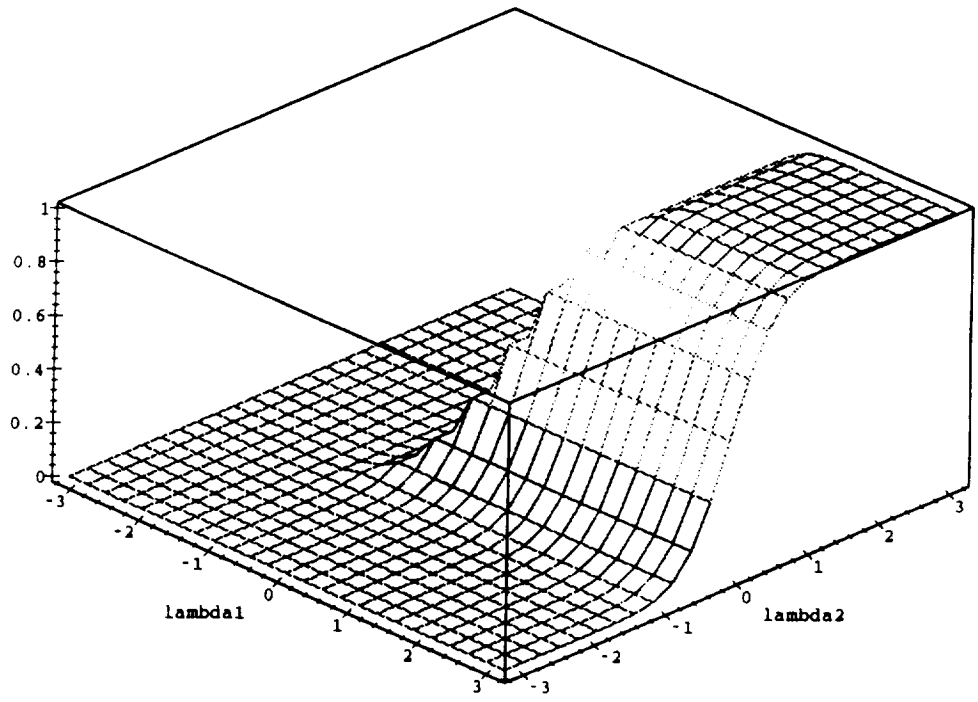

Figure 4.5. Plot of $F(\lambda_1, \lambda_2)$ for jointly stationary marginal sequences with Gaussian marginal spectra.

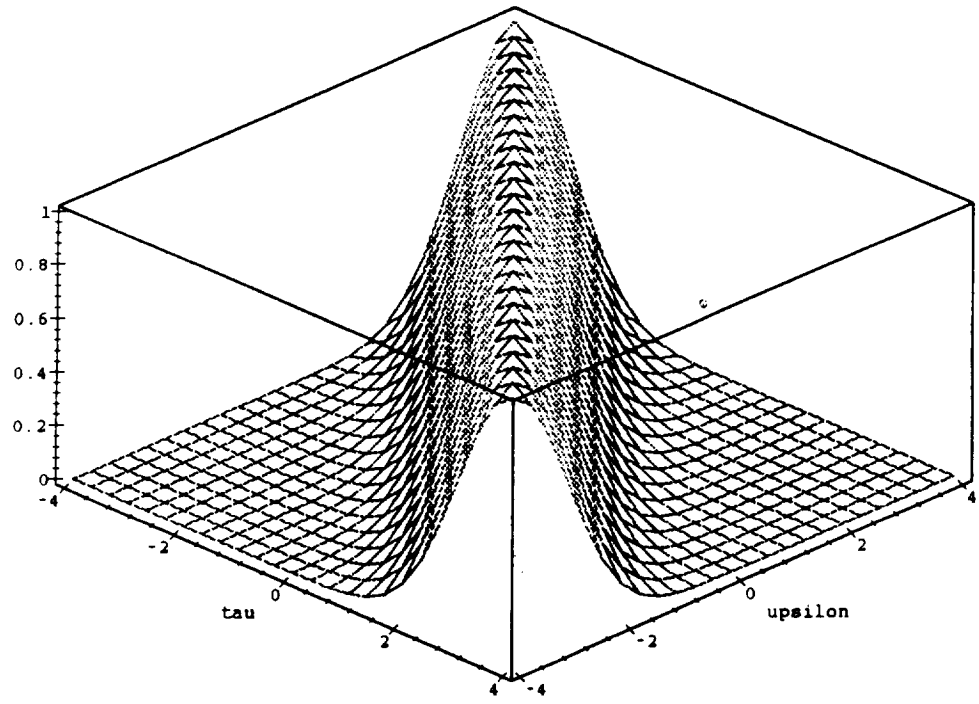

Figure 4.6. Plot of $R(\tau, v)$ for jointly stationary marginal sequences with Gaussian marginal spectra.

where $F(\lambda, \tilde{\lambda})$ is of bounded variation and in general complex valued. In the special case where $X(t)$ is stationary, (4.12) reduces to (2.6).

Suppose $X_1(t)$ and $X_2(t)$ are the marginal sequences of a stationary random field $X(t,u)$ and they are individually stationary, but jointly harmonizable. Then the autocorrelations will be one parameter functions of shifts in t, but the cross-correlations will be two parameter functions. Consider once again, the autocorrelation of the stationary random field

$$R(\tau, v) = E\left[X^*(t,u)X(t+\tau, u+v)\right].$$

Letting $u = 0$ and $t = -\tau$, gives

$$\begin{aligned} R(\tau, v) &= E\left[X^*(-\tau, 0)X(0, v)\right], \\ &= E\left[X_1^*(-\tau)X_2(0, v)\right], \\ &= R_{12}(-\tau, v). \end{aligned}$$

Similarly, letting $t = 0$ and $u = -v$, gives

$$\begin{aligned} R(\tau, v) &= E\left[X^*(0, v)X(\tau, 0)\right], \\ &= E\left[X_2^*(-v)X_1(\tau)\right], \\ &= R_{21}(-v, \tau). \end{aligned}$$

Note that the cross-correlations are non-stationary, so $R_{12}(\tau, v)$ and $R_{21}(\tau, v)$ can NOT be reduced to $R_{12}(v-\tau)$ and $R_{21}(v-\tau)$ respectively.

Recalling that the autocorrelation of a stationary random field is a Hermitian function i.e.,

$$R(\tau, v) = R^*(-\tau, -v),$$

it is easy to show that

$$\begin{aligned} R_{12}(-\tau, v) &= R_{12}^*(\tau, -v) \\ R_{21}(-v, \tau) &= R_{21}^*(v, -\tau). \end{aligned}$$

Therefore, the cross-spectral distributions are guaranteed to be real valued and non-decreasing.

4.7 Some Examples

4.7.1 Generalized Brownian Motion

Perhaps the simplest example of a two-dimensional orthogonal increment process is a variation on the Brownian sheet on $(-\pi, \pi] \times (-\pi, \pi]$. Let $\{Z(\lambda_1, \lambda_2),\ -\pi < \lambda_1, \lambda_2 \leq \pi\}$ be an orthogonal increment process that is Gaussian distributed, with

$$E[Z(\lambda_1, \lambda_2)] = 0, \tag{4.13}$$

$$Var[Z(\lambda_1, \lambda_2)] = \frac{\sigma^2}{4\pi^2}\left[(\lambda_1 + \pi)(\lambda_2 + \pi) + \frac{\kappa}{4\pi^2}\left(\lambda_1^2 - \pi^2\right)\left(\lambda_2^2 - \pi^2\right)\right]. \tag{4.14}$$

Then the corresponding distribution function, $F(\lambda_1, \lambda_2)$, is given by

$$F(\lambda_1, \lambda_2) = \begin{cases} 0 & \lambda_1 \text{ or } \lambda_2 \leq -\pi, \\ \frac{\sigma^2}{4\pi^2}\left[(\lambda_1 + \pi)(\lambda_2 + \pi) + \frac{\kappa}{4\pi^2}\left(\lambda_1^2 - \pi^2\right)\left(\lambda_2^2 - \pi^2\right)\right] & -\pi < \lambda_1, \lambda_2 \leq \pi, \\ \frac{\sigma^2(\lambda_1 + \pi)}{2\pi} & -\pi < \lambda_1 \leq \pi,\ \lambda_2 \geq \pi, \\ \frac{\sigma^2(\lambda_2 + \pi)}{2\pi} & \lambda_1 \geq \pi,\ -\pi < \lambda \leq \pi, \\ \sigma^2 & \lambda_1 \text{ and } \lambda_2 \geq \pi. \end{cases}$$

The marginal spectral distributions $F_{11}(\lambda_1)$ and $F_{22}(\lambda_2)$ are equal to each other and correspond to a uniform marginal spectral density. In general, the joint spectral distribution is not strictly uniform in the sense that the joint spectral density is not flat over $(-\pi, \pi] \times (-\pi, \pi]$ unless the parameter κ is set to zero. This particular family of generalized bivariate uniform density functions is called Morgenstern's bivariate uniform distribution (see Appendix A). It should be noted that other families of bivariate density functions exist that generate uniform marginal densities [26].

The joint spectral density is given by

$$\begin{aligned} f(\lambda_1, \lambda_2) &= \frac{dF(\lambda_1, \lambda_2)}{d\lambda_1 d\lambda_2} \\ &= \frac{\sigma^2}{4\pi^2}\left(1 + \frac{\kappa}{\pi^2}\lambda_1\lambda_2\right). \end{aligned}$$

Substituting into (3.19) gives

$$\begin{aligned} R(\tau,\upsilon) &= \frac{\sigma^2}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{j(\lambda_1\tau+\lambda_2\upsilon)}\left(1+\frac{\kappa}{\pi^2}\lambda_1\lambda_2\right)d\lambda_1 d\lambda_2, \\ &= \frac{\sigma^2}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{j(\lambda_1\tau+\lambda_2\upsilon)}\,d\lambda_1 d\lambda_2 + \frac{\sigma^2\kappa}{4\pi^4}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{j(\lambda_1\tau+\lambda_2\upsilon)}\lambda_1\lambda_2\,d\lambda_1 d\lambda_2 \\ &= \sigma^2\delta(\tau,\upsilon) + \frac{\sigma^2\kappa}{4\pi^4}\int_{-\pi}^{\pi}\lambda_1 e^{j\lambda_1\tau}\,d\lambda_1\int_{-\pi}^{\pi}\lambda_2 e^{j\lambda_2\upsilon}\,d\lambda_2. \end{aligned} \tag{4.15}$$

The two integrals on the right hand side of (4.15) evaluate to

$$\int_{-\pi}^{\pi}\lambda e^{j\lambda\tau}\,d\lambda = \frac{2j}{\tau^2}\left[\sin(\tau\pi) - \tau\pi\cos(\tau\pi)\right]. \tag{4.16}$$

A plot of this function for continuous τ is shown in Figure 4.7 If τ is restricted to integers, (4.16) reduces to

$$\begin{aligned} \int_{-\pi}^{\pi}\lambda e^{j\lambda\tau}\,d\lambda &= -\frac{2\pi j}{\tau}\cos(\tau\pi) \\ &= \frac{2\pi j}{\tau}(-1)^{\tau+1}. \end{aligned} \tag{4.17}$$

A plot of this function is shown in Figure 4.8.

Substituting (4.16) back into (4.15) gives

$$\begin{aligned} R(\tau,\upsilon) &= \sigma^2\delta(\tau,\upsilon) - \frac{\sigma^2\kappa}{4\pi^4\tau^2\upsilon^2}\left[\sin(\tau\pi) - \tau\pi\cos(\tau\pi)\right]\left[\sin(\upsilon\pi) - \upsilon\pi\cos(\upsilon\pi)\right] \\ &= \sigma^2\delta(\tau,\upsilon) - \frac{\sigma^2\kappa}{4\pi^4\tau^2\upsilon^2}\Big[\sin(\tau\pi)\sin(\upsilon\pi) - \tau\pi\cos(\tau\pi)\sin(\upsilon\pi) - \\ &\qquad \upsilon\pi\sin(\tau\pi)\cos(\upsilon\pi) + \tau\upsilon\pi^2\cos(\tau\pi)\cos(\upsilon\pi)\Big]. \end{aligned}$$

Restricting τ and υ to integers, reduces this to

$$R(\tau,\upsilon) = \sigma^2\delta(\tau,\upsilon) - (-1)^{\tau+\upsilon}\frac{\sigma^2\kappa}{4\pi^2\tau\upsilon}. \tag{4.18}$$

Evaluating slices in $R(\tau,\upsilon)$ corresponding to $\tau = 0$ and $\upsilon = 0$ gives

$$R_{11}(\tau) = R(\tau,0) = \sigma^2\delta(\tau),$$

$$R_{22}(\upsilon) = R(0,\upsilon) = \sigma^2\delta(\upsilon),$$

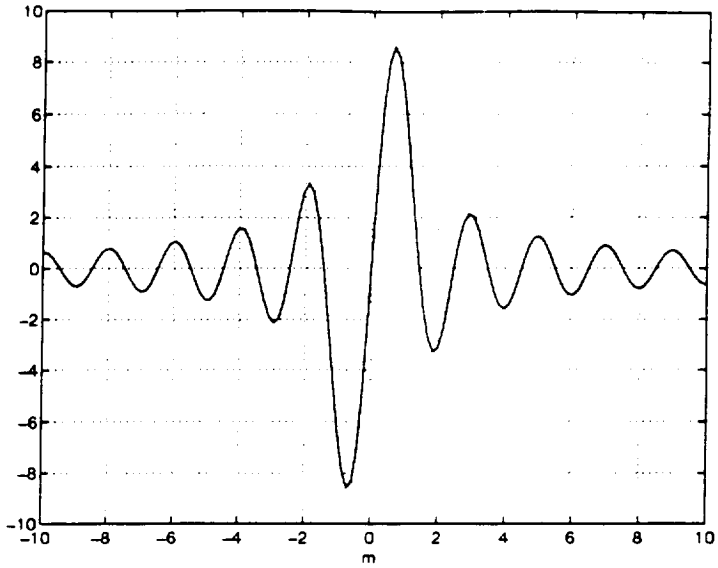

Figure 4.7. Plot of equation (4.16) for continuous τ.

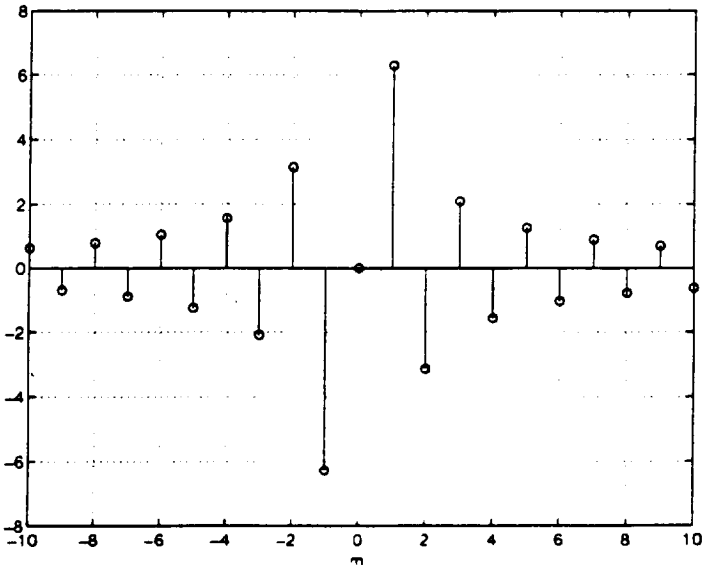

Figure 4.8. Plot of equation (4.17) for discrete τ.

and the cross-correlation is given by

$$R_{12}(\tau,\upsilon) = R(-\tau,\upsilon) = \sigma^2\delta(\tau,\upsilon) + (-1)^{\tau+\upsilon}\frac{\sigma^2\kappa}{4\pi^2\tau\upsilon}.$$

The two-dimensional autocorrelation for the generalized Brownian sheet is shown in Figure 4.9 with the exception of the impulse function at the origin. The parameter κ is proportional to the correlation coefficient between λ_1 and λ_2. Notice that as $\kappa \to 0$, the oscillations away from zero lag vanish and the autocorrelation becomes an impulse at the origin. This is to be expected because for $\kappa = 0$ the JSD is flat over $(-\pi,\pi] \times (-\pi,\pi]$.

4.7.2 Bivariate Gaussian

In the generalized Brownian motion example, even though the marginal sequences are not jointly stationary, the marginal spectra are identical. Let's consider a case where the marginal spectra are different from each other. If the joint spectral density is bivariate Gaussian then the marginal spectra will be univariate Gaussian, but in general they will have different means and variances. In the following example the means are set to zero for simplicity, but the variances are unequal.

Let $\{Z(\lambda_1,\lambda_2), -\pi \le \lambda_1,\lambda_2 \le \pi\}$ be an orthogonal increment process that is Gaussian distributed, with

$$E[Z(\lambda_1,\lambda_2)] = 0,$$

$$Var[Z(\lambda_1,\lambda_2)] = L\left(\frac{\lambda_1}{\sigma_1},\frac{\lambda_2}{\sigma_2},\rho\right) + Q\left(\frac{\lambda_1}{\sigma_1}\right) + Q\left(\frac{\lambda_2}{\sigma_2}\right) - L\left(\frac{\pi}{\sigma_1},\frac{\pi}{\sigma_2},\rho\right) - 1.$$

The corresponding spectral distribution, $F(\lambda_1,\lambda_2)$, is given by

$$F(\lambda_1,\lambda_2) = \begin{cases} 0 & \lambda_1 \text{ or } \lambda_2 \le -\pi, \\ L\left(\frac{\lambda_1}{\sigma_1},\frac{\lambda_2}{\sigma_2},\rho\right) + Q\left(\frac{\lambda_1}{\sigma_1}\right) + Q\left(\frac{\lambda_2}{\sigma_2}\right) - L\left(\frac{\pi}{\sigma_1},\frac{\pi}{\sigma_2},\rho\right) - 1 & -\pi \le \lambda_1,\lambda_2 \le \pi, \\ L\left(\frac{\lambda_1}{\sigma_1},\frac{\pi}{\sigma_2},\rho\right) + Q\left(\frac{\lambda_1}{\sigma_1}\right) + Q\left(\frac{\pi}{\sigma_2}\right) - L\left(\frac{\pi}{\sigma_1},\frac{\pi}{\sigma_2},\rho\right) - 1 & -\pi \le \lambda_1 \le \pi,\ \lambda_2 \\ L\left(\frac{\pi}{\sigma_1},\frac{\lambda_2}{\sigma_2},\rho\right) + Q\left(\frac{\pi}{\sigma_1}\right) - Q\left(\frac{\lambda_2}{\sigma_2}\right) - L\left(\frac{\pi}{\sigma_1},\frac{\pi}{\sigma_2},\rho\right) - 1 & \lambda_1 \ge \pi,\ -\pi \le \lambda \\ 1 & \lambda_1 \text{ and } \lambda_2 \ge \pi. \end{cases}$$

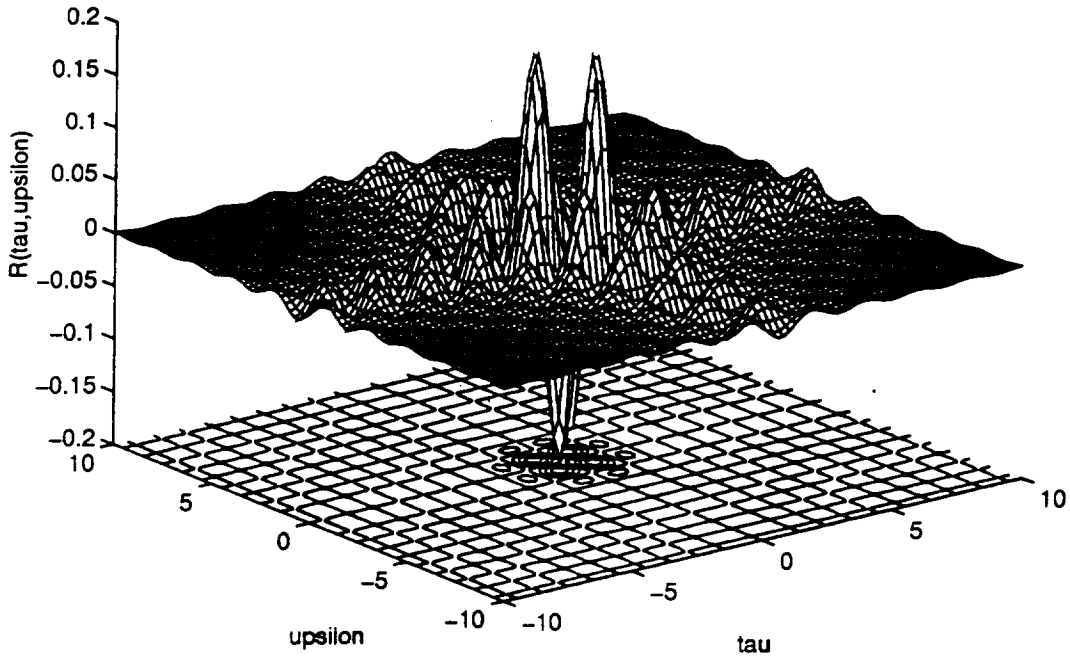

Figure 4.9. Plot of $R(\tau, v)$ for continuous τ and v.

See Appendix B for the definitions of the Q-function and the L-function. Unlike the uniform distribution, the bivariate Gaussian distribution is non-zero over all of R^2, so the tails of the Gaussian will be assumed to be sufficiently small at the boundaries of $(-\pi, \pi] \times (-\pi, \pi]$ that

$$\begin{aligned} Q\left(\frac{\pi}{\sigma_1}\right) &\approx 0, \\ Q\left(\frac{\pi}{\sigma_2}\right) &\approx 0, \\ L\left(\frac{\lambda_1}{\sigma_1}, \frac{\pi}{\sigma_2}, \rho\right) &\approx 0, \\ L\left(\frac{\pi}{\sigma_1}, \frac{\lambda_2}{\sigma_2}, \rho\right) &\approx 0, \\ L\left(\frac{\pi}{\sigma_1}, \frac{\pi}{\sigma_2}, \rho\right) &\approx 0. \end{aligned}$$

Otherwise the truncated bivariate Gaussian distribution should be used [27]. If this assumption is valid, then $F(\lambda_1, \lambda_2)$ reduces to

$$F(\lambda_1, \lambda_2) = \begin{cases} 0 & \lambda_1 \text{ or } \lambda_2 \leq -\pi, \\ L\left(\frac{\lambda_1}{\sigma_1}, \frac{\lambda_2}{\sigma_2}, \rho\right) + Q\left(\frac{\lambda_1}{\sigma_1}\right) + Q\left(\frac{\lambda_2}{\sigma_2}\right) - 1 & -\pi \leq \lambda_1, \lambda_2 \leq \pi, \\ Q\left(\frac{\lambda_1}{\sigma_1}\right) & -\pi \leq \lambda_1 \leq \pi,\ \lambda_2 \geq \pi, \\ Q\left(\frac{\lambda_2}{\sigma_2}\right) & \lambda_1 \geq \pi,\ -\pi \leq \lambda \leq \pi, \\ 1 & \lambda_1 \text{ and } \lambda_2 \geq \pi. \end{cases}$$

The joint spectral density is given by

$$f(\lambda_1, \lambda_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{\lambda_1^2}{\sigma_1^2} - \frac{2\rho\lambda_1\lambda_2}{\sigma_1\sigma_2} + \frac{\lambda_2^2}{\sigma_2^2}\right)\right], \quad -\pi \leq \lambda_1, \lambda_2 \leq \pi$$

Substituting into (3.19) gives

$$R(\tau, \upsilon) = \exp\left[-\frac{1}{2}(\sigma_1^2\tau^2 + 2\rho\sigma_1\sigma_2\tau\upsilon + \sigma_2^2\upsilon^2)\right].$$

The autocorrelations of the marginal sequences are given by

$$\begin{aligned} R_{11}(\tau) &= R(\tau, 0) = e^{-\frac{1}{2}\sigma_1^2\tau^2}, \\ R_{22}(\tau) &= R(0, \tau) = e^{-\frac{1}{2}\sigma_2^2\tau^2}, \end{aligned}$$

and the cross-correlation is given by

$$R_{12}(\tau, v) = R(-\tau, v) = \exp\left[-\frac{1}{2}(\sigma_1^2\tau^2 - 2\rho\sigma_1\sigma_2\tau v + \sigma_2^2 v^2)\right].$$

See Appendix B for plots of the bivariate Gaussian.

4.8 Simulating Marginal Spectra with Particular JSDs

Marginal sequences with specified joint spectral densities can be simulated using a variation on a technique proposed by Zrnić for generating I&Q sequences with weatherlike spectra [28]. Writing the complex valued random field in terms of the inphase and quadrature (I&Q) components,

$$X(t,u) = I(t,u) + jQ(t,u),$$

they can be modeled by

$$I(t,u) = s(t,u)\cos\phi(t,u) + n(t,u)\cos\psi(t,u) \tag{4.19}$$

$$Q(t,u) = s(t,u)\sin\phi(t,u) + n(t,u)\sin\psi(t,u) \tag{4.20}$$

where $s(t,u)$ and $n(t,u)$ are Rayleigh distributed signal and noise envelopes and $\phi(t,u)$ and $\psi(t,u)$ are uniformly distributed phases on $[0, 2\pi]$. The noise envelope, $n(t,u)$, is assumed to be broadband compared to $s(t,u)$. Expressing the I&Q fields in terms of a two dimensional Discrete Fourier Series (DFS) gives

$$I(t,u) + jQ(t,u) = \frac{1}{M^2}\sum_{k=1}^{M}\sum_{l=1}^{M}\sqrt{P(k,l)}\, e^{j\theta(k,l)} e^{-j\frac{2\pi}{M}(kt+lu)} \tag{4.21}$$

where $P(k,l)$ is the instantaneous power of the signal plus noise and $\theta(k,l)$ is a uniformly distributed phase on $[0, 2\pi]$. If the shape of the true JSD is given by $S(k,l)$ and the noise power per discrete frequency is a constant, N, then $P(k,l)$is given by

$$P(k,l) = -\left[S(k,l) + N\right]\ln U(k,l), \tag{4.22}$$

where $U(k,l)$ is a uniform random number on $[0,1]$. The steps for generating the I&Q field can be summarized as follows:

1. Select a desired spectral shape $S(k,l)$.
2. Set the signal-to-noise ratio (SNR).
3. Take the inverse discrete Fourier transform of the product of $\sqrt{P(k,l)}$ and an independent phase $e^{j\theta(k,l)}$.

The marginal sequences are easily extracted from the resulting I&Q field by setting

$$X_1(t) = X(t,0),$$

$$X_2(t) = X(0,t).$$

In the following two subsections, this simulation technique is used to generate the marginal sequences for the examples in subsections 4.5.2 and 4.7.2 respectively.

4.8.1 Jointly Stationary with Gaussian Marginal Spectra

In this example, the marginal sequences are jointly stationary with Gaussian marginal spectra. The ideal shape of the joint spectral density is

$$S(\lambda_1,\lambda_2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\lambda_1-\mu)^2}{2\sigma^2}} \delta(\lambda_1-\lambda_2).$$

Figure 4.10 shows the dB magnitude of the true JSD of the signal plus noise for a SNR of 100dB. The quadratic shape of the log of the Gaussian can be clearly distinguished rising out of the noise floor along $\lambda_1 = \lambda_2$.

Since the marginal sequences are jointly stationary, the auto-spectra are identical and equal to

$$f_{11}(\lambda) = f_{22}(\lambda) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\lambda_1-\mu)^2}{2\sigma^2}} \delta(\lambda_1-\lambda_2).$$

A plot of the auto-spectra of the simulated marginal sequences is shown in Figure 4.11. The joint spectral density is formed by setting $f(\lambda_1,\lambda_2)$ equal to $f_{11}(\lambda)$ along the diagonal and zero elsewhere.

If the assumption that the marginal sequences are jointly stationary is dropped, the non-stationary cross-correlation can be computed by taking the ensemble average

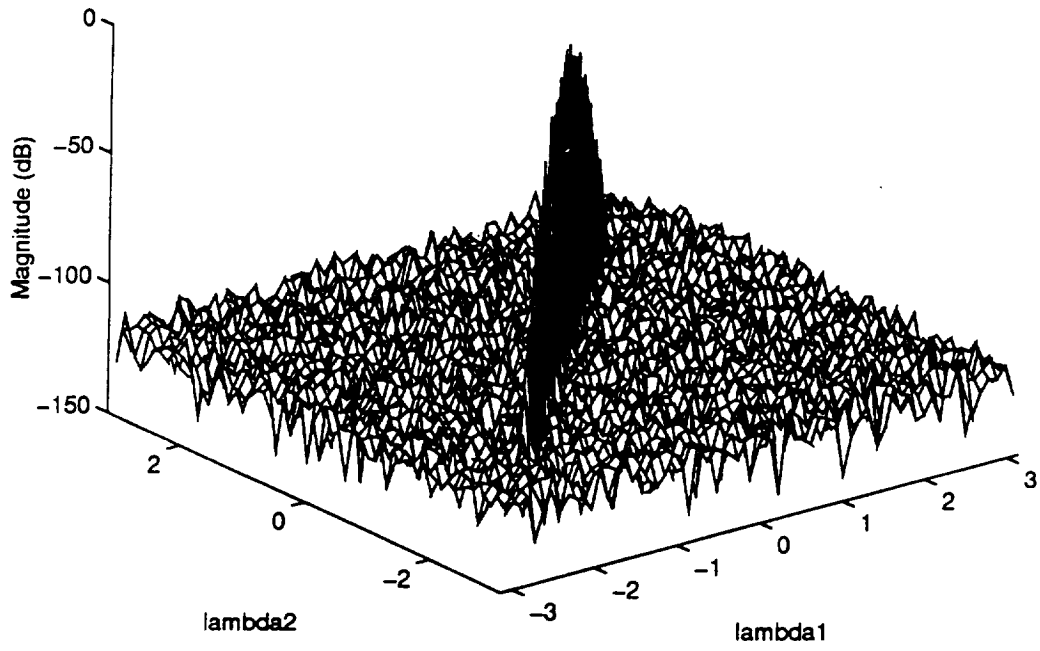

Figure 4.10. Plot of $P(\lambda_1, \lambda_2)$.

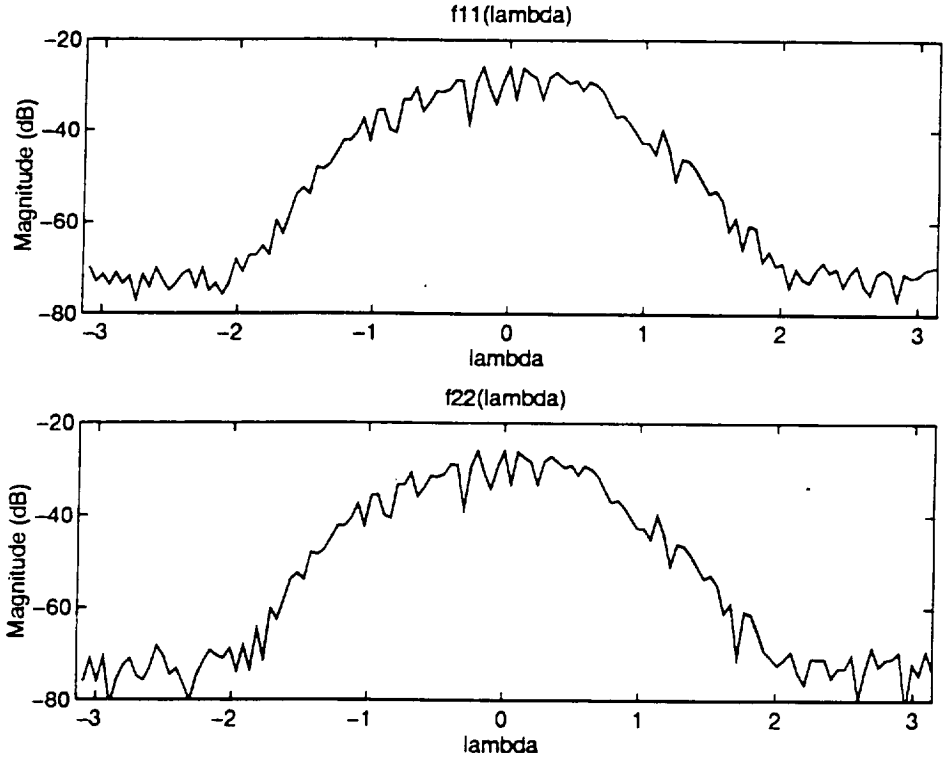

Figure 4.11. Plot of the marginal spectra $f_{11}(\lambda)$ and $f_{22}(\lambda)$.

over N independent realizations of the marginal sequences [29]

$$\hat{R}_{12}(t_1, t_2) = \frac{1}{N} \sum_{i}^{N} X^*_{1_i}(t_1) X_{2_i}(t_2). \tag{4.23}$$

Figure 4.12 shows the estimate of the JSD based on the two dimensional Fourier transform of the non-stationary cross-correlation where the ensemble average is taken over $N = 1000$ realizations. A ridge along $\lambda_1 = \lambda_2$ is still visible, but it is embedded in a structure dictated by the shape of the marginal spectra. For a single realization ($N = 1$), the non-stationary cross spectral estimate is given by

$$\hat{f}_{12}(\lambda_1, \lambda_2) = \hat{f}_{11}(\lambda_1) \hat{f}_{22}(\lambda_2).$$

The structure characteristic of these estimates of the non-stationary cross-correlation is an artifact of this outer-product of the sample auto-spectra. In the limit as $N \to \infty$, $\hat{f}(\lambda_1, \lambda_2)$ should approach the true JSD, but the convergence does not appear to be particularly fast.

4.8.2 Jointly Harmonizable with Bivariate Gaussian JSD

The case where the marginal sequences are jointly stationary is not a realistic one for estimating the JSD from the non-stationary cross-correlation since the JSD can be found much easier from the auto-spectra. Consider a more practical example where the marginal sequences are jointly harmonizable with a bivariate Gaussian shaped JSD. For this example the ideal shape of the joint spectral density is

$$S(\lambda_1, \lambda_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{\lambda_1^2}{\sigma_1^2} - \frac{2\rho\lambda_1\lambda_2}{\sigma_1\sigma_2} + \frac{\lambda_2^2}{\sigma_2^2}\right)\right].$$

In the simulations, $\sigma_1 = \sigma_2$ and $\rho = .95$, so that the presence of a correlation in the joint spectrum would be clearly visible. Figures 4.13 and 4.14 show a 3-D mesh plot and a contour plot respectively of the dB magnitude of the true JSD of the signal plus noise for a SNR of 100dB.

Using (4.23) to estimate the JSD, the plot shown in Figure 4.15 was generated. It is more difficult to see, but the desired gaussian shape is embedded in the central mound. It can be seen more clearly in the contour plot of Figure 4.16.

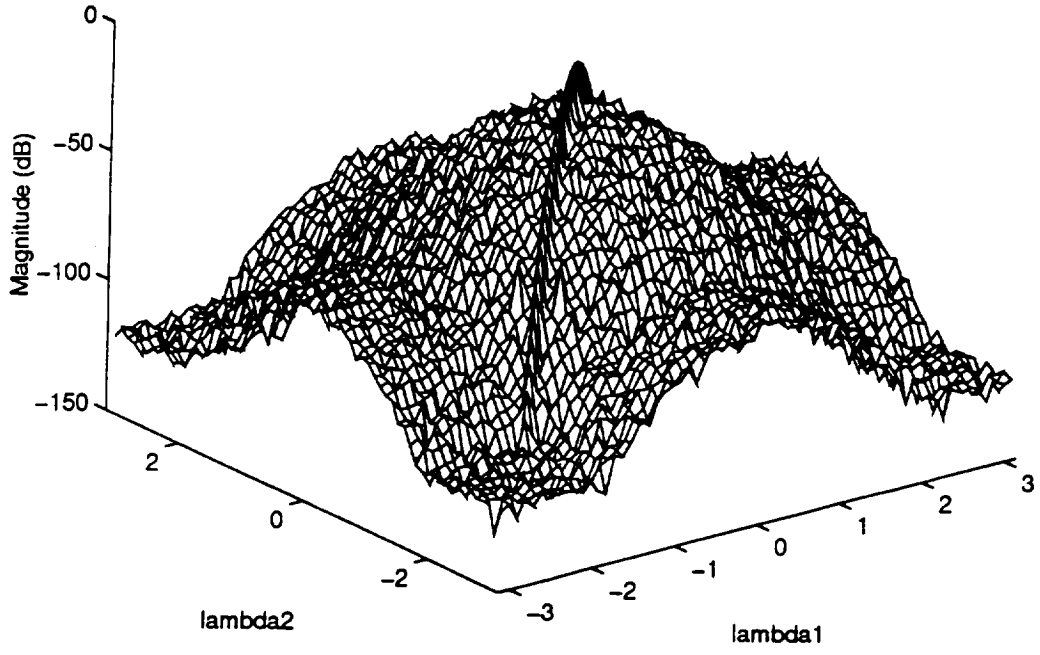

Figure 4.12. Plot of the JSD, $\hat{f}(\lambda_1, \lambda_2)$.

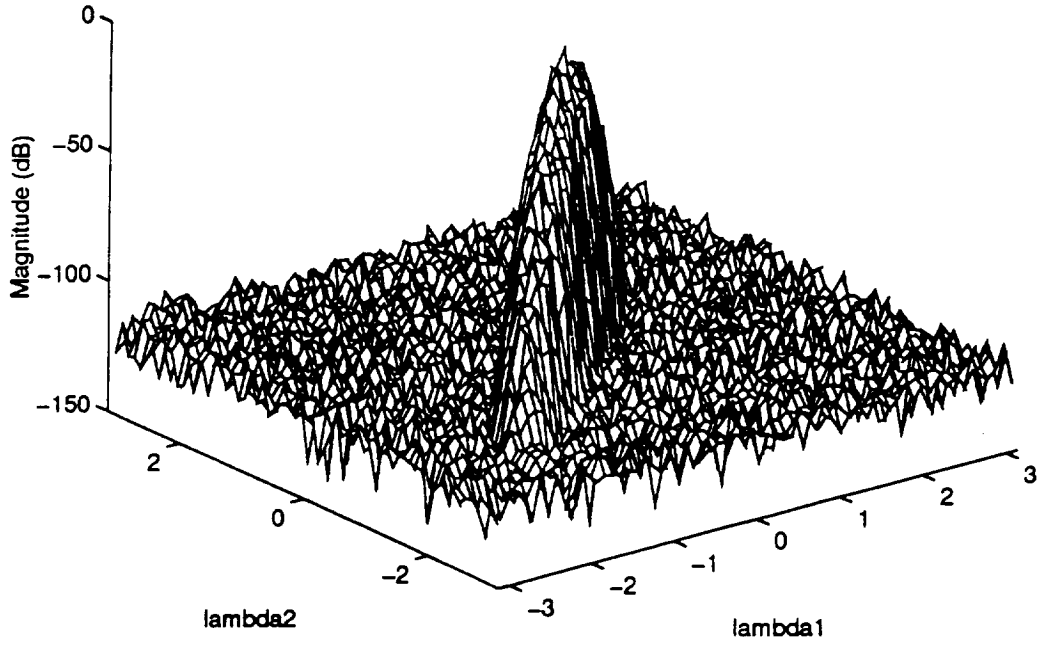

Figure 4.13. Plot of $P(\lambda_1, \lambda_2)$.

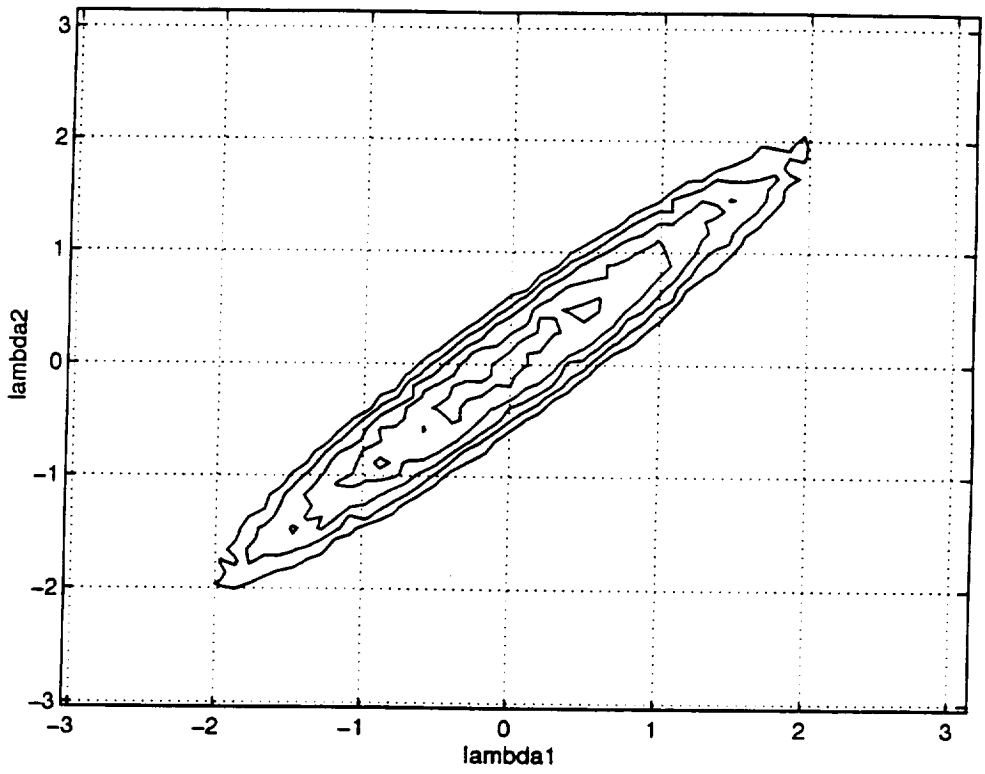

Figure 4.14. Contour plot of $P(\lambda_1, \lambda_2)$.

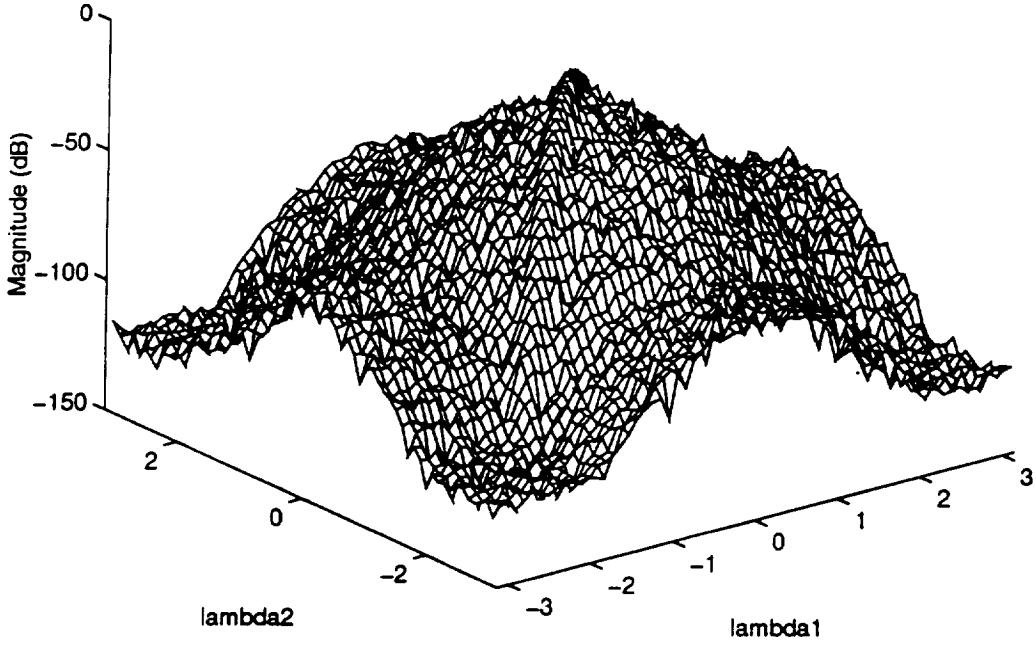

Figure 4.15. Plot of $\hat{f}(\lambda_1, \lambda_2)$.

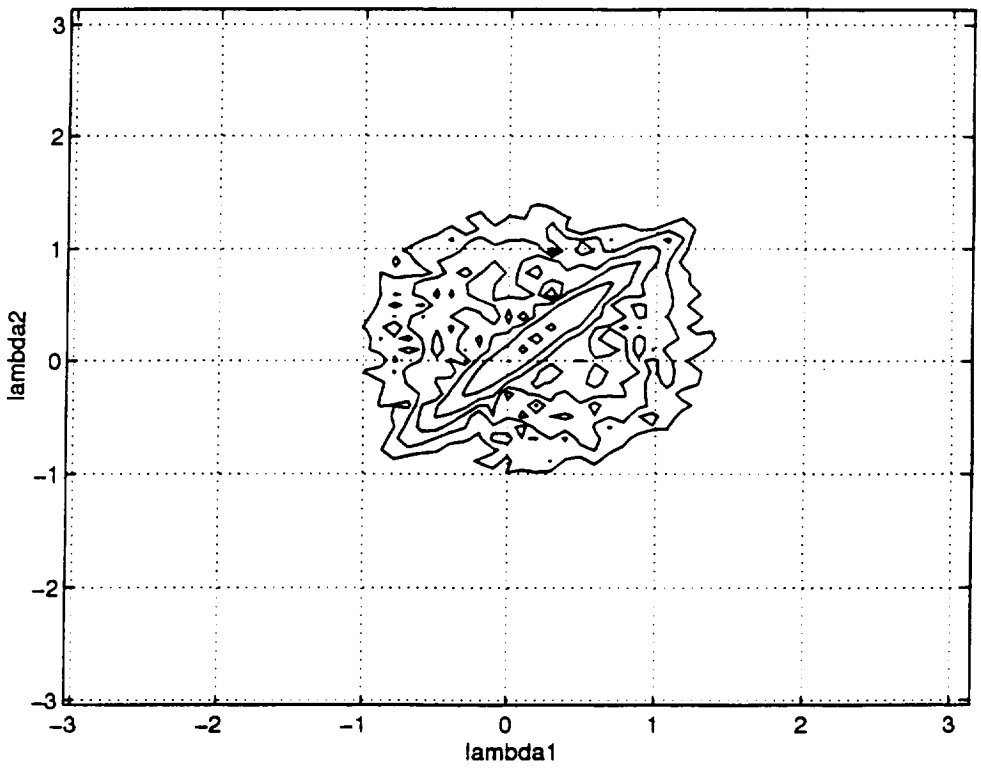

Figure 4.16. Contour plot of $\hat{f}(\lambda_1, \lambda_2)$.

4.9 Summary

In this chapter, marginal sequences are shown to be a class of sequences for which the JSD can be determined. Two drawbacks are immediately apparent. First, for the two sequences to be valid marginals, they must start at the same value. It might be tempting to try to relax this condition by setting the starting values of the two sequences to zero, or the average of the two starting values, or employing some other replacement scheme, but that would violate the condition that the two sequences be individual stationarity. Further research needs to be done applying these sequences to practical applications. The other drawback is that jointly stationary marginal sequences necessarily have identical PSDs. The resulting JSD is a degenerate case. A more general class of marginal sequences can be treated by assuming them to be jointly harmonizable. Then estimating the JSD involves estimating the non-stationary cross-correlation which is quite difficult in practice. In the simulations, it was assumed that an arbitrary number of realizations of the marginal sequences were available for computing ensemble averages.

CHAPTER 5

AN EXPONENTIAL MODEL

An alternative approach to constraining the sequences is to specify a signal model that explicitly incorporates the joint spectral density in its definition. Consider a pair of random sequences $\{X(t), Y(t)\}$ that can be modeled by

$$X(t) = \sum_{i=1}^{K} \alpha_i e^{j(\lambda_i t + \theta_i)} + \xi_t, \tag{5.1}$$

$$Y(t) = \sum_{i=1}^{K} \beta_i e^{j(\nu_i t + \phi_i)} + \zeta_t, \tag{5.2}$$

where $\{\alpha_i\}$ and $\{\beta_i\}$ are constant complex valued amplitudes, $\{\lambda_i\}$ and $\{\nu_i\}$ are random frequencies, $\{\theta_i\}$ and $\{\phi_i\}$ are independent uniform random phases, and ξ_t and ζ_t are zero mean independent white Gaussian noise sequences.

Using this model, a link can be established between the bivariate random sequence and a family of random fields for which the spectral representation is equal to the JSD of $\{X(t), Y(t)\}$. For this purpose, it is necessary to assume that pairs of frequencies $\{\lambda_i, \nu_i\}$ are independent and identically distributed according to a bivariate probability density function, $p_{\lambda\nu}(\lambda, \nu)$, that coincides with the normalized joint spectral density, $f_{\lambda\nu}(\lambda, \nu)$. Let the characteristic function of $p_{\lambda\nu}(\lambda, \nu)$ be denoted by $\Phi_{\lambda\nu}(\tau, \upsilon)$. Marginal functions for $f_{\lambda\nu}(\lambda, \nu)$, $p_{\lambda\nu}(\lambda, \nu)$, and $\Phi_{\lambda\nu}(\tau, \upsilon)$ are written in the usual manner.

A similar model is often encountered in the familiar parameter estimation problem in which devising "good" estimators for the amplitudes, frequencies, or phases is of interest [30]. This model is different in one significant respect. The frequency content in (5.1) and (5.2) is a random variable, and the goal here is not to estimate the frequency (or amplitude, or phase). Rather, it is shown below that if a pair of

random sequences can be modeled by (5.1) and (5.2), then it is possible, under certain conditions, to determine the joint probability density of λ and ν.

5.1 Correlation and Spectral Properties

The first question to answer is whether $X(t)$ and $Y(t)$ are individually and/or jointly stationary. The mean of $X(t)$ is given by

$$\begin{aligned} E[X(t)] &= E\left[\sum_{i=1}^{K} \alpha_i e^{j(\lambda_i t+\theta_i)} + \xi_t\right], \\ &= \sum_{i=1}^{K} \alpha_i E\left[e^{j(\lambda_i t+\theta_i)}\right] + E[\xi_t], \\ &= \sum_{i=1}^{K} \alpha_i E\left[e^{j\lambda_i t}\right] E\left[e^{j\theta_i}\right] \\ &= 0. \end{aligned}$$

The last step results from the fact that $E\left[e^{j\theta_i}\right] = 0$ for θ_i uniformly distributed on $(-\pi, \pi]$. Similarly, it is easy to show that $E[Y(t)] = 0$. The autocorrelation of $X(t)$ is given by

$$\begin{aligned} R_{xx}(t,u) &= E\left[X^*(t)X(u)\right], \\ &= E\left[\left(\sum_{i=1}^{K} \alpha_i^* e^{-j(\lambda_i t+\theta_i)} + \xi_t^*\right)\left(\sum_{k=1}^{K} \alpha_k e^{j(\lambda_k u+\theta_k)} + \xi_u\right)\right], \\ &= \sum_{i=1}^{K}\sum_{i=k}^{K} \alpha_i^*\alpha_k E\left[e^{-j\lambda_i t}e^{j\lambda_k u}\right] E\left[e^{j(\theta_k-\theta_i)}\right] + N_\xi\delta(u-t), \end{aligned}$$

where N_ξ is the mean white noise power for ξ_t. Since the phases are independent and uniformly distributed,

$$E\left[e^{j(\theta_k-\theta_i)}\right] = \begin{cases} 1 & k = i, \\ 0 & \text{otherwise}, \end{cases}$$

and the autocorrelation reduces to

$$R_{xx}(\tau) = \sum_{i=1}^{K} |\alpha_i|^2 E\left[e^{j\lambda_i\tau}\right] + N_\xi\delta(\tau), \tag{5.3}$$

where $\tau = t - s$ and therefore $X(t)$ is wide-sense stationary. Notice that the expectation on the right-hand side of (5.3) is, by definition, the characteristic function of λ_i [4]. Since the $\{\lambda_i\}$ are independent and identically distributed, the characteristic function can be pulled outside the summation and the autocorrelation becomes

$$R_{xx}(\tau) = A\Phi_\lambda(\tau) + N_\xi\delta(\tau),$$

and similarly,

$$R_{yy}(\tau) = B\Phi_\nu(\tau) + N_\zeta\delta(\tau),$$

where $A = \sum_{i=1}^{K} |\alpha_i|^2$, $B = \sum_{i=1}^{K} |\beta_i|^2$, and N_ζ is the white noise power of ζ_t.

The cross-correlation of $X(t)$ and $Y(t)$ is given by

$$\begin{aligned} R_{xy}(t,u) &= E\left[X^*(t)Y(u)\right], \\ &= E\left[\left(\sum_{i=1}^{K} \alpha_i^* e^{-j(\lambda_i t+\theta_i)} + \xi_t^*\right)\left(\sum_{k=1}^{K} \beta_k e^{j(\nu_k u+\phi_k)} + \zeta_u\right)\right], \\ &= \sum_{i=1}^{K}\sum_{k=1}^{K} \alpha_i^*\beta_k E\left[e^{-j\lambda_i t}e^{j\nu_k u}\right] E\left[e^{j(\phi_k-\theta_i)}\right]. \end{aligned} \tag{5.4}$$

If θ_i and ϕ_k are uncorrelated for all i, k then the cross-correlation is zero. If, on the other hand, θ_i and ϕ_k are allowed to be correlated for $i = k$, then the cross-correlation will be scaled by a factor determined by the joint probability density of θ_i and ϕ_i. For example, if the phases are jointly distributed according to Morgenstern's bivariate uniform (see Appendix A), then

$$E\left[e^{j(\phi_k-\theta_i)}\right] = \begin{cases} \frac{\kappa}{\pi^2} & k = i, \\ 0 & \text{otherwise.} \end{cases}$$

For our purposes, the constant scale factor can by assumed to be unity without loss of generality. This assumption corresponds to setting $\theta_i = \phi_i$.

Substituting back into (5.4) gives

$$R_{xy}(t,u) = \sum_{i=1}^{K} \alpha_i^*\beta_i E\left[e^{j(-\lambda_i t+\nu_i u)}\right].$$

Note that in general, the sequences are not jointly stationary for $\lambda_i \neq \nu_i$. The expectation on the right-hand side can be recognized as the joint characteristic function of λ_i and ν_i. Since the pairs $\{\lambda_i, \nu_i\}$ are independent and identically distributed, the characteristic function can be pulled outside the summation to give

$$R_{xy}(t,u) = \left(\sum_{i=1}^{K} \alpha_i^* \beta_i\right) \Phi_{\lambda\nu}(-t,u).$$

If $\Phi_{\lambda\nu}(-t,u) = \Phi_{\lambda\nu}(u-t)$ then the sequences are jointly stationary.

It is remarkable that the cross-correlation may or may not contain information about the joint probability structure of λ and ν depending on a fairly subtle detail of the signal model. Thus it is difficult to make general statements about a relationship between the correlation between sequences in the time domain and the correlation in their frequency content. As in the case of marginal sequences, if $X(t)$ and $Y(t)$ are jointly stationary, then the joint spectral density is concentrated along $\lambda = \nu$. In order to treat a more general class of problems the non-stationary cross-correlation must be used.

Taking the discrete time Fourier transform of $R_{xx}(\tau)$, $R_{yy}(\tau)$, and $R_{xy}(t,u)$ gives the auto-spectra as

$$f_{xx}(\lambda) = A f_\lambda(\lambda) + \frac{N_\xi}{2\pi},$$
$$f_{yy}(\lambda) = B f_\nu(\nu) + \frac{N_\zeta}{2\pi},$$

and the cross-spectral density as

$$f_{xy}(\lambda,\nu) = \left(\sum_{i=1}^{K} \alpha_i^* \beta_i\right) f_{\lambda\nu}(-\lambda,\nu).$$

Considering that the sequences are composed of exponentials with additive white noise, a question arises of whether spectra are discrete or continuous. Since the frequency content is a random variable, there is no reason to expect that the λ_i's (or ν_i's) will be harmonically related and thus the sequences will not be periodic. Nevertheless, they are "almost" periodic and taken at face value, a single realization of $X(t)$ (or

$Y(t)$) would be expected to have a discrete spectrum [1]. However, in the expected value, the true spectrum is continuous and equal to the probability density function of the frequency. This is a model for which the sample spectrum can be significantly different from the true spectrum. For small K, the sample spectrum consists of a flat noise component with peaks located at the frequencies of the exponentials. Most likely, this will bear little resemblance to the true spectrum. Therefore, in order for the sample auto-spectra to look anything like the true spectrum, K must be large.

5.2 Simulations

In the following simulations, the joint spectral density is bivariate Gaussian with zero mean and equal variances so that

$$f_{\lambda\nu}(\lambda,\nu) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2\sigma^2(1-\rho^2)}\left(\lambda^2 - 2\rho\lambda\nu + \nu^2\right)\right],$$

and the marginal spectra are given by

$$f_\lambda(\lambda) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\lambda^2}{2\sigma^2}},$$
$$f_\nu(\nu) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\nu^2}{2\sigma^2}}.$$

Figure 5.1 shows the power spectral density for $K = 256$ exponentials, and an SNR of 50dB. A roughly Gaussian shaped can be distinguished centered at $\lambda = \frac{\pi}{3}$.

In the next two subsections, estimates of the joint spectral density, $\hat{f}_{\lambda\nu}(\lambda,\nu)$, are based on estimates of the non-stationary cross-correlation computed as,

$$\hat{R}_{xy}(t,u) = \frac{1}{N}\sum_{i}^{N} X_i^*(t) Y_i(u),$$

where N is the number of realizations in the ensemble average. Taking the two-dimensional FFT of $\hat{R}_{xy}(-t,u)$ gives $\hat{f}_{\lambda\nu}(\lambda,\nu)$. The resulting estimates are displayed as contour plots so that correlations between λ and ν will be most visible.

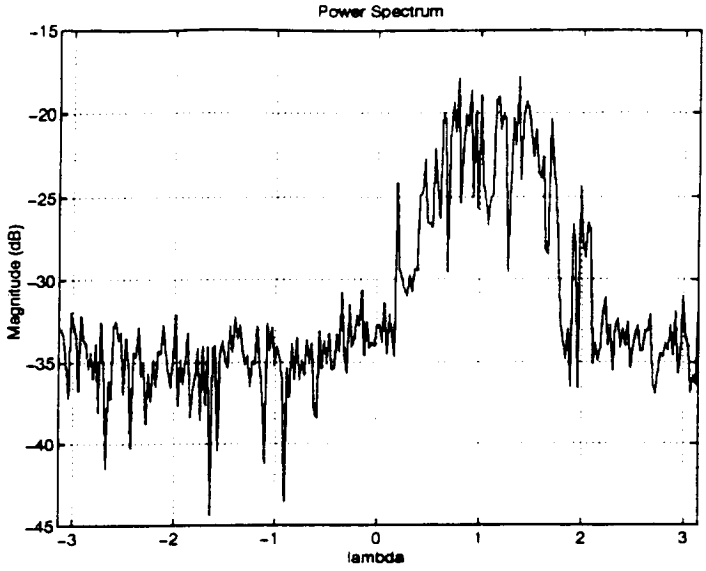

Figure 5.1. Power spectral density of $X(t)$ for K=256.

5.2.1 Single Exponential ($K = 1$)

The simplest case is when $X(t)$ and $Y(t)$ are the sum of a single exponential and a white random noise sequence and the amplitudes are set to unity.

$$X(t) = e^{j(\lambda t+\theta)} + \xi_t,$$

$$Y(t) = e^{j(\nu t+\phi)} + \zeta_t.$$

The the cross-correlation is given by

$$R_{xy}(s,t) = E\left[e^{j(\nu t-\lambda s)}\right]\ E\left[e^{j(\phi-\theta)}\right].$$

If $\theta = \phi$ then the phases exactly cancel, leaving

$$R_{xy}(s,t) = \Phi_{\lambda\nu}(-s,t).$$

Figures 5.2 thru 5.4 show contour plots of $\hat{f}_{\lambda\nu}(\lambda,\nu)$ for $\rho = 0$, 0.5, and 0.9 respectively. Ensemble averages are over 1,000 realizations in each case. What should be noticed in these plots is that as ρ increases, the plot become more prolate along $\lambda = \nu$.

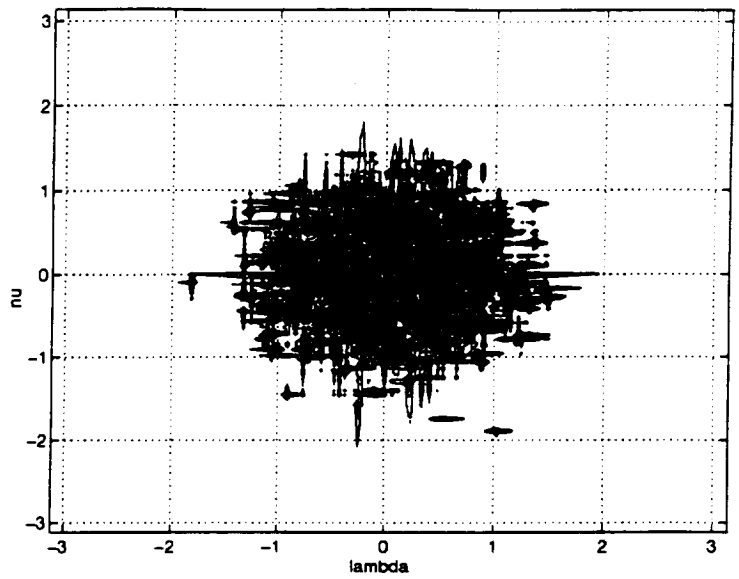

Figure 5.2. Contour plot of $\hat{f}_{\lambda\nu}(\lambda, \nu)$ $(K = 1, \rho = 0)$.

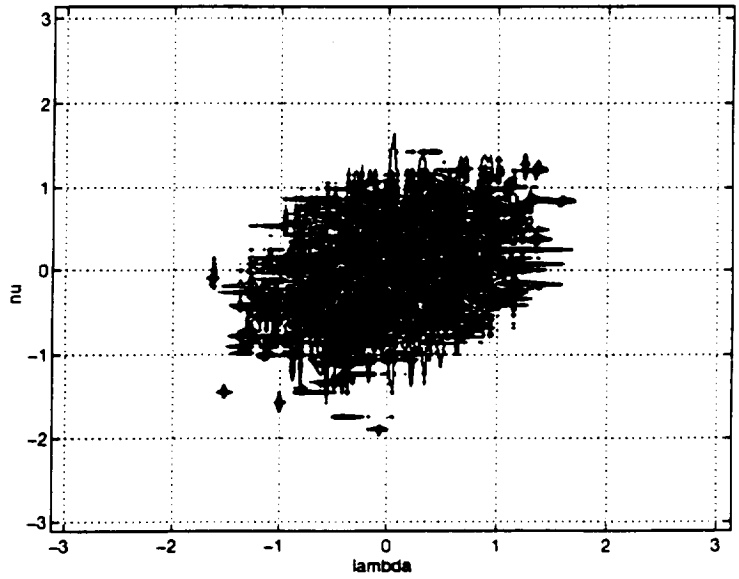

Figure 5.3. Contour plot of $\hat{f}_{\lambda\nu}(\lambda, \nu)$ $(K = 1, \rho = .5)$.

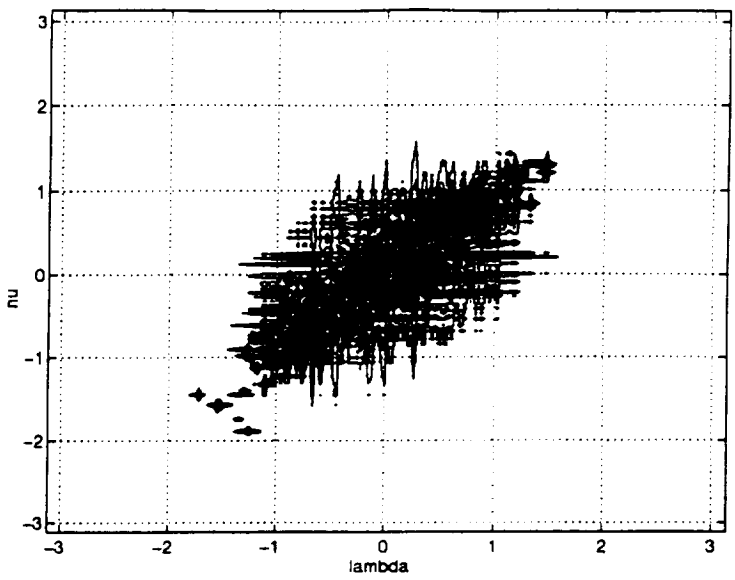

Figure 5.4. Contour plot of $\hat{f}_{\lambda\nu}(\lambda, \nu)$ $(K = 1, \rho = .9)$.

5.2.2 Multiple Exponentials $(K > 1)$

With more than one exponential in the sum, the situation becomes a bit more complicated. Consider the case of two exponentials.

$$X(t) = e^{j(\lambda_1 t + \theta_1)} + e^{j(\lambda_2 t + \theta_2)} + \xi_t,$$

$$Y(t) = e^{j(\nu_1 t + \phi_1)} + e^{j(\nu_2 t + \phi_2)} + \zeta_t.$$

The cross-correlation is given by

$$\begin{aligned} R_{xy}(s,t) &= E\left[\left(e^{-j(\lambda_1 s + \theta_1)} + e^{-j(\lambda_2 s + \theta_2)} + \xi_s\right)\left(e^{j(\nu_1 t + \phi_1)} + e^{j(\nu_2 t + \phi_2)} + \zeta_t\right)\right], \\ &= E\left[e^{j(\nu_1 t - \lambda_1 s)}\right] + E\left[e^{j(\nu_2 t - \lambda_1 s)}\right]\ E\left[e^{j(\theta_2 - \theta_1)}\right] + \\ &\quad E\left[e^{j(\nu_1 t - \lambda_2 s)}\right]\ E\left[e^{j(\theta_1 - \theta_2)}\right] + E\left[e^{j(\nu_2 t - \lambda_2 s)}\right] + \text{other terms.} \end{aligned}$$

The "other terms" include all the product terms with ξ_s and ζ_t. In the expected value, all the terms of the summation go to zero except $E\left[e^{j(\nu_1 t - \lambda_1 s)}\right]$ and $E\left[e^{j(\nu_2 t - \lambda_2 s)}\right]$. but when estimating the cross-correlation from a finite number of realizations, these

"vanishing" terms can contribute a significant error. The more exponentials that are included in the model (the higher K is) the more of these cross-product terms appear in the estimate. The upshot is that as K is increased, more realizations are needed to obtain a reasonable estimate of $R_{xy}(s,t)$. Figure 5.5 shows the contour plot of $\hat{f}_{\lambda\nu}(\lambda,\nu)$ for $K = 2$ and $\rho = 0.9$ averaged over 1,000 realizations. Notice that in Figure 5.5 the contours are not as prolate as they were for the single exponential case shown in Figure 5.4. Many other situations could be considered here, but this is not intended to be an exhaustive study of the estimation issues associated with the joint spectral density. The insights to be gained from these simulations are that

1. For this model, it is possible to get the joint spectral density from the non-stationary cross-correlation.
2. A large number of realizations are needed for meaningful estimates.

5.3 Summary

The exponential model proposed in this chapter, provides an alternative example of how a bivariate random sequence can be constrained so that its joint spectral density can be determined. If the phases, θ_i and ϕ_i, are uncorrelated then the cross-correlation is zero, and can reveal nothing about the joint probability structure of λ and ν. If the phases are correlated, then the cross-correlation is related to the joint characteristic function of λ and ν. Except for some special cases, the cross-correlation is non-stationary. When $\lambda_i = \nu_i$, the joint spectral density is concentrated along the diagonal just as was the case for jointly stationary marginal sequences in Chapter 4. In general, simulations showed that the shape of a joint spectral density estimate based on the cross-cross correlation, averaged over a sufficient number of realizations, does reflect the degree of correlation between λ and ν.

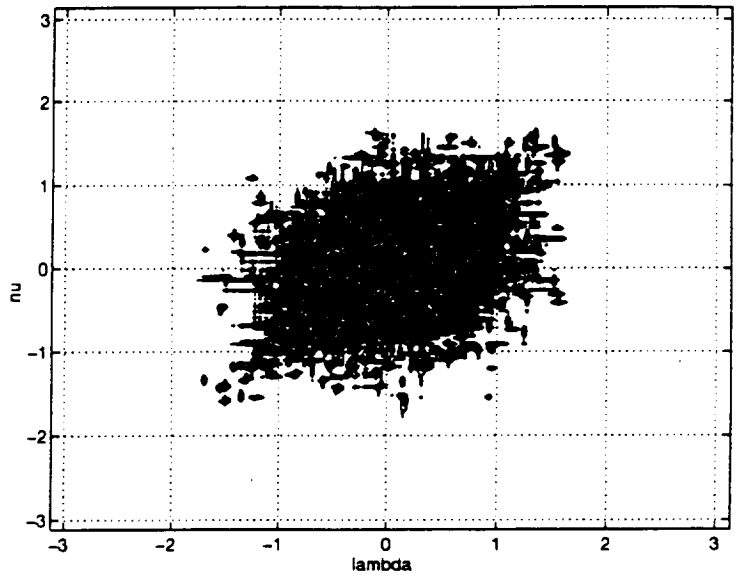

Figure 5.5. Contour plot of $\hat{f}_{\lambda\nu}(\lambda, \nu)$ $(N = 1000, K = 2, \rho = .9)$.

CHAPTER 6

CONCLUSIONS

6.1 Motivation

The original motivation for pursuing the definition of a joint spectral density was rooted in a problem from the radar signal processing of weather. In this application, random sequences represent the backscatter returns from range cells that extend radially from the radar antenna. The frequency content of each random sequence is related to windspeed in its corresponding range cell through the Doppler shift principle. For detecting hazardous weather conditions with a pulsed Doppler radar, the windspeed gradient is often a quantity of particular interest. When estimating a windspeed gradient based hazard index, the ability to also estimate the probability density of that index would be of considerable value. It is in this context that the need for the joint spectral density arose. If, through the joint spectral density, a joint probability density of velocity in adjacent range cells can be estimated, then random variable transformations can be used to get a PDF for windspeed gradient or elementary functions of the windspeed gradient. This method could be implemented either parametrically or non-parametrically. If prior knowledge of the form of the joint PDF of velocity is available, then its parameters could be estimated from the joint spectral density. Otherwise, the joint spectral density could be used directly. The contribution of this dissertation has been to develop the theoretical framework necessary to define and apply the joint spectral density.

6.2 Properties of the Joint Spectral Density

In Chapter 3, the joint spectral density is defined in terms of the two-dimensional spectral density of a stationary random field that has marginal spectra corresponding to the power spectral densities of the two sequences. There are two stages of "many-to-

one" relationships operating here. At the first stage, many different random sequences can have the same power spectral density. At the second stage, once a two-dimensional spectral density has been specified, there are many random fields that can have that same power spectral density function. Therefore it is possible to think of the joint spectral density as forming a bridge between families of random sequences and families of random fields. The spectral theory of multivariate random sequences and the spectral theory of random fields are both well established. The importance of the joint spectral density is that it pulls these two bodies of theory together in such a way that it is possible to talk about the joint probability density function of the frequency content of multiple sequences. Therefore all of the familiar techniques for interpreting and manipulating probability density functions can be applied to the analysis of the frequency content of multiple random sequences.

The projection-slice theorem provides a relationship between slices in the two-dimensional autocorrelation of this random field, $R(\tau, \upsilon)$, and projections in the joint spectral density. For spectral density functions with elliptical symmetry such as the bivariate Gaussian, knowing $R(\tau, 0)$, $R(0, \tau)$, and $R(\tau, \tau)$ is sufficient to determine the joint spectral density everywhere. The margins of $R(\tau, \upsilon)$ are recognized to be the autocorrelations of the bivariate random sequence, but further research needs to be done to determine if there are any conditions under which $R(\tau, \tau)$ can be computed from those sequences.

In the absence of any constraints on the random sequences, the joint spectral density is underdetermined. In other words, a gap exits between the conventional functions of bivariate spectral analysis, e.g. the auto- and cross-spectral densities, and the joint spectral density. In Chapters 4 and 5, two types of sequences are identified for which the JSD be determined by realizations of a bivariate random sequence.

6.3 Marginal Sequences

In Chapter 4 it is shown that if the bivariate random sequences are the marginal sequences of a stationary random field, then the JSD can be written in terms of the cross-spectral density. When the sequences are jointly stationary, it is shown that the autospectra of the two sequences are necessarily equal to each other and the JSD is concentrated along the diagonal and has the same shape as the autospectra. This is considered to be a degenerate case. By relaxing the condition of joint stationarity to require only that the sequences be jointly harmonizable, a more general class of sequences can be treated.

Examples are provided of the JSDs for the jointly stationary and jointly harmonizable cases. A simulation method is presented for generating marginal sequences that correspond to an arbitrary joint spectral density specified by the user. For a jointly stationary and a jointly harmonizable case, realizations of the simulated marginal sequences were used to estimate the JSD. Plots of the true and estimated JSD are included for comparison. It is clear from these plots that the estimated JSD does not closely resemble the true JSD. However, the effects of a correlation between the two frequency variables is visible so the potential for making parameterized estimates of the JSD is promising.

6.4 Exponential Model

In Chapter 5 the sequences are assumed to be modeled by a sum of complex exponentials and additive white Gaussian noise. For a particular sequence, the frequency of each exponential in the sum is an independent realization of a random variable. The random variables that generate the frequencies in the two sequences are distributed according to a joint probability density function that corresponds to the joint spectral density. If the phases are independent and uniformly distributed on $(0, 2\pi]$, the sequences are shown to be individually stationary and with autocorrelations proportional to the marginal characteristic function of the frequency. The

cross-correlation is found to be zero if the phases in one sequence are uncorrelated with the phases in the other sequence. If the phases are correlated, then the cross-correlation is proportional to a reflection of the joint characteristic function of the frequencies.

Simulations are presented for the joint spectral density distributed as a bivariate Gaussian. For the case of a single exponential, simulations were run for $\rho = 0$, 0.5, and 0.9 and contour plots of the estimated joint spectral density reflect the degree of correlation between the frequency content of the two sequences. As the number of exponentials is increased, it was found that more realizations need to be ensemble averaged to mitigate the effects of cross-product terms in the cross-correlation. These cross-product terms go to zero in the expected value, but when the expected value is estimated by averaging over a finite number of realizations they can introduce a significant error.

6.5 Future Work

For the marginal sequences and the exponential model, the joint spectral density is equal to the non-stationary cross-spectral density reflected about one of the frequency variables. As such, estimation of the cross-correlation is a primary concern. The fact that the cross-correlation is in general non-stationary complicates the estimation problem considerably. Since ergodicity does not hold, time averages can not be substituted for ensemble averages. This means that, except in certain special cases, multiple realizations are needed to get meaningful estimates. A potential remedy to this estimation problem is the concept of local stationarity which offers the possibility of estimating the non-stationary cross-correlation from a single realization [29, 31]. The cross-correlation is called locally stationary if it can be factored into a stationary part and a non-stationary part by introducing the substitution

$$t_1 = t - \frac{\tau}{2}, \quad t_2 = t + \frac{\tau}{2},$$

into $Rxy(t_1, t_2)$. Then the cross-correlation can be written as

$$Rxy(t_1, t_2) = C\Gamma_1\left(\frac{t_1 + t_2}{2}\right)\Gamma_2(t_2 - t_1),$$

where C is a constant, $\Gamma_1(t)$ is the instantaneous power, and $\Gamma_2(\tau)$ is a stationary sequence. If $\Gamma_1(t)$ is slowly varying with respect to $\Gamma_2(\tau)$ then they can be estimated separately [29]. This and other implementation issues should be a fruitful area for future research.

6.6 Final Assessments

In this dissertation a theoretical framework has been developed for considering the joint spectral density of bivariate random sequences. A significant contribution has been made by defining and exploring the properties of the joint spectral density in terms of a link between the spectral representation of bivariate random sequences and the spectral representation of stationary random fields. Since the joint spectral density was found, in general, to be underdetermined with respect to the two sequences, there is a need to devise constraints on the sequences that allow the JSD to be determined. Two such constraints are investigated here, but further research is called for to identify other classes of signals that yield unique JSDs.

APPENDICES

Appendix A

Families of Bivariate Uniform Distributions

There is more than one solution to the problem of determining a bivariate density function that yields uniform marginal densities. Two one-parameter families of bivariate distributions that satisfy the consistency requirement for the marginals are the Morgenstern's uniform distribution and Plackett's uniform distribution [26].

A.1 Morgenstern's Uniform Distribution

Morgenstern proposed a family of distribution functions given by

$$H(x,y) = F(x)G(y)\{1 + \kappa[1 - F(x)][1 - G(y)]\}, \quad -1 \leq \kappa \leq 1. \tag{A.1}$$

The corresponding density function is given by

$$h(x,y) = f(x)g(y)\{1 + \kappa[2F(x) - 1][2G(y) - 1]\}, \quad -1 \leq \kappa \leq 1. \tag{A.2}$$

If $F(x)$ and $G(y)$ are uniform distributions on $[0,1]$, then Morgenstern's uniform distribution is given by

$$h(x,y) = [1 + \kappa(2x - 1)(2y - 1)], \quad -1 \leq \kappa \leq 1. \tag{A.3}$$

Figures A-1 thru A-5 show $h(x,y)$ for various values of κ.

A.2 Plackett's Uniform Distribution

The density function for Plackett's bivariate uniform distribution is given by

$$h(x,y) = \frac{\psi[(\psi - 1)(x + y - 2xy) + 1]}{\{[1 + (\psi - 1)(x + y)]^2 - 4\psi(\psi - 1)xy\}^{\frac{3}{2}}}, \quad \psi \geq 0. \tag{A.4}$$

Figures A-6 thru A-10 show $h(x,y)$ for various values of ψ.

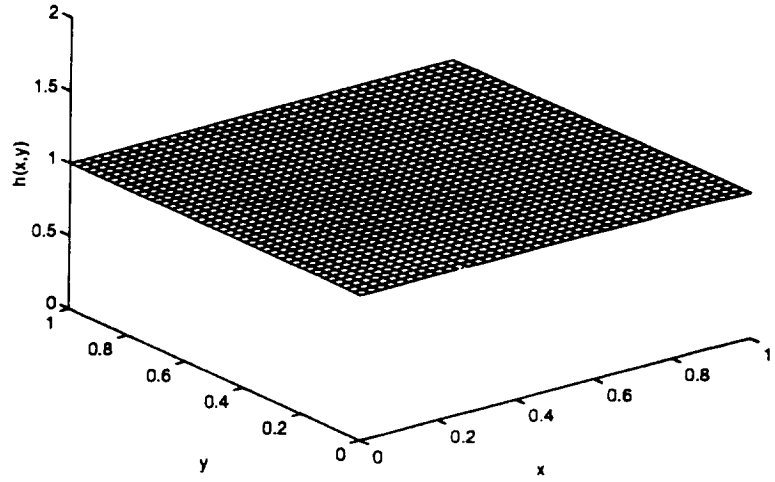

Figure A-1. Morgenstern's uniform density for $\kappa = 0$.

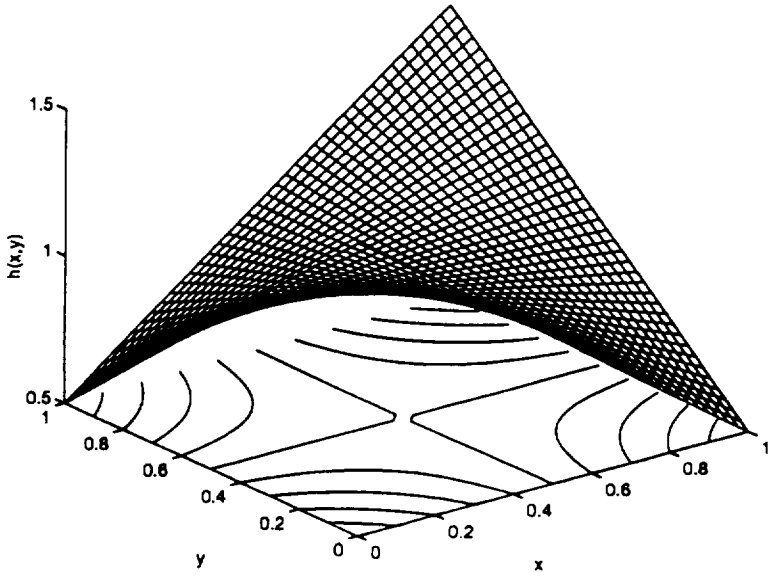

Figure A-2. Morgenstern's uniform density for $\kappa = 0.5$.

C-2.

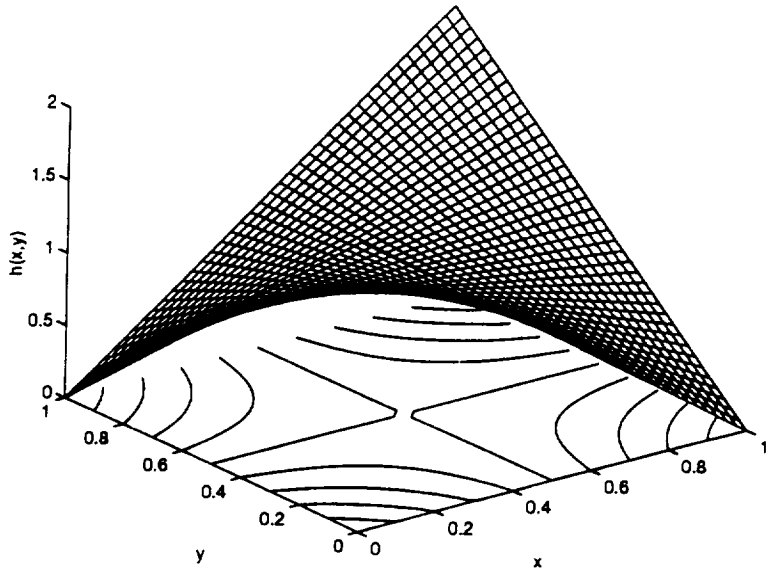

Figure A-3. Morgenstern's uniform density for $\kappa = 1$.

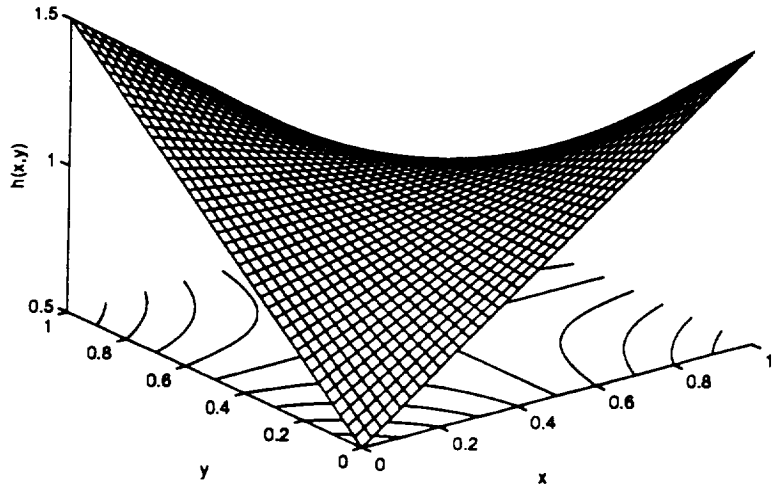

Figure A-4. Morgenstern's uniform density for $\kappa = -0.5$.

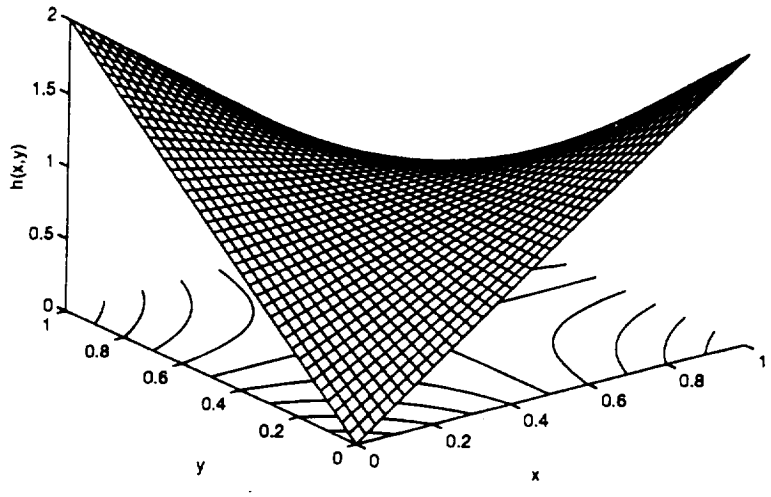

Figure A-5. Morgenstern's uniform density for $\kappa = -1$.

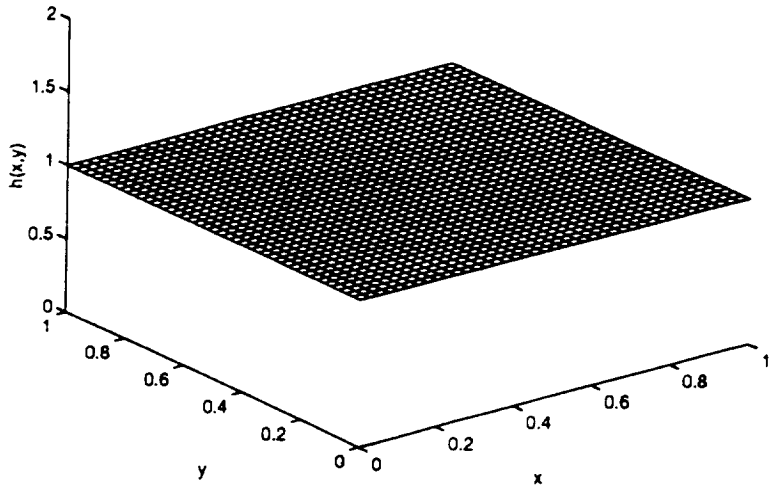

Figure A-6. Plackett's uniform density for $\psi = 0$.

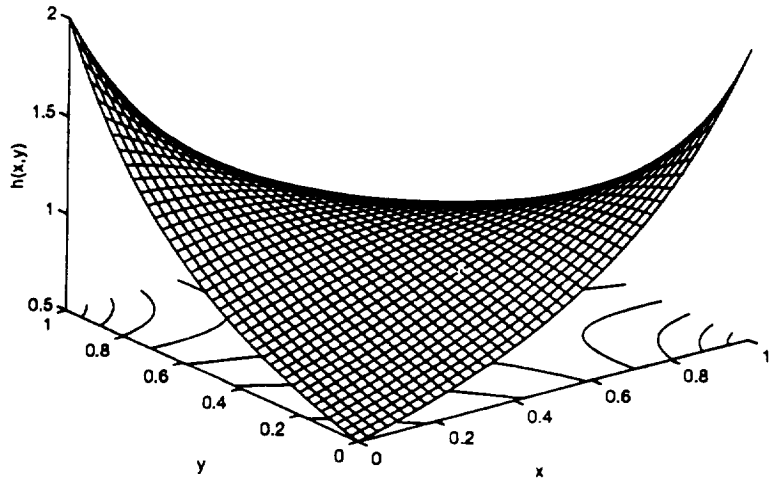

Figure A-7. Plackett's uniform density for $\psi = 0.5$.

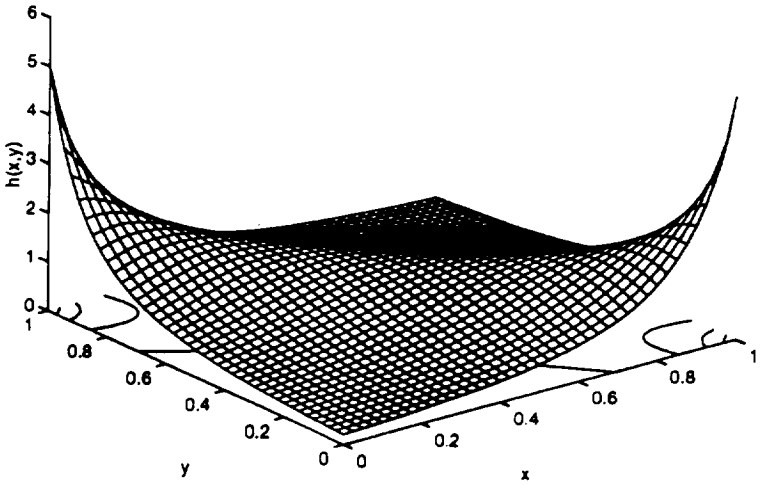

Figure A-8. Plackett's uniform density for $\psi = 0.2$.

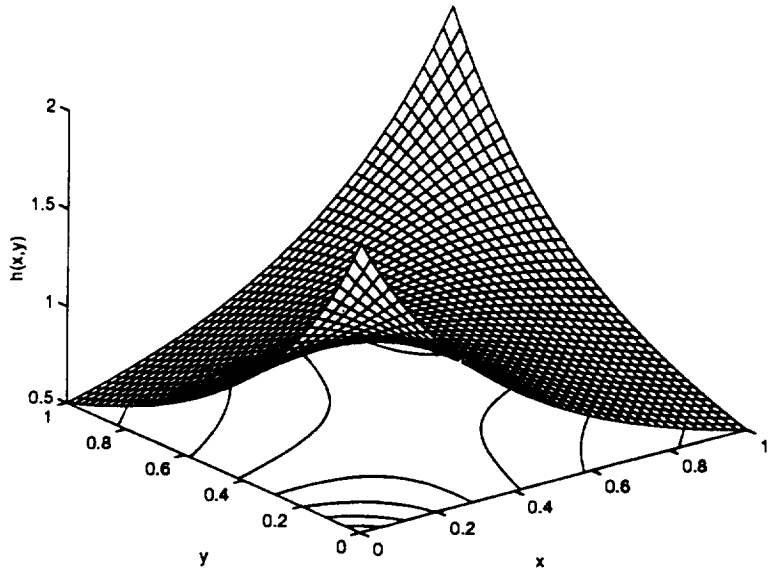

Figure A-9. Plackett's uniform density for $\psi = 2$.

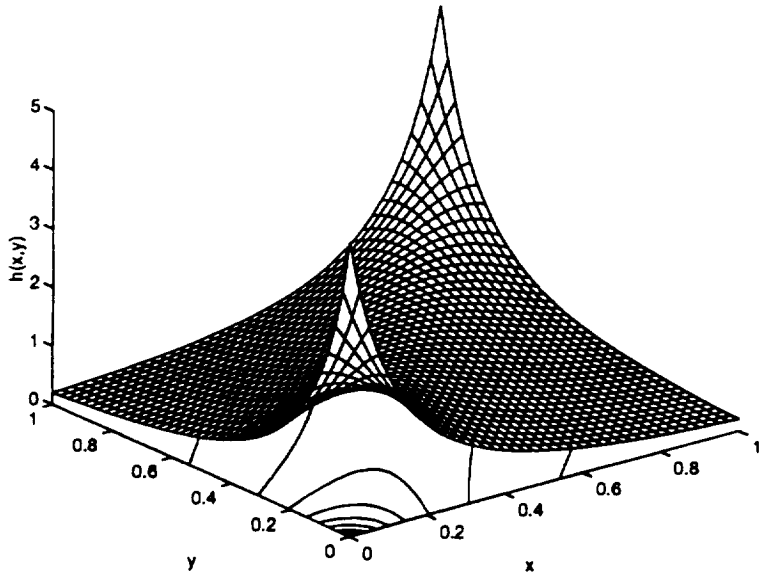

Figure A-10. Plackett's uniform density for $\psi = 5$.

Appendix B

Some Properties of the Gaussian Distribution

B.1 Univariate Gaussian

A standard univariate Gaussian is given by

$$G(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \tag{B.1}$$

and the distribution function is given by

$$F_X(x) = P(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}}\, dt, \tag{B.2}$$

where the integral on the right hand side can not be evaluated in closed form. Instead it is often tabulated in terms of the area under the tail of the Gaussian, the so called Q function [32]

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{\frac{-t^2}{2}}\, dt, \tag{B.3}$$

or in terms of the error function

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{\frac{-t^2}{2}}\, dt. \tag{B.4}$$

The relationships among $F(x), Q(x)$, and $\text{erf}(x)$ for a standard zero mean, unit variance Gaussian function are as follows:

$$F(x) + Q(x) = 1, \qquad F(x) = Q(-x), \qquad \text{erf}(x) = 2F(\sqrt{2}x) - 1.$$

For a Gaussian random variable y with mean μ and variance σ^2, the distribution function, $F(y)$, can be found from the Q and erf functions replacing x with $\frac{x-\mu}{\sigma}$. This gives

$$F(y) = Q\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{2}\left(1 + \text{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)\right). \tag{B.5}$$

B.2 Bivariate Gaussian

A standard bivariate Gaussian is given by

$$g(x,y,\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(x^2 - 2\rho xy + y^2\right)\right] \tag{B.6}$$

and can be written in terms of Z as

$$g(x,y,\rho) = \frac{1}{\sqrt{1-\rho^2}} Z(x) Z\left(\frac{y-\rho z}{\sqrt{1-\rho^2}}\right). \tag{B.7}$$

Figures B-1 thru B-3 show plots of the bivariate Gaussian for $\rho = 0$, $\rho < 0$, and $\rho > 0$ respectively. As in the univariate case, the integral of the bivariate Gaussian is tabulated in terms of the area under the tail called the L function [32]

$$L(h,k,\rho) = \int_h^\infty \int_k^\infty g(x,y,\rho)\, dx\, dy = \int_h^\infty Z(x)\, dx \int_k^\infty Z(w)\, dw \tag{B.8}$$

where

$$w = \left(\frac{k-\rho x}{\sqrt{1-\rho^2}}\right).$$

Some useful relationships among F, L, and Q are

$$F(h,k;\rho) = F(k,h;\rho) = L(-h,-k,\rho) = L(-k,-h,\rho) \tag{B.9}$$

$$= L(h,k,\rho) + Q(h) + Q(k) - 1 \tag{B.10}$$

$$F(-h,k,\rho) = 1 - Q(h) - L(h,k,-\rho) \tag{B.11}$$

$$F(h,-k,\rho) = 1 - Q(k) - L(h,k,-\rho). \tag{B.12}$$

B.3 Symmetry Properties of the Bivariate Gaussian

If a bivariate time series has a Gaussian shaped joint power spectral density, then the characteristic function is of the form

$$\Phi(\tau_1,\tau_2) = e^{-\frac{1}{2}(a^2\tau_1^2 + 2\rho ab\tau_1\tau_2 + b^2\tau_2^2)} \tag{B.13}$$

which has elliptical symmetry resulting from the

$$a^2x^2 + 2rabxy + b^2y^2 \tag{B.14}$$

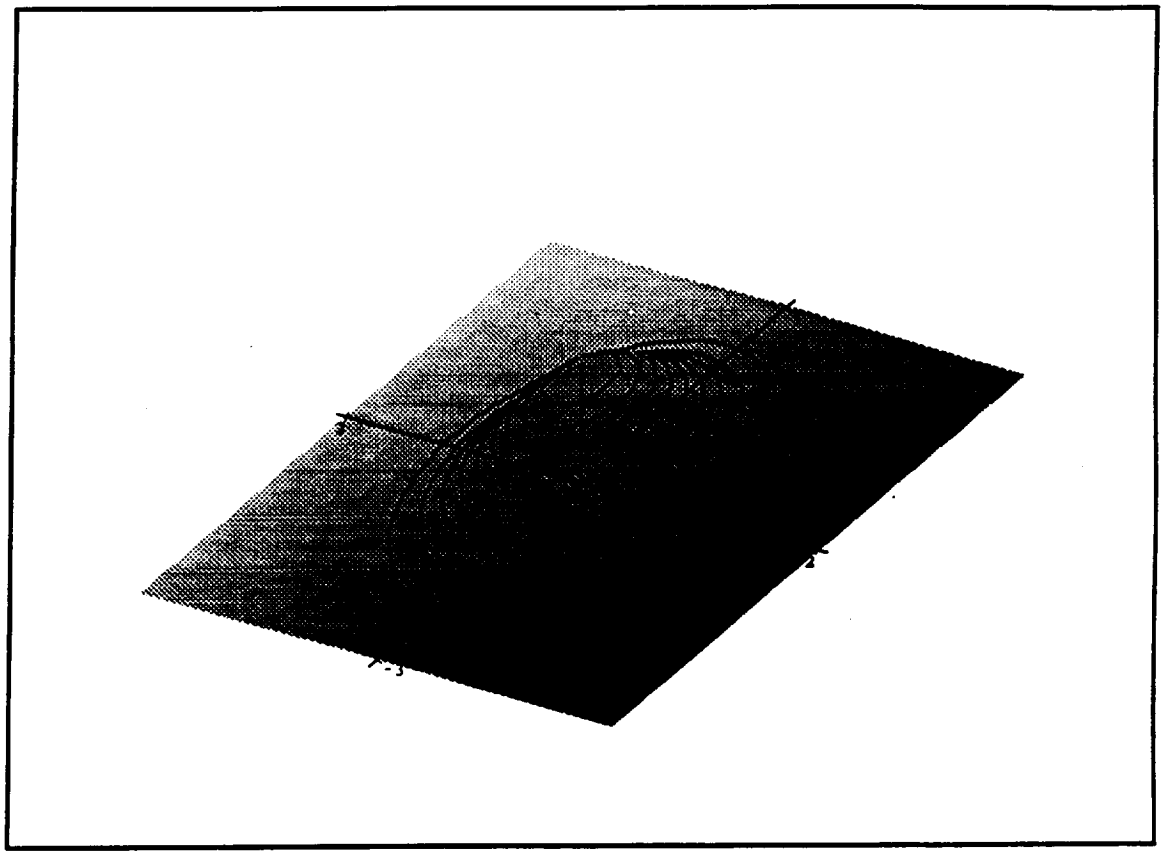

Figure B-1. 2-dimensional Gaussian with $\rho = 0$

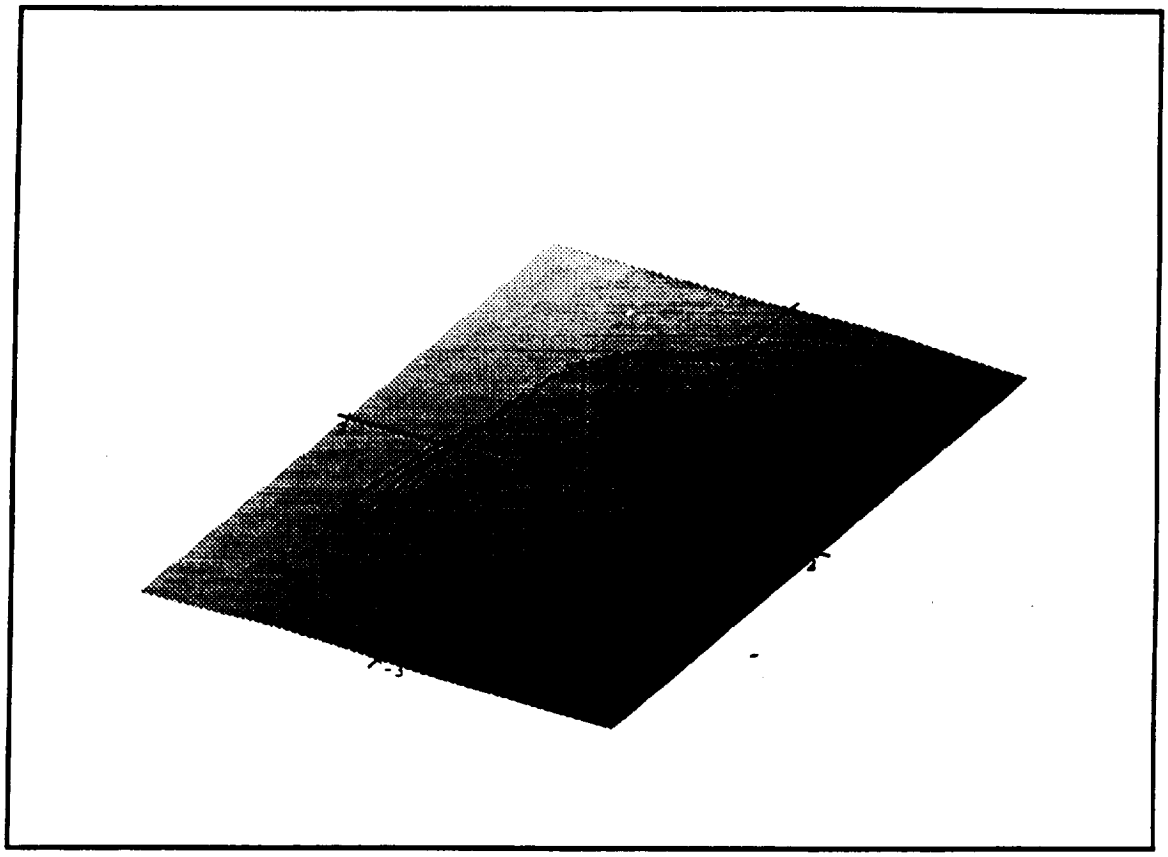

Figure B-2. 2-dimensional Gaussian with $\rho > 0$

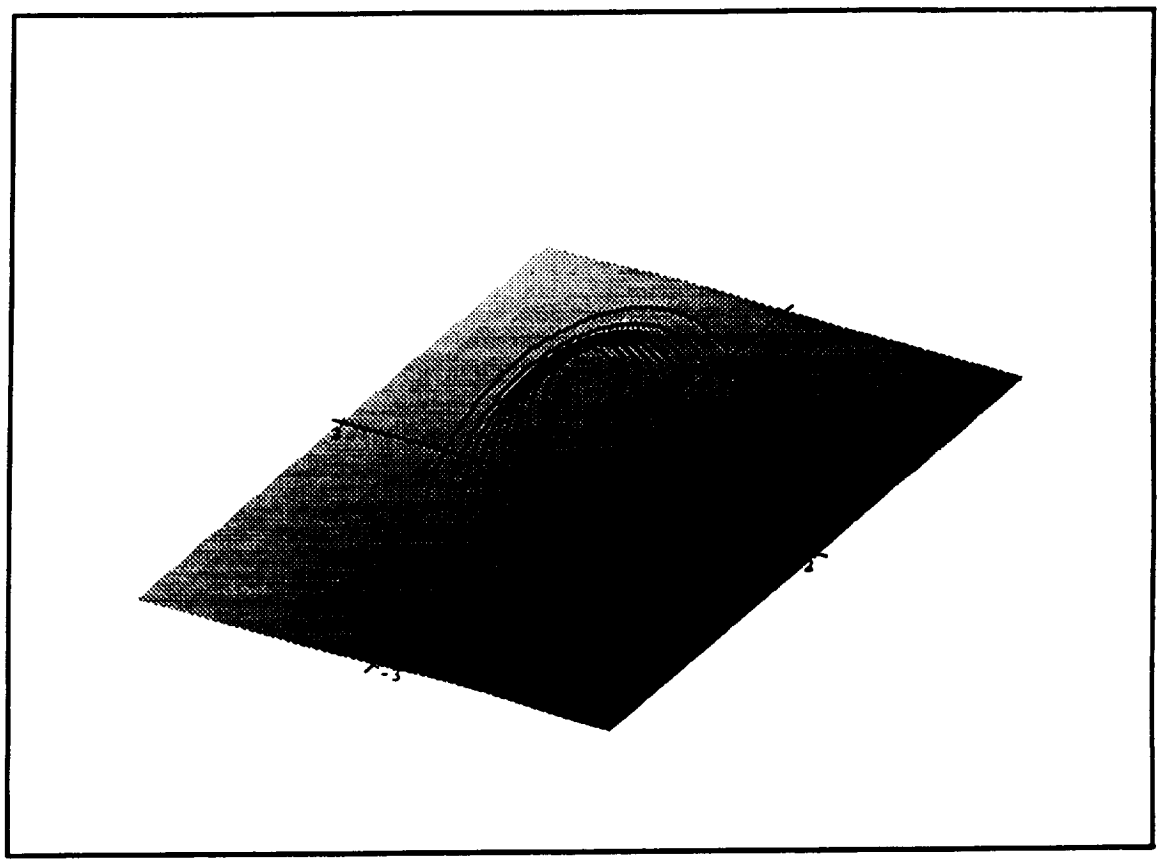

Figure B-3. 2-dimensional Gaussian with $\rho < 0$

form of the argument of the exponential. If $r = 0$ then (B.14) reduces to

$$\frac{x^2}{b^2} + \frac{y^2}{a^2} = c^2 \tag{B.15}$$

which describes a family of ellipses for which the major and minor axes coincide with the coordinate axes (see Figure B-4). A non-zero value for r corresponds to a rotation of axes by an angle of α, where α is given by

$$\cot \alpha = \frac{a^2 - b^2}{2rab}, \tag{B.16}$$

as shown in Figure B-5.

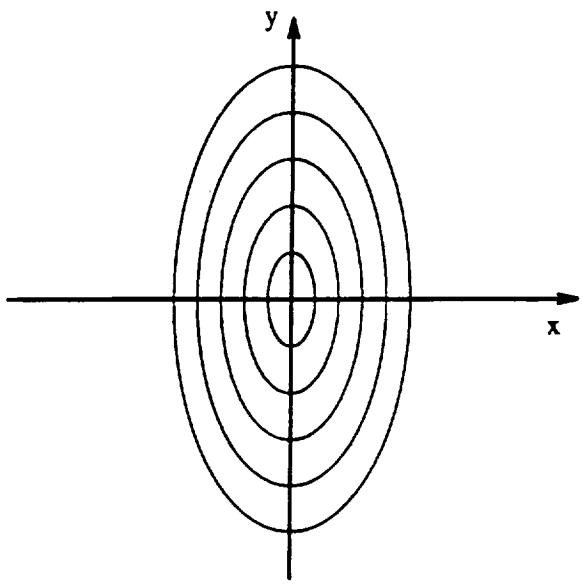

Figure B-4. Family of ellipses for $r = 0$.

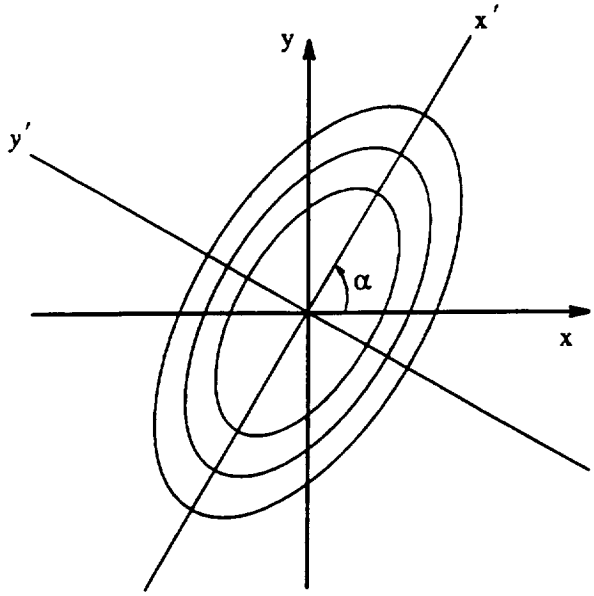

Figure B-5. Family of ellipses for $r \neq 0$.

REFERENCES

1. M. B. Priestley. *Spectral Analysis and Time Series.* Academic Press, New York, 1981.

2. R. J. Doviak and D. S. Zrnić. *Doppler Radar and Weather Observations.* Academic Press, Orlando, FL, 1984.

3. K. S. Miller and M. M. Rochwarger. "A Covariance Approach to Spectral Moment Estimation". *IEEE Transactions on Information Theory*, 18:588–596, 1972.

4. A. Papoulis. *Probability, Random Variables, and Stochastic Processes.* second edition, McGraw-Hill, New York, 1984.

5. D. S. Zrnić. "Estimation of Spectral Moments for Weather Echoes". *IEEE Trans. Geosci. Electron.*, GE-17(4):113–128, October 1979.

6. R. L. Bowles and R. Targ. "Windshear Detection and Avoidance: Airborne Systems Perspective". Presented at the 16^{th} Congress of the ICAS, Jerusalem, Israel, August 28 - September 2, 1988.

7. A. M. Yaglom. *Correlation Theory of Stationary and Related Random Functions.* Springer-Verlag, New York, 1987.

8. J. L. Doob. *Stochastic Processes.* Wiley, New York, 1953.

9. B. V. Gnedenko. *The Theory of Probability.* Chelsea, New York, 1962.

10. H. Cramér and M. R. Leadbetter. *Stationary and Related Stochastic Processes; Sample Function Properties and Their Applications.* Wiley, New York, 1967.

11. L. H. Koopmans. *The Spectral Analysis of Time Series.* Academic Press, New York, 1974.

12. P. J. Brockwell and R. A. Davis. *Time Series: Theory and Methods.* Springer-Verlag, New York, 1991.

13. H. Cramér. "On the Theory of Stationary Random Processes". *Annals of Mathematics*, 41(1):215–230, January 1940.

14. A. N. Kolmogorov. *Foundations of the Theory of Probability.* second English edition, Chelsea, New York, 1956.

15. M. Loève, editor. *Probability Theory.* third edition, Van Nostrand, Princeton, NJ, 1963.

16. A. N. Kolmogorov. "Stationary Sequences in Hilbert Space". *Bull. Math. Moscow Univ.*, 2(6):1–40, 1941. English translation in Kailath, 1977, pp. 66–89.

17. E. Parzen. "Statistical Inference on Time Series by Hilbert Space Methods, I". Technical Report No. 23, Department of Statistics, Stanford University, 1959. Published in "Time Series Analysis Papers" by E. Parzen. Holden-Day, San Francisco.

18. T. Kailath, editor. *Linear Least-Squares Estimation*, volume 17 of *Benchmark Papers in Electrical Engineering and Computer Sciences.* Dowden, Hutchinson & Ross, Stroudsburg, PA, 1977.

19. G. M. Jenkins and D. G. Watts, editors. *Spectral Analysis and its Applications.* Holden-Day, San Francisco, CA, 1968.

20. M. Rosenblatt. *Stationary Sequences and Random Fields.* Birkhauser, Boston, 1985.

21. D. E. Dudgeon and R. M. Mersereau. *Multidimensional Digital Signal Processing.* Prentice-Hall, Englewood Cliffs, NJ, 1984.

22. S. J. Orfanidis. *Optimum Signal Processing: An Introduction.* second edition, McGraw-Hill, New York, 1990.

23. A. V. Oppenheim, A. S. Willsky, and I. T. Young. *Signals and Systems.* Prentice-Hall, Englewood Cliffs, NJ, 1983.

24. A. Papoulis. *The Fourier Integral and its Applications.* McGraw-Hill, New York, 1962.

25. M. J. Lighthill, editor. *Introduction to Fourier Analysis and Generalised Functions.* Cambridge University Press, Cambridge, England, 1958.

26. K. V. Mardia. *Families of Bivariate Distributions.* Griffin's Statistical Monographs & Courses. Hafner, Darien, CT, 1970.

27. N. L. Johnson and S. Kotz, editors. *Distributions in Statistics: Continuous Multivariate Distributions.* Wiley, New York, 1972.

28. D. S. Zrnić. "Simulation of Weatherlike Doppler Spectra and Signals". *Journal of Applied Meteorology*, 14(4):619–620, June 1975.

29. J. S. Bendat and A. G. Piersol, editors. *Random Data; Analysis and Measurement Procedures.* Wiley, New York, 1971.

30. S. M. Kay, editor. *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice-Hall, Englewood Cliffs, NJ, 1993.

31. R. A. Silverman. "Locally Stationary Random Processes". *IRE Transactions on Information Theory*, pages 182–187, September 1957.

32. M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions.* Dover, New York, 1965.