การประยุกต์ของระเบียบวิธีเวฟเล็ตในฟีลด์สุ่ม

นายกิติพล นวลทอง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาคณิตศาสตร์ประยุกต์
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2554

APPLICATION OF WAVELET METHODS TO RANDOM FIELDS

Kitipol Nualtong

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Applied Mathematics
Suranaree University of Technology
Academic Year 2011

APPLICATION OF WAVELET METHODS TO RANDOM FIELDS

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Thesis Examining Committee

(Assoc. Prof. Dr. Prapasri Asawakun)
Chairperson

(Asst. Prof. Dr. Eckart Schulz)
Member (Thesis Advisor)

(Asst. Prof. Dr. Songkiat Sumetkijakan)
Member

(Prof. Dr. Pairote Sattayatham)
Member

(Prof. Dr. Sergey Meleshko)
Member

(Assoc. Prof. Dr. Nikolay Moshkin)
Member

(Prof. Dr. Sukit Limpijumnong)
Vice Rector for Academic Affairs

(Assoc. Prof. Dr. Prapun Manyum)
Dean of Institute of Science

กิติพล นวลทอง : การประยุกต์ของวิธีเวฟเล็ตในฟีลด์สุ่ม (APPLICATION OF WAVELET METHODS TO RANDOM FIELDS)
อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร. เอ็กคาร์ท ชูล์ซ, 148 หน้า.

การแปลงเวฟเล็ตต่อเนื่องของฟีลด์สุ่มมิติจำกัด ซึ่งมีความนิ่งอย่างอ่อนหรือส่วนเพิ่มมีความนิ่งอย่างเข้มหรือส่วนเพิ่มมีความนิ่งอย่างอ่อนโดยเมทริกซ์การเปลี่ยนขนาดใด ๆ เป็นฟีลด์สุ่มใหม่ซึ่งมีความนิ่งอย่างอ่อนร่วมเมื่อใช้เมทริกซ์การเปลี่ยนขนาดที่ต่างกัน นอกจากนี้ สามารถคำนวณหาฟังก์ชันสหสัมพันธ์ไขว้และฟังก์ชันความหนาแน่นเชิงสเปกตรัมกำลังไขว้ และอภิปรายสมบัติเออโกดิกของฟีลด์สุ่มใหม่เหล่านี้ด้วย

เราสามารถใช้ระเบียบวิธีเวฟเล็ตเชิงภินทนะสร้างบราวเนียนโมชัน และบราวเนียนชีทได้โดยการสร้างบราวเนียนชีทจากฐานหลักเวฟเล็ตของปริภูมิ $L^2([0,1]^d)$ ด้วยฟังก์ชันฮาร์ และขยายวิธีการนี้ไปยังฐานหลักเวฟเล็ตโดยฟังก์ชันค้ำจุนกระชับด้วย

สาขาวิชาคณิตศาสตร์ ลายมือชื่อนักศึกษา ____________________

ปีการศึกษา 2554 ลายมือชื่ออาจารย์ที่ปรึกษา ________________

KITIPOL NUALTONG : APPLICATION OF WAVELET METHODS TO RANDOM FIELDS. THESIS ADVISOR : ASST. PROF. ECKART SCHULZ, Ph.D. 148 PP.

CONTINUOUS WAVELET / DISCRETE WAVELET / RANDOM FIELD / ERGODICITY PROPERTY / POWER SPECTRAL DENSITY / FRACTIONAL BROWNIAN FIELD / BROWNIAN MOTION / BROWNIAN SHEET

The continuous wavelet transform of three types of finite dimensional random fields, namely weakly stationary random fields, random fields with stationary increments and random fields with weakly stationary increments are considered. It is shown that the transformed fields by different dilation matrices are jointly weakly stationary, and the cross-correlation function and cross power spectral density function are determined. In addition, ergodicity properties of the transformed fields are discussed.

The discrete wavelet method is used to construct Brownian motion and Brownian sheets. We employ a Haar wavelet basis of $L^2([0,1]^d)$ to construct a Brownian sheet, and then extend this framework to an arbitrary compactly supported wavelet basis of $L^2([0,1]^d)$ and obtain the representation of a Brownian sheet by wavelets.

School of Mathematics Student's Signature ________________

Academic Year 2011 Advisor's Signature ________________

ACKNOWLEDGEMENTS

My gratitude goes to Asst. Prof. Dr. Eckart Schulz for his outstanding guidance throughout the completion of this work and his commitment to excellent teaching at Suranaree University of Technology.

Heartfelt appreciation goes to Assoc. Prof. Dr. Prapasri Asawakun, Prof. Dr. Sergey V. Meleshko, and Assoc. Prof. Dr. Nikolay Moshkin for their quality lecturing and support during my stay at the School of Mathematics.

I would also like to thank Prof. Dr. Pairote Sattayatham and Asst. Prof. Dr. Songkiat Sumetkijakan for their asserted comments which significantly improved the quality of this dissertation.

Special mention deserve Asst. Prof. Dr. Jessada Tanthanuch for his assistance with the Latex software. I would like to thank all of my friends at the School for their companionship.

I express my sincere appreciation to both the Ministry of University Affairs of Thailand (MUA) and Prince of Songkla University for providing the necessary funds to support this study.

Finally, my deepest appreciation is to my parents for their support and encouragement during my studies.

Kitipol Nualtong

CONTENTS

CONTENTS (Continued)

CONTENTS (Continued)

CHAPTER I

INTRODUCTION

Wavelets are known to have intimate connections to several other parts of mathematics, notably phase-space analysis of signal processing, reproducing kernel Hilbert spaces etc. Random fields have found numerous applications in diverse areas such as image processing, signal processing, oceanography, geology, forestry, turbulence, geography, finance and engineering. This thesis explores the connection of wavelets with random fields in further detail. First, we continue from the work on the continuous wavelet transform of random fields accumulated over the last 10 years, to find the spectral representation and consider some properties of continuous wavelet transforms of random fields. Secondly, we use the discrete wavelet method to construct a wavelet representation of Brownian motion, which is an example of a stochastic process with particular properties, from a wavelet basis and then extend this construction to the higher dimensional setting of random fields.

1.1 Continuous Wavelet Transforms and Random Fields

1.1.1 Previous Research

The continuous wavelet transform, following Daubechies (1992), Walnut (2001) and Pinsky (2009), is a tool of analysing a square integrable function on $\mathbb{R}^d$ by correlating it with a two-parameter family of functions, obtained from translating and scaling of a single analysing function, called mother wavelet function. On the other hand, a random field, following Childers (1998), Grimmett and Strizaker (1998) and Ludeman (2003), is a family of random variables from a probability space to a Borel subset of $\mathbb{R}$ or $\mathbb{C}$ usually parametized in some continuous fashion by a position variable in Euclidean space. A random field in one dimensional (time) space is usually called a stochastic process. The connection of wavelets with stochastic processes has been exploited for over 20 years by Averkamp and Houdre (1998), Benedetto and Frazier (1994), Combanis and Houdre (1995), Flandrin (1989 and 1992), Kato and Masry (1999), Paulo and Rudolf (1999), Masry (1993), Meyer, Sellan and Taqqu (1999), Ramanathan and Zeitouni (1991) and Tewfik and Kim (1992) etc. On the other hand, the connection of wavelets with general random fields began to be established more recently, over the last 15 years by Averkamp and Houdre (1998), Heneghan, Lowen and Teich (1996) and Masry (1995 and 1998).

A random field is said to be weakly stationary if it has a constant mean and the auto-correlation is invariant under position shifts. It is called a weakly stationary increments random field, if the random fields composed of the increments are weakly stationary. Also, two random fields are jointly weakly stationary if they are individually weakly stationary and the cross-correlation is position shift invariant. One observes that the auto-correlation function of a weakly stationary random field is a function of position shift only. For this type of random field, a characterization of auto-correlation in the frequency domain is available as introduced by Childer (1998), Grimmett and Stirzaker (1998) and Ludeman (2003). Consequentially, one can define the power spectral density function as the generalized Fourier transform of the auto-correlation function. Similarly, if two random fields are jointly weakly stationary, then the cross-correlation function is again a function of only shift position, and the power cross spectral density function can be defined by the generalized Fourier transform of the cross-correlation function. Applying the inverse Fourier transform, one obtains the cross power spectral representation of the cross-correlation function.

Masry (1993) showed that the wavelet transform of a stochastic process with weakly stationary increments is a weakly stationary stochastic process whose auto-correlation function and spectral density function can be determined. Also, Cambanis and Houdre (1993) found a new proof that the wavelet transform of a stationary stochastic process as well as a stochastic process with stationary increments is a weakly stationary process, and then Averkamp and Houdre (1998) extended this viewpoint to random fields and obtained that the wavelet transforms of a random field, at different positive scaling parameters, are jointly weakly stationary random fields with zero mean. Furthermore, Masry (1998) determined the power spectral and cross power spectral representation of these random fields.

A fractional Brownian field is an example of a process which itself is not stationary, but whose increments are. This allows one to associate a well-defined spectral representation to such a process. Flandrin (1989) proposed how to obtain the spectral density function of the wavelet transform of fractional Brownian motion, which is the one dimensional case of a fractional Brownian field and Takeshi Kato and Elias Masry (1998) gave detailed proofs of this assertion. Furthermore, in 1996, Heneghand, Lowen and Teich considered the spectral density function of the wavelet transform of a two-dimensional fractional Brownian field, but their exposition is without proof.

The ergodic theorem, as presented in Viniotis (1998) for example states that the estimate for the mean converges to the true mean in the mean square sense. Grimmete and Stirzaker (1992) presented an ergodic theorem for weakly stationary random processes saying that given such a process, there exists some random variable with same mean to which the estimate for the mean of the process converges in the mean square sense.

1.1.2 The 1^{st} Objective of the Thesis

In this thesis, we discuss the classification of the continuous wavelet transform of three classes of d-dimensional random fields: weakly stationary random fields, stationary increments random fields and weakly stationary increments random fields. In each case we determine the spectral density function of the wavelet transform via arbitrary dilation matrix of the random field. Moreover, as an example, we obtain the spectral density function of the wavelet transform of a fractional Brownian field in the general d-dimensional case. We further investigate the ergodic property of the transformed random field, for both weakly stationary random fields as well as random fields with weakly stationary increments.

1.2 Discrete Wavelet Methods and Random Fields

1.2.1 Previous Research

Brownian motion is a stochastic process having continuous sample paths and independent increments (see also Section 3.5 and Appendix E). As shown in Childers (1998), Grimmett and Stirzaker (1998) and Michael (2000), Brownian motion is a self-similar stochastic process. Here, self-similarity of a stochastic process is a form of statistical scale invariance. Since wavelets are also naturally associated with scaling, there have been a number of attempts to represent Brownian motion in terms of wavelets. Following Michael (2000) and Pinsky (2009), Brownian motion can be constructed by means of the Haar wavelet.

1.2.2 The 2^{nd} Objective of the Thesis

In this thesis, we develop a framework for constructing the wavelet representation of Brownian motion by a wavelet basis of $L^2[0, 1]$ different from the Haar basis. We obtain a compactly supported wavelet function generating a representation of Brownian motion. We then extend this construction to the multidimensional case, using both Haar function and arbitrary compactly supported wavelet bases to generate a representation of a d-dimensional Brownian sheet.

1.3 Overview of the Thesis

This thesis is organized as follows. In Chapter II the basic notation is introduced and the concepts from the continuous Fourier transform, the Fourier transform of a measure, the continuous wavelet transform and the discrete wavelet methods as used in this thesis are reviewed. In Chapter III, random fields and their probabilistic properties such as the correlation and covariance function are reviewed, special classes of random fields are introduced and the power spectral density of some classes of random fields is reviewed. Moreover, the wavelet transform of random fields is discussed. Chapter IV is devoted to the discussion of the wavelet transform of random fields and the determination of the power spectral density function of the wavelet transform of weakly stationary random fields, stationary increment random fields and weakly stationary increment random fields. Also as an example, the power spectral density function of the wavelet transform of a fractional Brownian field is presented. In chapter V, mean ergodic random fields and the ergodic theorem for weakly stationary random fields are discussed and connected to the continuous wavelet transform of some classes of random fields. In Chapter VI, the discrete wavelet method is employed for obtaining a wavelet representation of some random fields such as Brownian motion and Brownian sheet.

CHAPTER II

FOURIER ANALYSIS AND WAVELETS

In this chapter, we review the mathematical concepts used in this thesis. We begin by discussing the Fourier transform of an integrable function, and also functions in Schwartz space and their basic properties. We then review the Fourier transform of a measure and Bochner's theorem. Finally, we review the continuous wavelet transform and the discrete wavelet method. Throughout, it is assumed that the reader is familiar with the foundations of real analysis, such as measure theory and function spaces.

2.1 The Continuous Fourier Transform

Throughout, $\mathbb{R}^d$ will denote the d-dimensional Euclidean space, and $\hat{\mathbb{R}}^d$ its algebraic dual. It is well known that $\hat{\mathbb{R}}^d$ can be identified with $\mathbb{R}^d$ itself through the usual inner product,

$$\langle x, \xi \rangle = x \cdot \xi \quad \text{for } x \in \mathbb{R}^d, \xi \in \hat{\mathbb{R}}^d.$$

In applications, $\mathbb{R}^d$ is often called the space domain (or time domain, if $d = 1$), and $\hat{\mathbb{R}}^d$ the Fourier domain, or frequency domain.

In this section we begin to develop the properties of the Fourier transform of integrable functions and of tempered distributions. Our main interest is in the basic rules for the transform. The inverse Fourier transform and properties involving convolution and some further operators will be studied later. Details can be found in Folland (1999), Gasquet and Witomski (1999) and Strichartz (1994).

2.1.1 L^1-Fourier Transform

We summarize the definition and elementary properties of the L^1-Fourier transform in the following definitions and remarks.

Definition 2.1. (L^1-Fourier Transform)

Let $f \in L^1(\mathbb{R}^d)$. Its Fourier transform is the function

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-i\xi \cdot x}\, dx \quad \text{for } \xi \in \hat{\mathbb{R}}^d.$$

As an illustration, the Fourier transform of $f(x) = e^{-\alpha|x|^2}$ ($x \in \mathbb{R}^d$ where $\alpha > 0$) is computed in Example A.1 (Appendix A).

Remark 2.1. The basic properties of the Fourier transform are as follows:

i) This integral make sense if and only if $f \in L^1(\mathbb{R}^d)$, since $|e^{-i\xi\cdot x}| = 1$.

ii) Consider the map $\mathcal{F}$ on $L^1(\mathbb{R}^d)$ given by $\mathcal{F}(f) = \hat{f}$. Then $\mathcal{F}$ is a linear, one-to-one and norm-reducing operator of $L^1(\mathbb{R}^d)$ into $C_0(\mathbb{R}^d)$: $\|\mathcal{F}(f)\|_\infty \leq \|f\|_1$. This operator is also called the Fourier transform. (We use the notation $\mathcal{F}$ for the Fourier transform only where it is needed for clearity.)

The following formula is essential for introducing the inverse Fourier transform.

Proposition 2.1. *Let f and g be two integrable functions. Then $f\hat{g}$ and $\hat{f}g$ are integrable functions on $\mathbb{R}^d$ and*

$$\int_{\mathbb{R}^d} f(t)\hat{g}(t)\, dt = \int_{\mathbb{R}^d} \hat{f}(x) g(x)\, dx$$

Proof. Details of the proof can be found in Gasquet and Witomski (1999), pp.156.

□

Definition 2.2. (L^1-Inverse Fourier Transform)

Let $f \in L^1(\mathbb{R}^d)$. Its inverse Fourier transform is

$$\check{f}(x) = \frac{1}{(2\pi)^d} \int_{\hat{\mathbb{R}}^d} f(\xi) e^{i\xi\cdot x}\, d\xi.$$

Remark 2.2. Since this definition is very similar to definition of $\hat{f}$, the basic properties of the inverse Fourier transform follow imediately:

i) This integral make sense if and only if $f \in L^1(\hat{\mathbb{R}}^d)$, since $|e^{i\xi\cdot x}| = 1$.

ii) The map $\mathcal{F}^{-1}$ on $L^1(\hat{\mathbb{R}}^d)$ given by $\mathcal{F}^{-1}(f) = \check{f}$ is a linear, one-to-one and norm-reducing operator of $L^1(\mathbb{R}^d)$ into $C_0(\mathbb{R}^d)$: $\|\mathcal{F}^{-1}(f)\|_\infty \leq \frac{1}{(2\pi)^d}\|f\|_1$. (We use the notation $\mathcal{F}^{-1}$ for the inverse Fourier transform only where it is needed for clarity.)

The Fourier transform is remarkable in that the inverse operator is obtained very simply from $\mathcal{F}$ itself. In fact, it is just $\mathcal{F}^{-1}$. However, one must be cautions, f being integrable does not imply that $\hat{f}$ is integrable. One needs an additional hypotheses on f to invert the Fourier transform $f \mapsto \hat{f}$, as in the following theorem.

Theorem 2.2. *If $f \in L^1(\mathbb{R}^d)$ and $\hat{f} \in L^1(\hat{\mathbb{R}}^d)$, then*

$$f(x) = \mathcal{F}^{-1}(\hat{f})(x) = \frac{1}{(2\pi)^d}\int_{\hat{\mathbb{R}}^d} \hat{f}(\xi)e^{i\xi\cdot x}\,d\xi \qquad a.e\ x \in \mathbb{R}^d.$$

Proof. Details of proof can be found in Gasquet and Witomski (1999), pp.163-165. □

One of the remarkable properties of the Fourier transform is the relation between derivation and multiplication by a monomial as in the following theorem.

Theorem 2.3. *Suppose that $f \in \mathbb{R}$. If f is continuous, piecewise smooth and $f' \in L^1(\mathbb{R})$, then*

$$\mathcal{F}(f')(\xi) = i\xi\mathcal{F}(\xi). \tag{2.1}$$

On the other hand, if $xf(x)$ is integrable, then

$$\mathcal{F}(xf(x))(\xi) = i[\mathcal{F}(f)]'(\xi) \tag{2.2}$$

Proof. Details of the proof can be found in Folland (1999), pp.250. □

2.1.2 The Schwartz Space and Tempered Distributions

Having presented the theory of the continuous Fourier transform for function that are integrable on $\mathbb{R}^d$, we wish to extend it to generalized functions, including functions that do not decay at infinity. If f and g are integrable functions on $\mathbb{R}^d$, by Proposition 2.1, one has $\int_{\mathbb{R}^d} \hat{f}(x)g(x)\,dx = \int_{\mathbb{R}^d} f(t)\hat{g}(t)\,dt$. That is, in distributional notation, $\langle \hat{f}, g\rangle = \langle f, \hat{g}\rangle$. This suggests to define the Fourier transform $\hat{F}$ of a distribution F by $\langle \hat{F}, \varphi\rangle = \langle F, \hat{\varphi}\rangle$. There is, however, the problem that sometimes $\langle \hat{F}, \varphi\rangle$ makes sense, while $\langle F, \hat{\varphi}\rangle$ may not make sense. For most purposes, a better solution to this problem is to reduce the class of special functions F and, correspondingly, to restrict the class of allowable functions φ. The nicest way of doing this was discovered by Laurent Schwartz.

Definition 2.3. (Schwartz Space)
Given $N \in \mathbb{N}_0$ and multi-index α, we let

$$\|f\|_{(N,\alpha)} = \sup_{x\in\mathbb{R}^d}(1+\|x\|)^N|(D^\alpha f)(x)|,$$

for $f \in C^\infty(\mathbb{R}^d)$. Set $S(\mathbb{R}^d) = \{f \in C^\infty(\mathbb{R}^d) \;:\; \|f\|_{(N,\alpha)} < \infty, \forall N \in \mathbb{N}_0, \forall \alpha \in \wedge\}$ where $\wedge$ is the set of all multi-indices.

It turns out that $S(\mathbb{R}^d)$ is a complex vector space, and the family of seminorms $\|\cdot\|_{(N,\alpha)}$ determines a complete metric on $S(\mathbb{R}^d)$. We call $S(\mathbb{R}^d)$ the Schwartz space, and its element are called Schwartz functions. Thus a Schwartz function is a function f in class C^∞ such that f and its derivatives vanish at infinity more rapidly than any power of $(1+\|x\|)^N$ for all $N \in \mathbb{N}_0$.

Clearly, $S(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$ for all $1 \le p \le \infty$. Furthermore, the topology in $L^p(\mathbb{R}^d)$ is weaker than that in $S(\mathbb{R}^d)$: If $f_n \longrightarrow f$ in $S(\mathbb{R}^d)$, then $f_n \longrightarrow f$ in $L^p(\mathbb{R}^d)$.

Theorem 2.4. *The Fourier transform $\mathcal{F} : f \longmapsto \hat{f}$ maps $S(\mathbb{R}^d)$ homeomorphically onto $S(\mathbb{R}^d)$ in the topology of $S(\mathbb{R}^d)$ and the same is true for the inverse Fourier transform.*

Proof. Details of the proof can be found in Pathak (2001), pp.65-67. □

Definition 2.4. (Tempered Distributions)
Consider the (topological) dual of $S(\mathbb{R}^d)$,

$$S'(\mathbb{R}^d) = \{\Phi : S(\mathbb{R}^d) \to \mathbb{C} | \Phi \text{ is linear and continuous}\}.$$

Elements of $S'(\mathbb{R}^d)$ are called tempered distributions.

Example 2.1. In the following we present some classes of tempered distribution. Throughout, f denotes an arbitrary element of $S(\mathbb{R}^d)$.

1) Let $\phi \in L^q(\mathbb{R}^d)$, $\;1 \leq q \leq \infty$. As $S(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$, where $\frac{1}{p} + \frac{1}{q} = 1$, and the topology on $L^p(\mathbb{R}^d)$ is weaker than that on $S(\mathbb{R}^d)$, then $\Phi(f) = \langle f, \phi \rangle = \displaystyle\int_{\mathbb{R}^d} f(x)\overline{\phi(x)}\, dx$ defines an element $\Phi \in S'(\mathbb{R}^d)$.

2) Let $\phi : \mathbb{R}^d \to \mathbb{C}$ be measurable and $|\phi(x)| \leq C(1 + \|x\|)^M$, C is a constant, for all $x \in \mathbb{R}^d$ (such a ϕ is called slowly increasing). Then $\Phi(f) = \langle f, \phi \rangle = \displaystyle\int_{\mathbb{R}^d} f(x)\overline{\phi(x)}\, dx$ defines an element $\Phi \in S'(\mathbb{R}^d)$.

3) Let μ be a finite measure on $\mathbb{R}^d$.
Then $\Phi(f) = \displaystyle\int_{\mathbb{R}^d} f(x)\, d\mu(x)$ defines an element $\Phi \in S'(\mathbb{R}^d)$.

In order to extend the Fourier transform to $S'(\mathbb{R}^d)$, let $\Phi \in S'(\mathbb{R}^d)$ be given. The map $f \mapsto \langle \hat{f}, \Phi \rangle$ for f belonging to $S(\mathbb{R}^d)$, is a continuous linear functional on $S(\mathbb{R}^d)$, since the Fourier transform $f \mapsto \hat{f}$ is continuous on $S(\mathbb{R}^d)$. Thus there exists a tempered distribution $\hat{\Phi} \in S'(\mathbb{R}^d)$ such that $\langle \hat{f}, \Phi \rangle = \langle f, \hat{\Phi} \rangle$, $\forall f \in S(\mathbb{R}^d)$. We define $\hat{\Phi}$ to be the Fourier transform of Φ. The reason $S(\mathbb{R}^d)$ is useful in studying Fourier transforms is that $\hat{f} \in S(\mathbb{R}^d)$ by Theorem 2.4. Since elements of

$S'(\mathbb{R}^d)$ are considered generalized function, we call the map $\mathcal{F} : \Phi \in S'(\mathbb{R}^d) \longmapsto \hat{\Phi} \in S'(\mathbb{R}^d)$ the generalized Fourier transform.

2.2 Convolution

We establish the existence of the convolution for integrable functions and consider its properties, especially with relation to the Fourier transform. First, we give the definition of the convolution of two functions.

Definition 2.5. (Convolution)
Let f, g be measurable functions on $\mathbb{R}^d$. The convolution of f and g denoted $f * g$ is defined by

$$(f * g)(x) = \int_{\mathbb{R}^d} f(y)g(x-y)\,dy.$$

By a change of variables, the preceding definition is equivalent to

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y)\,dy.$$

Remark 2.3.

(1) Unless additional assumptions are made about f and g, the convolution may not be defined.

(2) Suppose $f, g \in L^1(\mathbb{R}^d)$. Then $f * g$ is defined a.e., $f * g \in L^1(\mathbb{R}^d)$ and $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$.

We have the following theorem on the Fourier transform of the convolution of two functions.

Theorem 2.5. *Suppose f and g are two integrable functions. Then*

$$\mathcal{F}(f * g) = \mathcal{F}(f) \cdot \mathcal{F}(g)$$

$$\mathcal{F}^{-1}(\mathcal{F}(f) \cdot \mathcal{F}(g)) = f * g.$$

Proof. Details of the proof can be found in Folland (1999), pp.249 and 258. □

2.3 Operators on $L^p(\mathbb{R}^d)$

We introduce operators on $L^p(\mathbb{R}^d)$ that will be used frequently hereafter.

Definition 2.6. (Translation, Modulation and Dilation Operator)

Let $f \in L^p(\mathbb{R}^d)$, $1 \leq p < \infty$.

1) Translation by $y \in \mathbb{R}^d$ and denoted by T_y is the operator defined by

$$(T_y f)(x) = f(x - y) \text{ for all } x \in \mathbb{R}^d.$$

2) Modulation by $\xi \in \mathbb{R}^d$ and denoted by M_ξ is the operator defined by

$$(M_\xi f)(x) = e^{2i\pi\xi\cdot x} f(x) \text{ for all } x \in \mathbb{R}^d.$$

3) Dilation by $A \in M_d(\mathbb{R})$ (such that $\det A \neq 0$) and denoted by D_A is the operator defined by

$$(D_A f)(x) = |\det A|^{-\frac{p}{2}} f(A^{-1}x) \text{ for all } x \in \mathbb{R}^d.$$

It is easy to verify that these are all surjective isometries.

2.4 The Fourier Transform of a Measure

We next define the Fourier transform of a finite Borel measure. Bochner's Theorem characterizes these transforms, and involves the notion of positive definiteness. Before describing it, let us review the analogous concept for matrices.

Let $A = [A_{ij}]_{N\times N}$ denote an $N \times N$ matrix with complex entries. Associated to this matrix is a quadratic form on $\mathbb{C}$, defined by $\langle Au, u\rangle = \sum_{i,j=1}^{N} A_{ij}\overline{u_i}u_j$ for a vector $u = (u_1, u_2, ...u_N)$ in $\mathbb{C}^N$. We say the matrix is positive semi definite if the quadratic form is always non negative, $\langle Au, u\rangle \geq 0$ for all u. Now we can define what is meant by a positive definite function on $\mathbb{R}^d$.

Definition 2.7. (Positive Definite Function)

A function φ on $\mathbb{R}^d$ is a positive definite function if for every finite set of $\{\xi_i\}_{i=1}^N$

in $\mathbb{R}^d$ and every finite set of complex numbers $\{c_n\}_{n=1}^N$ we have

$$\sum_{m,n=1}^{N} c_m \overline{c_n} \varphi(\xi_m - \xi_n) \geq 0.$$

Definition 2.8. (Fourier Transform of Measure)

Let μ be a finite Borel measure on $\mathbb{R}^d$. The Fourier transform of the measure μ is the function

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^d} e^{-i\xi\cdot x}\, d\mu(x).$$

If μ is a probability measure then $\hat{\mu}$ is referred to as the characteristic function of the measure μ. The map $\mu \mapsto \hat{\mu}$ from the set of finite Borel measures is additive, positive homogeneous and one-to-one.

The characteristic function of the measure arising from some random variables are computed in Example A.2 in Appendix A and Remark D.3 in Appendix D.

Theorem 2.6. *(Bochner's Theorem)*

A function φ is a positive definite function on $\mathbb{R}^d$ if and only if there exists a nonnegative Borel measure μ on $\mathbb{R}^d$ such that

$$\varphi(x) \quad = \quad \int_{\mathbb{R}^d} e^{-ix\cdot\xi}\, d\mu(\xi), \quad a.e.\ x \in \mathbb{R}^d. \tag{2.3}$$

Proof. A detailed proof can be found in Appendix B. □

2.5 Continuous Wavelet Transforms

In this section, we review the continuous wavelet transform in d-dimensions. In general, the dilation-parameter in the continuous wavelet transform is a matrix belonging to a closed subgroup of the group of invertible $d \times d$ matrices, as introduced in the following.

2.5.1 Matrix Groups

Definition 2.9. (General Linear Group)

Let $M_d(\mathbb{R})$ be the set of real $d \times d$ matrices. The general linear group denoted by $GL_d(\mathbb{R})$ is the set of invertible elements,

$$GL_d(\mathbb{R}) = \{A \in M_d(\mathbb{R}) | \det(A) \neq 0\} .$$

Remark 2.4. We can see that

1. $GL_d(\mathbb{R})$ is a group under matrix multiplication.

2. $GL_d(\mathbb{R}) = \det^{-1}(\mathbb{R} - \{0\})$ is open in $M_d(\mathbb{R})$. In particular, $GL_1(\mathbb{R}) = \mathbb{R} - \{0\}$.

2.5.2 The Continuous Wavelet Transform

We are now ready to state the definition of the continuous wavelet transform.

Definition 2.10. (Continuous Wavelet Transform)

Let H be a closed subgroup of $GL_d(\mathbb{R})$. Fix $\varphi \in L^2(\mathbb{R}^d)$, called the mother wavelet. For each $a \in H$, called the dilation parameter, and $b \in \mathbb{R}^d$, called the translation parameter, we set

$$\varphi_{a,b} = T_b D_a \varphi.$$

That is

$$\varphi_{a,b}(x) = T_b D_a \varphi(x) = D_a \varphi(x-b) = |\det a|^{-\frac{1}{2}} \varphi(a^{-1}(x-b)) \qquad (2.4)$$

defines a 2-parameter family of functions in $L^2(\mathbb{R}^d)$.

Define the wavelet transform of $f \in L^2(\mathbb{R}^d)$ associated with mother wavelet φ by the inner product

$$CW_f^a(b) \quad = \quad \langle f, \varphi_{a,b} \rangle.$$

That is

$$CW_f^a(b) = |\det a|^{-\frac{1}{2}} \int_{\mathbb{R}^d} f(x)\overline{\varphi(a^{-1}(x-b))}\, dx. \tag{2.5}$$

Example 2.2. Let $d = 1$ and $H = \mathbb{R}^+$ or $H = \{c^k : k \in \mathbb{Z}, \text{ for fixed } c > 1\}$.

Then for fixed $a \in H, b \in \mathbb{R}$

$$WC_f^a(b) = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} f(x)\overline{\varphi\left(\frac{x-b}{a}\right)}\, dx$$

gives information of $f(x)$ at scale a and location (time) determined by b.

Example 2.3. Let $d = 2$, $H = \left\{ \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} : a \neq 0 \right\}$

Fix $\varphi \in L^2(\mathbb{R}^2)$. For each $h = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \in H, b \in \mathbb{R}^2$, set

$$\begin{aligned} \varphi_{a,b}(x) &= |\det h|^{-\frac{1}{2}} \varphi(h^{-1}(x-b)) = |a|^{-1}\varphi\left(\begin{bmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{a} \end{bmatrix} \begin{bmatrix} x_1 - b_1 \\ x_2 - b_2 \end{bmatrix} \right) \\ &= |a|^{-1}\varphi\left(\frac{x_1 - b_1}{a}, \frac{x_2 - b_2}{a} \right). \end{aligned}$$

Thus the wavelet transform of $f \in L^2(\mathbb{R}^2)$ is

$$CW_f^a(b) = \frac{1}{a} \int_{\mathbb{R}^2} f(x_1, x_2)\overline{\varphi\left(\frac{x_1 - b_1}{a}, \frac{x_2 - b_2}{a} \right)}\, dx_1 dx_2.$$

In the case of general $\mathbb{R}^d$, set $H = \{aI_d : a \neq 0\}$. Then for each $h \in H, b \in \mathbb{R}^d$ we have

$$\varphi_{a,b}(x) = |\det h|^{-\frac{1}{2}} \varphi(h^{-1}(x-b)) = |a|^{-\frac{d}{2}} \varphi\left(\frac{x-b}{a}\right).$$

Thus the continuous wavelet transform of $f \in L^2(\mathbb{R}^d)$ is

$$CW_f^a(b) = |a|^{-\frac{d}{2}} \int_{\mathbb{R}^d} f(x)\overline{\varphi\left(\frac{x-b}{a}\right)}\, dx.$$

2.6 Discrete Wavelet Methods

In the continuous wavelet transform of Example 2.2 we consider the family $\varphi_{a,b}(x) = |a|^{-\frac{1}{2}}\varphi(\frac{x-b}{a})$, where $a, b \in \mathbb{R}, a \neq 0$. We would like to restrict a, b to discrete values only. Now we choose $a = a_0$, $b = na_0^m$, where m, n range over $\mathbb{Z}$ and $a_0 > 1$ is fixed. This corresponds to

$$\varphi_{m,n}(x) = a_0^{-\frac{m}{2}}\varphi\left(\frac{x - na_0^m}{a_0^m}\right) = a_0^{-\frac{m}{2}}\varphi\left(a_0^{-m}x - n\right).$$

There are some questions which arise naturally: Do the discrete wavelet coefficients $\langle f, \varphi_{m,n}\rangle$ completely characterize $f \in L^2(\mathbb{R})$? Furthermore, can every function f be expressed in terms of the $\varphi_{m,n}$? In the present discrete case there is no analogue of the resolution of the identity, so one has to attack the problem in some other way.

2.6.1 Multiresolution Analysis on $\mathbb{R}$

Daubechies (1992) opined that the first construction of smooth orthonormal wavelet bases by Meyer seemed a bit miraculous in that the Meyer wavelets constitute an orthonormal basis. This situation changed with the advent of multiresolution analysis, formulated in the fall of 1986 by Mallat and Meyer. Multiresolution analysis provides a natural framework for the understanding of wavelet bases, and for the construction of new examples. The construction of most wavelet bases of square integrable functions on the interval $[0, 1]$ derives from a multiresolution analysis on $L^2(\mathbb{R})$. We therefore review the concept of multiresolution analysis, as outlined in detail in Daubechies (1992), Meyer (1986) and Walnut (2001).

Definition 2.11. (Multiresolution Analysis on $L^2(\mathbb{R})$)

A multiresolution analysis (MRA) on $L^2(\mathbb{R})$ consists of sequence of a closed subspaces V_j where $j \in \mathbb{Z}$ of $L^2(\mathbb{R})$, satisfying

(M1) $V_j \subset V_{j+1}$ for all $j \in \mathbb{Z}$

(M2) $f \in V_j$ if and only if $f(2\cdot) \in V_{j+1}$ for all $j \in \mathbb{Z}$

(M3) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$

(M4) $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$, and

(M5) there exists a function $\varphi \in V_0$ such that $\{\varphi(\cdot - k) : k \in \mathbb{Z}\}$ is a complete orthonormal basis for V_0.

Remark 2.5.

1) The function φ whose existence is asserted in (M5) is called the scaling function of the given multiresolution analysis.

2) Sometimes condition (M5) is relaxed by assuming that $\{\varphi(\cdot - n) : n \in \mathbb{Z}\}$ is a Riesz basis for V_0. That is, for every $f \in V_0$ there exists a unique sequence $\{\alpha_n\}_{n\in\mathbb{Z}} \in l^2(\mathbb{Z})$ such that

$$f(x) = \sum_{n\in\mathbb{Z}} \alpha_n \varphi(x - n), \tag{2.6}$$

with convergence in $L^2(\mathbb{R})$ and

$$A \sum_{n\in\mathbb{Z}} |\alpha_n|^2 \leq \|f\|_2^2 \leq B \sum_{n\in\mathbb{Z}} |\alpha_n|^2 \tag{2.7}$$

with constants $0 < A \leq B < \infty$, independent of f. Observe that (M5) implies that $\{\varphi(\cdot - n) : n \in \mathbb{Z}\}$ is a Riesz basis for V_0 with $A = B = 1$.

Remark 2.6.

1) Usually, a multiresolution analysis is defined by first identifying the closed subspace V_0 and the scaling function φ, and then setting

$$V_j = \{f(2^j\cdot) : f \in V_0\}$$

so that (M2) holds.

2) Let $f \in V_0 = \overline{\text{span}\{T_k\varphi\}}_{k\in\mathbb{Z}}$. Then $T_k f \in V_0$ for all $k \in \mathbb{Z}$.

Similarly, since $T_k f \in V_0$, then $f = T_{-k}(T_k f) \in V_0$. Hence $f \in V_0$ if and only if $T_k f \in V_0$ for all $k \in \mathbb{Z}$. This property is called translation invariance, so V_0 is a translation invariant subspace of $L^2(\mathbb{R})$.

3) It follows from $(M2)$ that $D_{2^{-j}}$ is an isomorphism of V_0 onto V_j (this is proved in Walnut (2001)).

4) It follows from 3) , (M5) and 2) above that $\{\varphi_{j,k} : \varphi_{j,k}(x) = 2^{\frac{j}{2}}\varphi(2^j x - k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis of V_j.

5) By 4) we have $V_j = \overline{\text{span}\{\varphi_{j,k}\}}_{k\in\mathbb{Z}} \subset \overline{\text{span}\{\varphi_{j,k}\}}_{j,k\in\mathbb{Z}}$, and hence the inclusion that $\bigcup_{j\in\mathbb{Z}} V_j \subset \overline{\text{span}\{\varphi_{j,k}\}}_{j,k\in\mathbb{Z}}$, so that $L^2(\mathbb{R}) = \overline{\bigcup_{j\in\mathbb{Z}} V_j} \subset \overline{\text{span}\{\varphi_{j,k}\}}_{j,k\in\mathbb{Z}}$, that is

$$\overline{\text{span}\{\varphi_{j,k}\}}_{j,k\in\mathbb{Z}} = L^2(\mathbb{R}). \tag{2.8}$$

6) Since $\varphi \in V_0 \subset V_1$ and $\{\varphi_{1,k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis of V_1, we have

$$\varphi = \sum_{k\in\mathbb{Z}} < \varphi, \varphi_{1,k} > \varphi_{1,k}. \tag{2.9}$$

Setting $h_k =< \varphi, \varphi_{1,k} >= \sqrt{2}\int_{\mathbb{R}} \varphi(x)\overline{\varphi(2x-k)}\, dx$, then we rewrite equation (2.9) as

$$\varphi = \sum_{k\in\mathbb{Z}} h_k \varphi_{1,k} \tag{2.10}$$

which often is written as

$$\varphi(x) = \sum_{k\in\mathbb{Z}} \sqrt{2} h_k \varphi(2x - k),$$

converges in $L^2(\mathbb{R})$, and is called the scaling relation, and $\{h_k\}_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})$ is called the scaling filter associated with φ. By Parseval's identity

$$\sum_{k\in\mathbb{Z}} |h_k|^2 = \sum_{k\in\mathbb{Z}} | < \varphi, \varphi_{1,k} > |^2 = \|\varphi\|_2^2.$$

7) Equation (2.9) can be formulated for every $f \in V_1$,

$$f = \sum_{k\in\mathbb{Z}} \langle f, \varphi_{1,k}\rangle \varphi_{1,k}.$$

Definition 2.12. (Orthogonal Projection)

Let H be a Hilbert space, V a closed subspace of H. Then for each $x \in H$, there exists unique element $y \in V$ such that $\|x-y\| = \inf_{z \in V} \|x-z\|$, we define this point as $Px = y$. If $x \in V$ then $Px = x$. The mapping P is called the orthogonal projection of H onto V.

Now suppose that we have a multiresolution analysis $\{V_j\}_{j\in\mathbb{Z}}$ with scaling function φ. Let P_j denote the orthogonal projection of $L^2(\mathbb{R})$ onto V_j, $j \in \mathbb{Z}$, then $P_j f = f$ for $f \in V_j$, and $P_j g = 0$ for $g \in V_j^{\perp}$. The projections P_j are called the approximation operators.

Note that if H is a Hilbert space, V a closed subspace of H and $\{e_n\}_{n\in J}$ an orthonormal basis for V, then for each $x \in H$, the projection of H onto V is given by

$$Px = \sum_{n\in J} \langle x, e_n\rangle e_n.$$

According to Remark 2.6 (4), $\{\varphi_{j,k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis for V_j. Thus, the approximation operators $P_j f \in V_j$ are given by

$$P_j f = \sum_{k\in\mathbb{Z}} \langle f, \varphi_{j,k}\rangle \varphi_{j,k} \tag{2.11}$$

for all $f \in V_j$. Let W_j denote the orthogonal complement of V_j in V_{j+1} then $V_{j+1} = V_j \oplus W_j$. The orthogonal projection Q_j of $L^2(\mathbb{R})$ onto W_j is called the detail operator. By straightforward computation, we have $Q_j = P_{j+1} - P_j$ so that for $f \in L^2(\mathbb{R}^d)$,

$$\lim_{j\to\infty} \|f - P_j f\|_2 = 0 \text{ and } \lim_{j\to-\infty} P_j f = 0, \tag{2.12}$$

details of the proof are given in Walnut (2001), pp.171 - 173. By these and a telescoping technique it follows that, for each $f \in L^2(\mathbb{R})$,

$$f = \sum_{j\in\mathbb{Z}} Q_j f.$$

Start with some closed subspace V_l of $L^2(\mathbb{R})$ then $L^2(\mathbb{R}) = V_l \oplus V_l^{\perp}$.

As $V_{l+1} = V_l \oplus W_l$ then

$$L^2(\mathbb{R}) = V_{l+1} \oplus V_{l+1}^{\perp} = V_l \oplus W_l \oplus V_{l+1}^{\perp}.$$

Continuing by induction we have, for all $m > 1$,

$$L^2(\mathbb{R}) = V_l \oplus W_l \oplus W_{l+1} \oplus ... \oplus V_{l+m} \oplus V_{l+m+1}^{\perp}.$$

Given $n > 0$ choose $l = -n, m = 2n$

$$L^2(\mathbb{R}) = V_{-n} \oplus W_{-n} \oplus W_{-n+1} \oplus ... \oplus W_{n-1} \oplus W_n \oplus V_{n+1}^{\perp}.$$

From 2.12 it follows, letting $n \to \infty$ that

$$L^2(\mathbb{R}) = \oplus_{j\in\mathbb{Z}} W_j.$$

Beginning with the scaling filter $\{h_k\}_{k\in\mathbb{Z}}$ we define a sequence $\{g_k\} \in l^2(\mathbb{Z})$ by $g_k = (-1)^k\overline{h_{1-k}}$, called the wavelet filter, and define an associated function $\psi \in W_1$ called mother wavelet, by

$$\psi = \sum_{k\in\mathbb{Z}} g_k \varphi_{1,k}. \tag{2.13}$$

We have an important theorem, a detailed proof of which is given in Walnut (2001).

Theorem 2.7. *Let φ be a scaling function of a MRA $\{V_j\}_{j\in\mathbb{Z}}$ on $L^2(\mathbb{R})$ and ψ the associated wavelet. Then*

1) $\{T_k\psi\}_{k\in\mathbb{Z}}$ is an orthonormal basis of W_0.

2) $\{\psi_{j,k} : \psi_{j,k}(x) = 2^{\frac{j}{2}}\psi(2^j x - k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis of W_j.

3) $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is an orthonormal (wavelet) basis of $L^2(\mathbb{R})$.

Remark 2.7.

1) From Theorem 2.7 (2) we have that the detail operators $Q_j f \in W_j$ are

given by

$$Q_j f = \sum_{k\in\mathbb{Z}} \langle f, \psi_{j,k}\rangle \psi_{j,k}. \tag{2.14}$$

2) From relation (2.10), applying the operator $D_{2^{-(j-1)}}T_m$ we have

$$\begin{aligned} D_{2^{-(j-1)}}T_m\varphi &= \sum_{k\in\mathbb{Z}} h_k D_{2^{-(j-1)}}T_m\varphi_{1,k} = \sum_{k\in\mathbb{Z}} h_k D_{2^{-(j-1)}}D_{2^{-1}}T_{2m+k}\varphi \\ &= \sum_{k\in\mathbb{Z}} h_k D_{2^{-j}}T_{2m+k}\varphi = \sum_{k\in\mathbb{Z}} h_{k-2m} D_{2^{-j}}T_k\varphi. \end{aligned}$$

So that

$$\varphi_{j-1,m} = \sum_{k\in\mathbb{Z}} h_{k-2m}\varphi_{j,k}. \tag{2.15}$$

Next, $\psi = \sum_{k\in\mathbb{Z}} g_k \varphi_{1,k}$ gives in the same way that

$$\psi_{j-1,m} = \sum_{k\in\mathbb{Z}} g_{k-2m}\varphi_{j,k}. \tag{2.16}$$

Since $V_j(\mathbb{R}) = V_{j-1} \oplus W_{j-1}$, we can see that for $f \in V_j$, $P_j f = P_{j-1}f + Q_{j-1}f$ thus by the relations (2.11) and (2.14)

$$\sum_{k\in\mathbb{Z}} \langle f, \varphi_{j,k}\rangle \varphi_{j,k} = \sum_{m\in\mathbb{Z}} \langle f, \varphi_{j-1,m}\rangle \varphi_{j-1,m} + \sum_{m\in\mathbb{Z}} \langle f, \psi_{j-1,m}\rangle \psi_{j-1,m}.$$

Then by equation (2.15) and (2.16) we have

$$\begin{aligned} \sum_{k\in\mathbb{Z}} \langle f, \varphi_{j,k}\rangle \varphi_{j,k} &= \sum_{m\in\mathbb{Z}} \left[\langle f, \varphi_{j-1,m}\rangle \sum_{k\in\mathbb{Z}} h_{k-2m}\varphi_{j,k} \right] + \sum_{m\in\mathbb{Z}} \left[\langle f, \psi_{j-1,m}\rangle \sum_{k\in\mathbb{Z}} g_{k-2m}\varphi_{j,k} \right] \\ &= \sum_{m\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \left[\langle f, \varphi_{j-1,m}\rangle h_{k-2m} + \langle f, \psi_{j-1,m}\rangle g_{k-2m} \right] \varphi_{j,k} \\ &= \sum_{k\in\mathbb{Z}} \left(\sum_{m\in\mathbb{Z}} \left[\langle f, \varphi_{j-1,m}\rangle h_{k-2m} + \langle f, \psi_{j-1,m}\rangle g_{k-2m} \right] \right) \varphi_{j,k}. \end{aligned}$$

Since $\{\varphi_{j,k}\}_{k\in\mathbb{Z}}$ is orthonormal, the coefficients on both sides must be identical. Then, for each $k \in \mathbb{Z}$ and $f \in V_j$,

$$\begin{aligned} \langle f, \varphi_{j,k}\rangle &= \sum_{m\in\mathbb{Z}} \left[\langle f, \varphi_{j-1,m}\rangle h_{k-2m} + \langle f, \psi_{j-1,m}\rangle g_{k-2m} \right] \\ &= \langle f, \sum_{m\in\mathbb{Z}} \left[\overline{h}_{k-2m}\varphi_{j-1,m} + \overline{g}_{k-2m}\psi_{j-1,m} \right] \rangle. \end{aligned}$$

Hence we have

$$\varphi_{j,k} = \sum_{m\in\mathbb{Z}} \left[\overline{h}_{k-2m}\varphi_{j-1,m} + \overline{g}_{k-2m}\psi_{j-1,m}\right].$$

In particular, if $j = 1, k = 0$ we get

$$\varphi_{1,0} = \sum_{m\in\mathbb{Z}} \left[\overline{h}_{-2m}\varphi_{0,m} + \overline{g}_{-2m}\psi_{0,m}\right],$$

that is

$$\sqrt{2}\varphi(2x) = \sum_{m\in\mathbb{Z}} \left[\overline{h}_{-2m}\varphi(x-m) + \overline{g}_{-2m}\psi(x-m)\right]. \tag{2.17}$$

If $j = 1, k = 1$ we get

$$\varphi_{1,1} = \sum_{m\in\mathbb{Z}} \left[\overline{h}_{1-2m}\varphi_{0,m} + \overline{g}_{1-2m}\psi_{0,m}\right],$$

that is

$$\sqrt{2}\varphi(2x-1) = \sum_{m\in\mathbb{Z}} \left[\overline{h}_{1-2m}\varphi(x-m) + \overline{g}_{1-2m}\psi(x-m)\right]. \tag{2.18}$$

2.6.2 Wavelet Bases with Compact Support

In this subsection we are mainly interested in constructing a wavelet basis consisting of compactly supported wavelets which is important for the construction of wavelet bases on the interval, in the next subsection.

Recall equations (2.9) and (2.10) defining the scaling filter $\{h_m\}_{m\in\mathbb{Z}}$ associated with the scaling function φ. If this sequence has finite length, $h_m = 0$ for $m < 0$ or $m > 2N-1$, $\quad N \in \mathbb{Z}^+$, then the corresponding basis wavelet has compact support. This can be checked from the recursive definition of the η_l, see more details of this recursive relation in Meyer (1987) and Daubechies (1988),

$$\varphi(x) = \lim_{l\to\infty} \eta_l(x)$$

$$\eta_l(x) = \sqrt{2}\sum_{m\in\mathbb{Z}} h_m\eta_{l-1}(2x-m), \eta_0 = \chi_{[-\frac{1}{2},\frac{1}{2})}.$$

Consider $\eta_1 = \sqrt{2}\sum_{m\in\mathbb{Z}} h_m \eta_0(2x-m)$. Recall that $h_m \neq 0$ only for $0 \leq m \leq 2N-1$, and $\text{supp}(\eta_0) \subset [-\frac{1}{2}, \frac{1}{2}]$. If $x \in \text{supp}(\eta_1)$, therefore, then $-\dfrac{1}{2} \leq 2x - m < \dfrac{1}{2}$ for all $0 \leq m \leq 2N-1$. In particular, $-\dfrac{1}{2} \leq 2x < \dfrac{1}{2} + (2N-1)$ or equivalently, $-\dfrac{1}{4} \leq x < N - \dfrac{1}{4}$.

Continue by induction. We obtain, when $x \in \text{supp}(\eta_2)$, then $-\dfrac{1}{4} \leq 2x - m \leq N - \dfrac{1}{4}$ for all $0 \leq m \leq 2N-1$ and hence $-\dfrac{1}{4} \leq 2x \leq (N - \dfrac{1}{4}) + (2N-1)$ or $-\dfrac{1}{8} \leq x \leq \dfrac{3N}{2} - \dfrac{5}{8}$.

Continuing, when $x \in \text{supp}(\eta_l)$ we have

$$-\frac{1}{2^{l+1}} \leq x \leq \frac{1}{2^{l+1}} + \frac{2^l - 1}{2^l}(2N-1).$$

Let $l \to \infty$, as $\varphi(x) = \lim_{l\to\infty} \eta_l(x)$ we obtain that $\text{supp}(\varphi) \subset [0, 2N-1)$

Therefore in this case, we can rewrite equation (2.10) as the following equation

$$\varphi(\frac{x}{2}) = \sum_{m=0}^{2N-1} \sqrt{2} h_m \varphi(x-m), \quad h_0 h_{2N-1} \neq 0.$$

Then the associated wavelet ψ has the same support (by a simple translation) and we can rewrite equation (2.13) by

$$\psi(\frac{x}{2}) = \sum_{m=0}^{2N-1} \sqrt{2} g_m \varphi(x-m), \quad g_0 g_{2N-1} \neq 0.$$

Recall that equations (2.17) and (2.18) can be written as

$$\sqrt{2}\varphi(2x) = \sum_{m=-N+1}^{0} \left[\overline{h}_{-2m}\varphi(x-m) + \overline{g}_{-2m}\psi(x-m)\right],$$

and

$$\sqrt{2}\varphi(2x-1) = \sum_{m=-N+1}^{0} \left[\overline{h}_{1-2m}\varphi(x-m) + \overline{g}_{1-2m}\psi(x-m)\right].$$

Hence

$$\begin{aligned}\sqrt{2}\varphi(2x) = &\underbrace{\overline{h}_0\varphi(x) + \overline{g}_0\psi(x)}_{m=0} + \underbrace{\overline{h}_2\varphi(x+1) + \overline{g}_2\psi(x+1)}_{m=-1} \\ &+ \ldots + \underbrace{\overline{h}_{2N-2}\varphi(x+N-1) + \overline{g}_{2N-2}\psi(x+N-1)}_{m=-N+1}.\end{aligned}$$

That is

$$\sqrt{2}\varphi(2x) = \overline{h}_0\varphi(x) + \overline{h}_2\varphi(x+1) + ... + \overline{h}_{2N-2}\varphi(x+N-1)$$
$$+ \overline{g}_0\psi(x) + \overline{g}_2\psi(x+1) + ... + \overline{g}_{2N-2}\psi(x+N-1), \tag{2.19}$$

and similarly,

$$\sqrt{2}\varphi(2x-1) = \underbrace{\overline{h}_1\varphi(x) + \overline{g}_1\psi(x)}_{m=0} + \underbrace{\overline{h}_3\varphi(x+1) + \overline{g}_3\psi(x+1)}_{m=-1}$$
$$+ ... + \underbrace{\overline{h}_{2N-1}\varphi(x+N-1) + \overline{g}_{2N-1}\psi(x+N-1)}_{m=-N+1}.$$

That is

$$\sqrt{2}\varphi(2x-1) = \overline{h}_1\varphi(x) + \overline{h}_3\varphi(x+1) + ... + \overline{h}_{2N-1}\varphi(x+N-1)$$
$$+ \overline{g}_1\psi(x) + \overline{g}_3\psi(x+1) + ... + \overline{g}_{2N-1}\psi(x+N-1). \tag{2.20}$$

2.6.3 Wavelet Bases of $L^2[0,1]$ and $L^2([0,1]^d)$

Throughout this section, we let $j_0 \in \mathbb{Z}$ be such that $2^{j_0} \geq 4N-4$ and $\{V_j(\mathbb{R})\}_{j=1}^{\infty}$ be a multiresolution analysis on $L^2(\mathbb{R})$ with scaling function φ and associated wavelet ψ such that the scaling filter associated with φ, $\{h_m\}_{m\in\mathbb{Z}}$, has finite length, $h_m = 0$ for $m < 0$ or $m > 2N-1$. Hence $\text{supp}\varphi, \text{supp}\psi \subset [0, 2N-1]$. Now set

$$V_j([0,1]) = \text{span}\{\varphi_{j,k}\big|_{[0,1]} : \varphi_{j,k} \in V_j(\mathbb{R})\}$$

and

$$v_j([0,1]) = \text{span}\{\varphi_{j,k} : \text{supp}\varphi_{j,k} \subset [0,1]\}.$$

Since φ is a function of compact support, it is obvious that $V_j([0,1])$ is finite dimensional. Note that the collection of function $\varphi_{j,k}\big|_{[0,1]}$ which do not vanish on $[0,1]$ may not be orthogonal on $[0,1]$. Our goal is to show that this system is linearly independent, and thus can be made orthonormal by the Gram-Schmidt method. We now proceed to present elementary lemmas which will be useful in the analysis of functions defined on the interval.

Definition 2.13. (Multiresolution Analysis on $L^2([0,1])$)
Let $\{V_j(\mathbb{R})\}_{j=1}^{\infty}$ be a multiresolution analysis on $L^2(\mathbb{R})$. A sequence $\{V_j\}_{j\geq j_0}$ of closed subspaces of $L^2([0,1])$ is called a multiresolution analysis on $L^2([0,1])$ associated with $\{V_j(\mathbb{R})\}$ if

i) $\forall j \geq j_0,\, v_j([0,1]) \subset V_j \subset V_j([0,1])$

ii) $\forall j \geq j_0,\, V_j \subset V_{j+1}$.

Lemma 2.8. *If* $f(x) = \sum_{k\in\mathbb{Z}} c_k\varphi(x-k)$ *where* $c_k = \int_{-\infty}^{\infty} f(x)\overline{\varphi(x-k)}\,dx$, *is a function in* $V_0(\mathbb{R})$ *such that* $f(x) = 0$ *for* $x \leq 0$, *then* $c_k = 0$ *for* $k \leq -1$.

Proof. See Appendix C. □

For each j, denote by $S(j)$ the range of all translation parameters k so that the support of $\varphi_{j,k}(x) = 2^{\frac{j}{2}}\varphi(2^j x - k)$ intersects the interval $(0,1)$, that is, $\varphi_{j,k}\big|_{[0,1]} \neq 0$. Since supp$\varphi_{j,k} = [2^{-j}k, 2^{-j}(k+2N-1)]$ it follows that

$$S(j) = [-2N+2, 2^j - 1] \cap \mathbb{Z}. \tag{2.21}$$

The set $S(j)$ can be divided into three disjoint subsets $S_1(j), S_2(j)$ and $S_3(j)$, according to whether the interior of the support of $\varphi_{j,k}$ contains 0, the support of $\varphi_{j,k}$ is complete contained in $[0,1]$, or the support of $\varphi_{j,k}$ contains 1, respectively. By (2.21) it follow that

$S_1(j) = \mathbb{Z}\cap[-2N+2,-1]$, and in fact, $\text{supp}(\varphi_{j,k}) \subset (-\infty, \frac{2N-2}{2^j}] \subset (-\infty, \frac{1}{2}]$ for $k \in S_1(j)$,

$S_2(j) = \mathbb{Z}\cap[0, 2^j - 2N + 1]$, and in fact, $\text{supp}(\varphi_{j,k}) \subset [0,1]$ for $k \in S_2(j)$,

$S_3(j) = \mathbb{Z}\cap[2^j - 2N + 2, 2^j - 1]$, and in fact, $\text{supp}(\varphi_{j,k}) \subset [1 - \frac{2N-2}{2^j}, \infty) \subset [\frac{1}{2}, \infty)$ for $k \in S_3(j)$.

It is shown in Appendix C that for $j \geq j_0$ and any function $f(x) = \sum_{k\in\mathbb{Z}} c_k\varphi(2^j x - k)$ in $V_j(\mathbb{R})$, $f(x) = 0$ for $0 \leq x \leq 1$ implies that $c_k = 0$ for $-2N+2 \leq k \leq 2^j-1$. This implies that $\{\varphi(2^j x - k) : k \in S(j)\}$ is a linearly independent system. Moreover, we can show in Appendix C that there exist constants $C_2 \geq C_1 > 0$ such that for any sequence $\alpha_{j,k}, k \in S(j)$ of coefficients,

$$C_1 \left(\sum_{k\in S(j)} |a_{j,k}|^2 \right)^{\frac{1}{2}} \leq \left\| \sum_{k\in S(j)} a_{j,k}\varphi_{j,k}(x) \right\|_{L^2[0,1]} \leq C_2 \left(\sum_{k\in S(j)} |a_{j,k}|^2 \right)^{\frac{1}{2}}.$$

This will establish the following fundamental theorem.

Theorem 2.9. *For $j \geq j_0$, the functions $\varphi_{j,k}\big|_{[0,1]}, k \in S(j) = [-2N+2, 2^j - 1]$, form a Riesz basis of the space $V_j([0,1])$.*

Proof. See Appendix C. □

Note that, we have the following orthogonality relations for elements of the collection $\{\varphi_{j,k}\big|_{[0,1]}\}_{k\in S(j)}$.

1) If $k \in S_1(j)$ and $l \in S_2(j)$ we have as supp$\varphi_{j,k} \subset (-\infty, \frac{1}{2}]$ and supp$\varphi_{j,l} \subset [0,1]$,

$$\int_0^1 \varphi_{j,k}(x)\overline{\varphi_{j,l}(x)}\,dx = 2^j \int_{-\infty}^{\infty} \varphi(2^j x - k)\overline{\varphi(2^j x - l)}\,dx = 0$$

since $\{\varphi_{0,m}\}_{m\in\mathbb{Z}}$ is orthonormal in V_0.

2) If $k \in S_2(j)$ and $l \in S_3(j)$ we have supp$\varphi_{j,k} \subset [0,1]$ and supp$\varphi_{j,l} \subset [\frac{1}{2}, \infty)$, and similarly to 1),

$$\int_0^1 \varphi_{j,k}(x)\overline{\varphi_{j,l}(x)}\,dx = 0$$

3) If $k \in S_1(j)$ and $l \in S_3(j)$ we have as supp$\varphi_{j,k} \subset (-\infty, \frac{1}{2}]$ and supp$\varphi_{j,l} \subset [\frac{1}{2}, \infty)$, that $\varphi_{j,k}\overline{\varphi_{j,k}} = 0$ a.e., and hence

$$\int_0^1 \varphi_{j,k}(x)\overline{\varphi_{j,l}(x)}\,dx = 0.$$

4) If $k \in S_2(j)$ and $l \in S_2(j)$ we have $\int_0^1 \varphi_{j,k}(x)\overline{\varphi_{j,l}(x)}\, dx = \delta_{k,l}$ by orthogonality and since $\text{supp}\varphi_{j,k}, \text{supp}\varphi_{j,l} \subset [0,1], k \in S_2(j)$.

In order to transform the basis $\varphi_{j,k}\big|_{[0,1]}, k \in S(j)$ to an orthonormal basis it is thus only necessary to make the function in $\varphi_{j,k}\big|_{[0,1]}, k \in S_1(j)$ mutually orthonormal and similarly, to render the function in $\varphi_{j,k}\big|_{[0,1]}, k \in S_3(j)$ mutually orthonormal. At the first for $\varphi_{j,k}\big|_{[0,1]}, k \in S(j)$, we apply Gram-Schmidt to these functions one thus obtains, in lieu of $\varphi_{j,k}(x)\big|_{[0,1]}, k \in S_1(j)$, new functions $2^{\frac{j}{2}}\varphi^{\alpha}_{-2N+2}(2^j x), ..., 2^{\frac{j}{2}}\varphi^{\alpha}_{-1}(2^j x)$, near the boundary 0. Now, for functions $\varphi_{j,k}\big|_{[0,1]}, k \in S_3(j)$, we reflect and translate by 1 to the right to obtain functions near the boundary 1. Next, we thus apply Gram-Schmidt to these functions. One thus obtains, in lieu of $\varphi_{j,k}(x), k \in S_3(j)$, new functions $2^{\frac{j}{2}}\varphi^{\beta}_{2^j-2N+2}(2^j(1-x)), ..., 2^{\frac{j}{2}}\varphi^{\beta}_{2^j-1}(2^j(1-x))$, near the boundary 1. Hence we have the following proposition.

Proposition 2.10. *The collection of function*

$2^{\frac{j}{2}}\varphi^{\alpha}_{-2N+2}(2^j x), ..., 2^{\frac{j}{2}}\varphi^{\alpha}_{-1}(2^j x),$

$2^{\frac{j}{2}}\varphi(2^j x - k)\big|_{[0,1]}, 0 \le k \le 2^j - 2N + 1$

and $2^{\frac{j}{2}}\varphi^{\beta}_{2^j-2N+2}(2^j(1-x)), ..., 2^{\frac{j}{2}}\varphi^{\beta}_{2^j-1}(2^j(1-x)),$

is an orthonormal basis of $V_j[0,1]$

It is easy to see that the space V_j contains the orthonormal system $\varphi_{j,k}|_{[0,1]}, 0 \le k \le 2^j - 2N + 1$, and we add boundaries functions near 0 and 1 from the collections $2^{\frac{j}{2}}\varphi^{\alpha}_{-2N+2}(2^j x), ..., 2^{\frac{j}{2}}\varphi^{\alpha}_{-1}(2^j x)$ and $2^{\frac{j}{2}}\varphi^{\beta}_{2^j-2N+2}(2^j(1-x)), ..., 2^{\frac{j}{2}}\varphi^{\beta}_{2^j-1}(2^j(1-x))$, respectively.

The construction of an orthonormal wavelet basis on $[0,1]$ follows thus the classical scheme of multiresolution analysis. One has a nested sequence $V_j[0,1]$ of subspaces of $L^2[0,1]$, $j \ge j_0$. The union of the $V_j[0,1]$ is dense in $L^2[0,1]$ just as the union of the $V_j(\mathbb{R})$ is dense in $L^2(\mathbb{R})$. One denotes by $W_j[0,1]$ the orthogonal complement of $V_{j+1}[0,1]$ in $V_j[0,1]$, that is

$$W_j[0,1] = (V_{j+1}[0,1]) \cap (V_j[0,1])^{\perp}.$$

Observe that $W_j[0,1]$ is not the space of restrictions of functions in W_j to $[0,1]$ thus the next lemma will be useful in what follows.

Now set $V_j[0,\infty) = \text{span}\{\varphi_{j,k}\big|_{[0,\infty)} : \varphi_{j,k} \in V_j(\mathbb{R})\}$. In the following we establish the second goal of this section, which is to construct a wavelet basis of the space $W_j([0,1])$. For this purpose, we need some lemmas.

Lemma 2.11. *The functions* $\psi(x-k)\big|_{[0,\infty)}, -2N+2 \le k \le -N+1$, *belong to* $V_0[0,\infty)$.

Proof. See Appendix C. □

The following lemma is an almost immediate consequence of the previous lemma by changing variables.

Lemma 2.12. *The functions* $\psi(2^j x-k)|_{[0,1]}, -2N+2 \le k \le -N$ *or* $2^j - N + 1 \le k \le 2^j - 1$, *belong to* $V_j[0,1]$.

Proof. See Appendix C. □

Applying Theorem 2.9 and Lemma 2.12, we reach the following important theorem.

Theorem 2.13. *For each* $j \ge j_0$ *a basis of* $V_{j+1}[0,1]$ *is formed by* $\{\varphi_{j,k}|_{[0,1]}, -2N+2 \le k \le 2^j-1\} \cup \{\psi_{j,k}|_{[0,1]}, -N+1 \le k \le 2^j-N\}$.

Proof. See Appendix C. □

We now show that each $\psi_{j,k}|_{[0,1]}, -N+1 \le k \le 2^j-N$ is orthogonal to $\varphi_{j,k}|_{[0,1]}, -2N+2 \le k \le 2^j-1$ so that $\psi_{j,k}|_{[0,1]}, -N+1 \le k \le 2^j-N$ forms elements of the space $W_j([0,1])$.

Lemma 2.12 tells us that the restrictions of the $\psi_{j,k}$ to $[0,1]$ belong to $V_j[0,1]$ if $-2N+2 \leq k \leq -N$ or $2^j - N + 1 \leq k \leq 2^j - 1$. When $-N+1 \leq k \leq -1$ or $2^j - 2N + 2 \leq k \leq 2^j - N$, the restrictions of the functions $\psi_{j,k}$ certainly don't belong to $V_j[0,1]$, but are not orthogonal to $V_j[0,1]$ either.

In any case, one has

$$L^2[0,1] = V_{j_0}[0,1] \oplus W_{j_0}[0,1] \oplus W_{j_0+1}[0,1] \oplus ... \tag{2.22}$$

We already have an orthonormal basis of $V_{j_0}[0,1]$ by Proposition 2.10.

Now, let $-N+1 \leq k \leq 2^j - N$.

We distinguish the cases $-N+1 \leq k \leq -1$, $0 \leq k \leq 2^j - 2N + 1$, and $2^j - 2N + 2 \leq k \leq 2^j - N$ as follows.

Case I: If $-N+1 \leq k \leq -1$, then $\mathrm{supp}(\psi_{j,k}|_{[0,1]}) \subset [-\frac{1}{4}, \frac{1}{2}]$. We have

i) $\psi_{j,k}|_{[0,1]}$ are orthogonal to $\varphi_{j,l}|_{[0,1]}, 0 \leq l \leq 2^j - 1$, in fact $\mathrm{supp}(\varphi_{j,l}|_{[0,1]}) \subset [0, \frac{5}{4}]$ so that $\mathrm{supp}(\psi_{j,k}\overline{\varphi_{j,l}}|_{[0,1]}) \subset [0, \frac{1}{2}]$ and hence

$$\int_0^1 \psi_{j,k}(x)\overline{\varphi_{j,l}(x)}\, dx = \int_{-\infty}^{\infty} \psi_{j,k}(x)\overline{\varphi_{j,l}(x)}\, dx = 0$$

as $\varphi_{j,l} \in V_j(\mathbb{R})$ and $\psi_{j,l} \in W_j(\mathbb{R})$.

ii) Similarly, $\psi_{j,k}|_{[0,1]}$ are orthogonal to $\psi_{j,l}|_{[0,1]}, 0 \leq l \leq 2^j - N$, in fact $\mathrm{supp}(\psi_{j,l}|_{[0,1]}) \subset [0,1]$ so that $\mathrm{supp}(\psi_{j,k}\overline{\psi_{j,l}}|_{[0,1]}) \subset [0, \frac{1}{2}]$ and hence

$$\int_0^1 \psi_{j,k}(x)\overline{\psi_{j,l}(x)}\, dx = \int_{-\infty}^{\infty} \psi_{j,k}(x)\overline{\psi_{j,l}(x)}\, dx = 0$$

as $\psi_{j,l}$ and $\psi_{j,l}$ belong to $W_j(\mathbb{R})$ and $k \neq l$.

iii) What the functions $\psi_{j,k}$ still look is orthogonality to the $N-1$ functions $2^{\frac{j}{2}}\varphi^{\alpha}_{-2N+2}(2^j x), ..., 2^{\frac{j}{2}}\varphi^{\alpha}_{-1}(2^j x)|_{[0,1]}$. Since these $N-1$ functions form an orthogonal sequence, the corrections which make the $\psi_{j,l}, -N+ \leq k \leq -1$ orthogonal to $V_j[0,1]$ are obvious, by Gram-Schmidt process. One thus obtains $N-1$ functions

$2^{\frac{j}{2}}h_{-N+1}(2^jx), ..., 2^{\frac{j}{2}}h_{-1}(2^jx)$ where $h_{-N+1}, ..., h_{-1}$ are independent of j.

By i) and iii) we have that $\psi_{j,k}\big|_{[0,1]}$ are orthogonal to $V_j[0,1]$ and ii) says that $\psi_{j,k}\big|_{[0,1]}$ for $0 \le k \le 2^j - N$, are mutually orthogonal it suffices to make the functions $2^{\frac{j}{2}}h_{-N+1}(2^jx), ..., 2^{\frac{j}{2}}h_{-1}(2^jx)$ mutually orthogonal, by means of the Gram-Schmidt process, to obtain the $N-1$ wavelets clustered at 0 namely $2^{\frac{j}{2}}\psi^{\alpha}_{-N+1}(2^jx)$, ..., $2^{\frac{j}{2}}\psi^{\alpha}_{-1}(2^jx)$

Case II: If $0 \le k \le 2^j - 2N + 1, \text{supp}\psi_{j,k} \subset [0,1]$, then obviously $\psi_{j,k}\big|_{[0,1]}$ belong to the orthogonal complement of $V_j[0,1]$ in $V_{j+1}[0,1]$, and moreover $\psi_{j,k}\big|_{[0,1]}$ are orthogonal already.

Case III: If $2^j - 2N + 2 \le k \le 2^j - N$, $\text{supp}\psi_{j,k} \subset [\frac{1}{2}, \frac{5}{4}]$. Then we have

i) $\psi_{j,k}\big|_{[0,1]}$ are orthogonal to $\varphi_{j,l}\big|_{[0,1]}, -2N + 2 \le l \le 2^j - 2N + 1$, in fact $\text{supp}\varphi_{j,l}\big|_{[0,1]} \subset [-\frac{1}{2}, 1]$ so that $\text{supp}\psi_{j,k}\overline{\varphi_{j,l}}\big|_{[0,1]} \subset [\frac{1}{2}, 1]$ and hence

$$\int_0^1 \psi_{j,k}(x)\overline{\varphi_{j,l}(x)}\,dx = \int_{-\infty}^{\infty} \psi_{j,k}(x)\overline{\varphi_{j,l}(x)}\,dx = 0$$

as $\varphi_{j,l} \in V_j(\mathbb{R})$ and $\psi_{j,l} \in W_j(\mathbb{R})$.

ii) Similarly, $\psi_{j,k}\big|_{[0,1]}$ are orthogonal to $\psi_{j,l}\big|_{[0,1]}, -2N+2 \le l \le 2^j - 2N + 1$, in fact $\text{supp}\psi_{j,k}\overline{\psi_{j,l}}\big|_{[0,1]} \subset [\frac{1}{2}, 1]$ and hence

$$\int_0^1 \psi_{j,k}(x)\overline{\psi_{j,l}(x)}\,dx = \int_{-\infty}^{\infty} \psi_{j,k}(x)\overline{\psi_{j,l}(x)}\,dx = 0$$

as $\psi_{j,l}$ and $\psi_{j,l}$ belong to $W_j(\mathbb{R})$ and $k \ne l$.

iii) What the functions $\psi_{j,k}\big|_{[0,1]}$ still look is orthogonality to the $N-1$ functions $2^{\frac{j}{2}}\varphi^{\beta}_{2^j-2N+2}(2^j(1-x)), ..., 2^{\frac{j}{2}}\varphi^{\beta}_{2^j-1}(2^j(1-x))$. Since these $N-1$ functions form an orthogonal sequence, the corrections which make the $\psi_{j,k}\big|_{[0,1]}$ orthogonal to $V_j[0,1]$ are obvious, by Gram-Schmidt process. One thus obtains $N-1$ functions $2^{\frac{j}{2}}\tilde{h}_{2^j-2N+2}(2^j(1-x)), ..., 2^{\frac{j}{2}}\tilde{h}_{2^j-1}(2^j(1-x))$ where $\tilde{h}_{2^j-2N+2}, ..., \tilde{h}_{2^j-1}$ are independent of j.

By i) and iii) we have that $\psi_{j,k}\big|_{[0,1]}$ are orthogonal to $V_j[0,1]$, and ii) says that $\psi_{j,k}\big|_{[0,1]}$ for $-N+1 \leq l \leq 2^j - 2N + 1$, are mutually orthogonal it suffices to make the function in $2^{\frac{j}{2}}\tilde{h}_{2^j-2N+2}(2^j(1-x)), ..., 2^{\frac{j}{2}}\tilde{h}_{2^j-1}(2^j(1-x))$ mutually orthogonal, by means of the Gram-Schmidt process, to obtain the $N-1$ wavelets clustered at 1, namely $2^{\frac{j}{2}}\psi^{\beta}_{2^j-2N+2}(2^j(1-x)), ..., 2^{\frac{j}{2}}\psi^{\beta}_{2^j-N}(2^j(1-x))$.

We have obtained the following theorem concluding the construction of a wavelet basis of $L^2[0,1]$.

Theorem 2.14. *Let $V_j(\mathbb{R})$ be multiresolution analysis on $\mathbb{R}$ with scaling function φ and associated wavelet ψ. Let $j_0 \in \mathbb{Z}$ be such that $2^{j_0} \geq 4N-4$. The following collection*

$$
\begin{aligned}
&2^{\frac{j_0}{2}}\varphi^{\alpha}_{-2N+2}(2^{j_0}x), ..., 2^{\frac{j_0}{2}}\varphi^{\alpha}_{-1}(2^{j_0}x),\\
&\varphi_{j_0,k}(x)\big|_{[0,1]}, 0 \leq k \leq 2^{j_0} - 2N + 1,\\
&2^{\frac{j_0}{2}}\varphi^{\beta}_{2^{j_0}-2N+2}(2^{j_0}(1-x)), ..., 2^{\frac{j_0}{2}}\varphi^{\beta}_{2^{j_0}-1}(2^{j_0}(1-x))\\
&2^{\frac{j}{2}}\psi^{\alpha}_{-N+1}(2^j x), ..., 2^{\frac{j}{2}}\psi^{\alpha}_{-1}(2^j x),\\
&\psi_{j,k}(x)\big|_{[0,1]}, 0 \leq k \leq 2^j - 2N + 1,\\
&2^{\frac{j}{2}}\psi^{\beta}_{2^j-2N+2}(2^j(1-x)), ..., 2^{\frac{j}{2}}\psi^{\beta}_{2^j-N}(2^j(1-x)), j \geq j_0
\end{aligned}
$$

is an orthonormal basis of $L^2[0,1]$.

Remark 2.8. Recall the construction of the boundary wavelets using the Gram-Schmidt process. This construction gives, for $j \geq j_0$ and $x \in [0,1]$,

$$
\begin{aligned}
2^{\frac{j}{2}}\psi^{\alpha}_{-N+1}(2^j x) &= C_{-N+1}\psi_{j,-N+1}(x_i)\\
2^{\frac{j}{2}}\psi^{\alpha}_{-N+2}(2^j x) &= C^1_{-N+2}\psi_{j,-N+2}(x) + C^2_{-N+1}\psi_{j,-N+1}(x)\\
&\vdots\\
2^{\frac{j}{2}}\psi^{\alpha}_{-1}(2^j x) &= C^1_{-1}\psi_{j,-1}(x) + C^2_{-2}\psi_{j,-2}(x) + ... + C^{N-1}_{-N+1}\psi_{j,-N+1}(x)
\end{aligned}
$$

and

$$
\begin{aligned}
2^{\frac{j}{2}}\psi^{\beta}_{2^j-2N+2}(2^j(1-x)) &= C_{2^j-2N+2}\psi_{j,2^j-2N+2}(x_i)\\
2^{\frac{j}{2}}\psi^{\beta}_{2^j-2N+3}(2^j(1-x)) &= C^1_{2^j-2N+3}\psi_{j,2^j-2N+3}(x) + C^2_{2^j-2N+2}\psi_{j,2^j-2N+2}(x)\\
&\vdots\\
2^{\frac{j}{2}}\psi^{\beta}_{2^j-N}(2^j(1-x)) &= C^1_{2^j-N}\psi_{j,2^j-N}(x) + C^2_{2^j-N-1}\psi_{j,2^j-N-1}(x) + ...\\
&\quad + C^{N-1}_{2^j-2N+2}\psi_{j,2^j-2N+2}(x)
\end{aligned}
$$

Remark 2.9. If $\{f_j\}_{j\in J}$ is an orthonormal basis of $L^2(I)$, then

(1) elementary tensors, $\{f_j \otimes f_{\tilde{j}}\}_{j,\tilde{j}\in J} = \{f_j f_{\tilde{j}}\}_{j,\tilde{j}\in J}$ form an orthonormal basis of $L^2(I) \otimes L^2(I) \cong L^2(I^2)$.

(2) From (1) and induction we obtain that, for any multi-index set $J^d = \{(j_i, ..., j_d) : j_i \in J\}$, the collection $\{\prod_{i=1}^{d} f_{j_i}\}_{(j_1,...,j_d)\in J^d}$ is an orthonormal basis of $L^2(I^d)$.

CHAPTER III

RANDOM FIELDS

In this chapter, we introduce random fields and related notation. We also review the concept of correlation function and covariance function of random fields at different sample paths and then discuss some special classes of random fields with continuous sample paths. Furthermore, we review the power spectral density function of a random field. For further details, see Childers (1997) and Grimmett and Stirzaker (1992). Finally, we discuss the wavelet transform of random fields, additional details can be found in Cambanis and Houdre (1995).

3.1 Random Fields

A random field $X(t, \omega)$ is a function of two variables t and ω where t is called the coordinate position (or time for $t \in \mathbb{R}^+$) variable in standard terminology, and ω is the outcome variable which has several meanings according to the application under cosideration.

Definition 3.1. (Random Field)
Let $(\Omega, \mathcal{F}, P)$ be a probability (sample) space, T a parameter set and S a Borel subset of $\mathbb{R}$ or $\mathbb{C}$. A family $\{X_t\}_{t\in T}$ where $X_t : \Omega \to S$ is an $\mathcal{F}$-measurable random variable for all $t \in T$, is called a random field.

If $T = \mathbb{R}$ or $T = [0, \infty)$, the random field is usually called a stochastic process or random process. The term random field is usually used to stress that the dimension of the coordinate space is higher than one.

In this thesis we consider random fields whose coordinate space is $T = \mathbb{R}^d$, that is X_t is a real or complex random variable for any $t \in \mathbb{R}^d$, and require that $X(t, \omega)$ is measurable on $\mathbb{R}^d \times \Omega$.

Remark 3.1. We can look at a random field in different ways.

(1) Fix $t \in T$. Then the map $\omega \longmapsto X_t(\omega)$ is a random variable, as in the definition, it is a function on the sample space. Thus, a random field is an ensemble of random variables over some coordinate space T.

(2) Fix $\omega \in \Omega$. Then the map $t \longmapsto X_t(\omega)$ is a deterministic function from T to S, usually called the sample path of X at ω.

Definition 3.2. (Expected Value, Variance and Covariance Values)
Let $(\Omega, \mathcal{F}, P)$ be a probability space and X, Y real or complex random variables. The expected value of $X \in L^1(\Omega)$ is defined by

$$E[X] = \int_{\Omega} X(\omega)\, dP(\omega)$$

The variance of $X \in L^2(\Omega)$ is defined by

$$\mathrm{Var}[X] = E[|X - E[X]|^2] = \int_{\Omega} |X(\omega) - E[X]|^2\, dP(\omega) = E[|X|^2] - |E[X]|^2.$$

The covariance of $X, Y \in L^2(\Omega)$ is defined by

$$\mathrm{Cov}[X, Y] = E[(X - E[X])\overline{(Y - E[Y])}].$$

Variance and covariance possess the following properties:

1) $\mathrm{Cov}[X, Y] = E[X\overline{Y}] - E[X]\overline{E[Y]}$

2) $\mathrm{Cov}[X, Y] = \overline{\mathrm{Cov}[Y, X]}$

3) $\mathrm{Cov}[aX + bY, Z] = a\mathrm{Cov}[X, Z] + b\mathrm{Cov}[Y, Z]$ for all scalars a, b

4) $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{Re}\mathrm{Cov}[X, Y]$.

Definition 3.3. (Second Order Random Field)

A random field $\{X_t\}_{t\in\mathbb{R}^d}$ is said to be a second order random field if $E[|X_t|^2] < \infty$ for all $t \in \mathbb{R}^d$.

Definition 3.4. (Distribution Function)

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and X a vector valued random variable, $X : \Omega \to \mathbb{R}^d$. Set $\mu_X(A) = P\{\omega \in \Omega : X(\omega) \in A\} = P(X^{-1}(A))$ for $A \in \mathcal{B}(\mathbb{R}^d)$. Then μ_X is a probability measure on $\mathcal{B}(\mathbb{R}^d)$, and

$$\int_{\mathbb{R}^d} h\, d\mu_X = \int_\Omega h \circ X\, dP(\omega) \quad \text{for all } h : \mathbb{R}^d \longrightarrow \mathbb{R} \tag{3.1}$$

whenever one of these integrals is defined. μ_X is called the distribution of X. More details can be found in Capiński and Marek (2004).

Define

$$\begin{aligned} F_X(\lambda) &= \mu_X((-\infty, \lambda]) \\ &= \mu_X((-\infty, \lambda_1] \times ... \times (-\infty, \lambda_d]) \quad \text{ for all } \ \lambda = (\lambda_1, ..., \lambda_d) \in \mathbb{R}^d. \end{aligned}$$

Then $F_X : \mathbb{R}^d \longrightarrow [0, 1]$ is called the distribution function of X.

Definition 3.5. (Density Function)

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and $X : \Omega \longrightarrow \mathbb{R}^d$ a random variable with distribution μ_X. Suppose that μ_X is absolutely continuous with respect to the Lebesgue measure λ. Then by Radon Nikodym's theorem, there exists a nonnegative integrable function f_X such that

$$\mu_X(A) = \int_A f_X(x)\, d\lambda(x) \qquad \text{for all } \ A \in \mathcal{B}(\mathbb{R}^d). \tag{3.2}$$

If F_X is the distribution function of X, we have for $t = (t_1, t_2, ..., t_d) \in \mathbb{R}^d$,

$$F_X(t) = \int_{(-\infty,t_1]} ... \int_{(-\infty,t_d]} f_X(x_1, ..., x_d)\, d\lambda(x_1)...d\lambda(x_d),$$

and $f = f_X$ is called the (joint) density function of μ_X. One can show that there is a one-to-one correspondence between distribution functions F_X and distribution measures μ_X.

Example 3.1. (Normal Distribution)

Let us start with parameters $m \in \mathbb{R}$ and $s > 0$, and consider the density function $f : \mathbb{R} \longrightarrow \mathbb{R}^+$ defined by

$$f(t) = \frac{1}{\sqrt{2\pi}s} e^{-\frac{(t-m)^2}{2s^2}}.$$

Set $F(t) = \int_{-\infty}^{t} f(x)\,dx$; then F is a distribution function, and determines a distribution μ by $\mu(-\infty, t] = F(t)$. Using (3.1) and (3.2) one obtains that any random variable X whose distribution function is of this type will have mean m and variance s^2; we therefore call μ the normal or Gaussian distribution with mean m and variance s^2, and write "X is $N(m, s^2)$".

Definition 3.6. (Identically Distributed)

Let X and Y be random variables on probability spaces $(\Omega, \mathcal{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, respectively. If $\mu_X = \mu_Y$, then X and Y are said to be identically distributed, written $X \stackrel{d}{=} Y$.

There are various ways of interpreting the statement $X_n \to X$ as $n \to \infty$, some can be found in the next definition.

Definition 3.7. (Modes of Convergence of Random Variables)

Let $X_1, X_2, ...$ and X be random variables on some probability space $(\Omega, \mathcal{F}, P)$. We say that:

(a) X_n converges to X almost surely, written as $X_n \stackrel{a.s.}{\to} X$, if $\{\omega \in \Omega : X_n(\omega) \to X(\omega) \text{ as } n \to \infty\}$ is an event whose probability is one.

(b) X_n converges to X in r^{th} mean, where $r \geq 1$, written as $X_n \stackrel{r}{\to} X$, if $E[|X|^r], E[|X_n|^r] < \infty$ for all n and $E[|X_n - X|^r] \to 0$ as $n \to \infty$.

Remark 3.2. Minkowski's inequality states that

$$\|X\|_r = (E[|X|^r])^{\frac{1}{r}}$$

defines a norm on the collection of random variables with finite r^{th} mean, for any value of $r \geq 1$. Of most use are the values $r = 1$ and $r = 2$.

If $X_n \xrightarrow{1} X$ we say that X_n converges to X in mean,

and

if $X_n \xrightarrow{2} X$ we say that X_n converges to X in mean square.

Definition 3.8. (Independent Random Variables)
A collection of random variables $\{X_n\}_{n=1}^N$ on a probability space $(\Omega, \mathcal{F}, P)$ is called independent, if for any k, $1 \leq k \leq N$, and any choice of Borel sets $B_{n_1}, ..., B_{n_k}$

$$P\left(X_{n_1}^{-1}(B_{n_1}) \cap ... \cap X_{n_k}^{-1}(B_{n_k})\right) = P\left(X_{n_1}^{-1}(B_{n_1})\right) ... P\left(X_{n_k}^{-1}(B_{n_k})\right).$$

Remark 3.3. Let X_1, X_2 be two independent vector valued random variables on $(\Omega, \mathcal{F}, P)$, say $X_1 : \Omega \longrightarrow \mathbb{R}^d$, $X_2 : \Omega \longrightarrow \mathbb{R}^l$. Consider the product of the measurable spaces $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $(\mathbb{R}^l, \mathcal{B}(\mathbb{R}^l))$, denoted by $(\mathbb{R}^d \times \mathbb{R}^l, \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R}^l))$. Next, consider $X : \Omega \to \mathbb{R}^{d+l}$ given by $X(\omega) = (X_1(\omega), X_2(\omega))$. It is easy to see that X is a random variable, that is $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathcal{R}^l) = \mathcal{B}(\mathbb{R}^{d+l})$-measurable. Let $\mu_{X_1}, \mu_{X_2}, \mu_X$ denote the distributions of X_1, X_2 and X respectively.
Then X_1 and X_2 are independent

if and only if $P(X_1^{-1}(B_1) \cap X_2^{-1}(B_2)) = P(X_1^{-1}(B_1))P(X_2^{-1}(B_2))$

if and only if $P(X^{-1}(B_1 \times B_2)) = P(X_1^{-1}(B_1))P(X_2^{-1}(B_2))$

if and only if $\mu_X(B_1 \times B_2) = \mu_{X_1}(B_1)\mu_{X_2}(B_2)$,

for all $B_1 \in \mathcal{B}(\mathbb{R}^d), B_2 \in \mathcal{B}(\mathbb{R}^l)$. That is, X_1 and X_2 are independent if and only if $\mu_X = \mu_{X_1} \times \mu_{X_2}$.

Next, let $f(x) : \mathbb{R}^d \longrightarrow \mathbb{R}$, $g(y) : \mathbb{R}^l \longrightarrow \mathbb{R}$ be Borel functions and consider the Borel function $f(x)g(y) : \mathbb{R}^{d+l} \to \mathbb{R}$. Suppose that either $f \geq 0$, $g \geq 0$

or $f(X_1), g(X_2) \in L^1(\Omega)$. Then we can see that $f(X_1)g(X_2) \geq 0$, respectively $f(X_1)g(X_2) \in L^1(\Omega)$. In fact, by Tonelli's Theorem and (3.1),

$$\begin{aligned} E[|f(X_1)g(X_2)|] &= \int_{\mathbb{R}^d \times \mathbb{R}^l} |f(x_1)g(x_2)|\, d_{\mu_{X_1} \times \mu_{X_2}}(x_1, x_2) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^l} |f(x_1)||g(x_2)|\, d_{\mu_{X_1}}(x_1) d_{\mu_{X_2}}(x_2) \\ &= E[|f(X_1)|]E[|g(X_2)|]. \end{aligned}$$

which shows that $f(X_1)g(X_2) \in L^1(\Omega)$ provided that $f(X_1), g(x_2) \in L^1(\Omega)$. Then we have have by Fubini's Theorem, in a similar way,

$$E[f(X_1)g(X_2)] = \int_{\mathbb{R}^d} \int_{\mathbb{R}^l} f(x_1)g(x_2)\, d_{\mu_{X_1}}(x_1)\, d_{\mu_{X_2}}(x_2) = E[f(X_1)]E[g(X_2)].$$

We summarize as following result:

Theorem 3.1. *Let X_1 and X_2 be independent random variables on $(\Omega, \mathcal{F}, P)$, f and g Borel functions. If $f \geq 0$, $g \geq 0$ or $f(X_1), g(X_2) \in L^1(\Omega)$, then $f(X)g(Y) \geq 0$, respectively $f(X)g(Y) \in L^1(\Omega)$ and $E[f(X_1)g(X_2)] = E[f(X_1)]E[g(X_2)]$.*

The following result is an immediate consequence of the above theorem by choosing $f(x_1) = x_1$ and $g(x_2) = x_2$.

Corollary 3.2. *Let X_1, X_2 be independent random variables $X_i : \Omega \to \mathbb{R}$ such that $X_1, X_2 \in L^1(\Omega)$. Then $X_1X_2 \in L^1(\Omega)$ and $E[X_1X_2] = E[X_1]E[X_2]$.*

3.2 The Correlation and Covariance Functions

In many applications of random fields, it is necessary to consider the relationships between sample paths starting at different positions. To obtain useful tools for analysing a pair of random variables, we recall that the correlation and the covariance of a pair of random variables is random at different positions t and $t + \tau$. To use this information to understand a pair of random variables, we therefore work with the correlation and covariance of the random variables X_t and $X_{t+\tau}$.

Definition 3.9. (Auto-correlation and Auto-covariance Functions)
Let $\{X_t\}_{t\in\mathbb{R}^d}$ be a random field such that $X_t \in L^2(\Omega)$ for all $t \in \mathbb{R}^d$.
The auto-correlation of X_t is defined by

$$R_X(t, t+\tau) = E[X_t\overline{X_{t+\tau}}] \quad \text{for} \quad t, \tau \in \mathbb{R}^d. \tag{3.3}$$

The auto-covariance of X_t is defined by

$$C_X(t, t+\tau) = E[(X_t - E[X_t])\,\overline{(X_{t+\tau} - E[X_{t+\tau}])}] \quad \text{for} \quad t, \tau \in \mathbb{R}^d. \tag{3.4}$$

The auto-correlation and auto-covariance functions of a random field are related to each other in the following way

$$C_X(t, t+\tau) = R_X(t, t+\tau) - E[X_t]\overline{E[X_{t+\tau}]}. \tag{3.5}$$

This property can be easily derived from equations (3.3) and (3.4). We can see that for a zero mean random field, the auto-covariance function coincides with the auto-correlation function.

So far we have considered the relationship between two sample paths of the same random field at different positions. We now address a similar relationship between two sample paths of different random field at different positions.

Definition 3.10. (The Cross-Correlation Function)
Let $\{X_t\}_{t\in\mathbb{R}^d}$ and $\{Y_t\}_{t\in\mathbb{R}^d}$ be random fields such that $X_t, Y_t \in L^2(\Omega)$ for all $t \in \mathbb{R}^d$.
The cross-correlation function of X_t and Y_t is defined by

$$R_{XY}(t, t+\tau) = E[X_t\overline{Y_{t+\tau}}]. \tag{3.6}$$

The cross-covariance is defined as

$$C_{XY}(t, t+\tau) = E\left[(X_t - E[X_t])\,\overline{(Y_{t+\tau} - E[Y_{t+\tau}])}\right]. \tag{3.7}$$

The cross-correlation and cross-covariance functions of two random fields are related to each other in the following way, see Childers (1999) for example.

$$C_{XY}(t, t+\tau) = R_{XY}(t, t+\tau) - E[X_t]\overline{E[Y_{t+\tau}]}. \tag{3.8}$$

This property can be easily derived from equations (3.6) and (3.7).

3.3 Classification of Random Fields

The mean and correlation functions can provide information about the spatial structure of a random field. In this section we examine three particular classes of stationary random fields with continuous sample paths.

Definition 3.11. (Strongly Stationary Random Field)
A random field $\{X_t\}_{t\in\mathbb{R}^d}$ is called strongly (or narrow sense or first-order) stationary if the families of random variables $\{X_{t_1}, X_{t_2}, ..., X_{t_n}\}$ and $\{X_{t_1+h}, ..., X_{t_n+h}\}$ for all $n \in \mathbb{N}$ for all $t_1, t_2, ..., t_n \in \mathbb{R}^d$ and $h \in \mathbb{R}^d$, have the same joint distribution.

In particular, mean, variance, auto-correlation etc. are independent of t. This motivates the following definition:

Definition 3.12. (Weakly Stationary Random Field)
A random field $\{X_t\}_{t\in\mathbb{R}^d}$ is called weakly (or wide sense or second order) stationary if $E[X_t] = m$, where m is a constant for all $t \in \mathbb{R}^d$ and

$$R_X(t, t+\tau) = R_X(\tau) \quad \text{for all} \quad t, \tau \in \mathbb{R}^d. \tag{3.9}$$

All strongly stationary random fields are also weakly stationary, provided the mean and auto-correlation functions exist. We say that if a random field is not weakly stationary, then it is non-stationary. The following definition introduces interesting examples of non-stationary random fields.

Definition 3.13. (Strongly / Weakly Stationary Increments)
A random field $\{X_t\}_{t\in\mathbb{R}^d}$ is said to have **strongly stationary increments** if the probability distribution of any increment $\Delta X(t;h) = X_{t+h} - X_t$ depends only on h for all $h \in \mathbb{R}^d$, to possess **weakly stationary increments** if for all t, $E[\Delta X(t;h)]$ depends on h only, and $R_{\Delta X(t;h_1)\Delta X(t+\tau;h_2)} = E[(X_{t+h_1} - X_t)(X_{t+\tau+h_2} - X_{t+\tau})]$ depends only on h_1, h_2 and τ, for all $h_1, h_2, \tau \in \mathbb{R}^d$.

For example, a fractional Brownian field $\{B_t^H\}_{t\in\mathbb{R}^d}$ with Hurst index $0 < H < 1$ (for a precise definition, see section 3.5) is a random field with zero mean and autocorrelation

$$R_{B_t^H}(t, t+\tau) = \frac{V_H}{2}\left[\|t\|^{2H} + \|t+\tau\|^{2H} - \|\tau\|^{2H}\right] \tag{3.10}$$

where $V_H = E[(B_1^H)^2]$. We can see that the fractional Brownian field is not a stationary random field because the auto-correlation function $R_{B_t^H}(t, t+\tau)$ depends not only on $\|\tau\|$ alone. We can however, show that a fractional Brownian field is a stationary increments random field, see more details in Appendix F. In a similar way, one defines:

Definition 3.14. (Jointly Strongly Stationary Random Field)
Two random fields, $\{X_t\}_{t\in\mathbb{R}^d}$ and $\{Y_t\}_{t\in\mathbb{R}^d}$, are jointly strongly stationary if the families $\{X_{t_1}, X_{t_2}, ..., X_{t_n}\}$ and $\{Y_{t_1+h}, Y_{t_2+h}, ..., Y_{t_n+h}\}$, for all $n \in \mathbb{N}$, $t_1, t_2, ..., t_n \in \mathbb{R}^d$ and $h \in \mathbb{R}^d$, have the same joint distribution.

Note that if two random fields are jointly strongly stationary, then each is individually strongly stationary.

Definition 3.15. (Jointly Weakly Stationary Random Field)
Two random fields, $\{X_t\}_{t\in\mathbb{R}^d}$ and $\{Y_t\}_{t\in\mathbb{R}^d}$, are jointly weakly stationary if they satisfy the following relationships:

(1) Both X_t and Y_t are individually weakly stationary with identical means.

(2) $R_{XY}(t, t+\tau) = R_{XY}(\tau)$ for all $t \in \mathbb{R}^d$ and $\tau \in \mathbb{R}^d$.

One can establish that if two random fields are jointly strongly stationary, then they are jointly weakly stationary. But, the converse is not true. Also by definition, if two random fields are jointly weakly stationary, then they are individually weakly stationary. However, the converse is not true in general.

3.4 The Power Spectral Density Function

The class of weakly stationary random fields is important as the autocorrelation function $R_X(t, t+\tau)$ is simply a function of the position difference τ, $R_X(t, t+\tau) = R_X(\tau)$. Next we remark on properties of the auto-correlation function $R_X(\tau)$ of a weakly stationary random field which will be required later, for an application of Bochner's Theorem.

Remark 3.4. Let $\{X_t\}_{t\in\mathbb{R}^d}$ be a weakly stationary random field. The following properties can be shown to be true :

(a) $R_X(-\tau) = \overline{R_X(\tau)}$. Simply apply definition 3.9 and equation (3.3) and change the variable t to $t - \tau$.

(b) $|R_X(\tau)| \leq R_X(0) = E[|X_t|^2]$. In fact, using Hölder's inequality $E[|X\overline{Y}|]^2 \leq E[|X|^2]E[|Y|^2]$, we have $|R_X(\tau)|^2 \leq E[|X_t\overline{X_{t+\tau}}|]^2 \leq E[|X|_t^2]E[|X_{t+\tau}|^2] = R_X^2(0)$.

(c) By equation (3.5) we have $R_X(\tau) = C_X(\tau) + |E[X_t]|^2$.

(d) $R_X(\tau)$ is a positive definite function. In fact, for $\{z_i\}_{i=1}^n \subset \mathbb{C}$,

$$\sum_{j,k=1}^{n} z_j\overline{z_k}R_X(\tau_k - \tau_j) = \sum_{j,k=1}^{n} z_j\overline{z_k}E[X_t\overline{X_{t+\tau_k-\tau_j}}] = \sum_{j,k=1}^{n} z_j\overline{z_k}E[X_{t+\tau_j}\overline{X_{t+\tau_k}}]$$

$$= \sum_{j,k=1}^{n} z_j\overline{z_k}\left(\mathrm{Cov}[X_{t+\tau_j}, X_{t+\tau_k}] + E[X_{t+\tau_j}]\overline{E[X_{t+\tau_k}]}\right),$$

that is

$$\sum_{j,k=1}^{n} z_j\overline{z_k}R_X(\tau_k-\tau_j) = \sum_{j,k=1}^{n} \text{Cov}[z_jX_{t+\tau_j}, z_kX_{t+\tau_k}] + \sum_{j,k=1}^{n} z_j\overline{z_k}E[X_{t+\tau_j}]\overline{E[X_{t+\tau_k}]}$$
$$= \text{Cov}[\sum_{j=1}^{n} z_jX_{t+\tau_j}, \sum_{k=1}^{n} z_kX_{t+\tau_k}] + \left|\sum_{j=1}^{n} z_jE[X_{t+\tau_j}]\right|^2 \geq 0.$$

Let $\{X_t\}_{t\in\mathbb{R}^d}$ be a weakly stationary random field. By Remark 3.4 (b) and (d), we may apply Bochner's Theorem (Theorem 2.6), to obtain

$$R_X(\tau) = \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}\, dF_X(\lambda) \quad \text{a.e.} \tag{3.11}$$

for some bounded Borel measure $F_X(\lambda)$, which is called spectral measure. Now if the measure F_X is absolutely continuous with regards to the Lebesgue measure, then we define the power spectral density $S_X(\lambda)$ as the generalized Fourier transform as follows.

Definition 3.16. Let $\{X_t\}_{t\in\mathbb{R}^d}$ be a weakly stationary random field with autocorrelation function $R_X(\tau)$. The power spectral density function $S_X(\lambda)$ is defined by the generalised inverse Fourier transform and its generalized Fourier transform as follows,

$$S_X(\lambda) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i\tau\cdot\lambda}R_X(\tau)\, d\tau. \tag{3.12}$$

and

$$R_X(\tau) = \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}S_X(\lambda)\, d\lambda \tag{3.13}$$

Thus, by (3.11) the power spectral density function is in fact a finite Borel measure. However, if R_X is integrable, then S_X is a continuous, non-negative function.

The power spectral density function is known by several names, including energy spectrum, spectral density, spectrum, and perhaps most commonly as simply the power spectrum.

Remark 3.5. The definition of the power spectral density function as above is known as the Winner-Khintchine Theorem or the Einstein-Wiener-Khintchine Theorem. We shall refer to this method of calculating the spectral density function as the correlation method.

An alternative definition for the power spectral density function is arrived at by noting the following. Assuming that the Fourier transform of the sample paths of a random variable X_t exists over a range $[-T,T]^d$, then we have the random variable

$$\mathcal{F}_T[X_t](\xi) = \int_{[-T,T]^d} X(t)e^{-i\xi\cdot t}\,dt$$

where $\mathcal{F}_T$ denotes the Fourier transform over the range $[-T,T]^d$.
The magnitude squared of this random variable is

$$\begin{aligned}|\mathcal{F}_T[X_t](\xi)|^2 &= \mathcal{F}_T[X_t](\xi)\overline{\mathcal{F}_T[X_t](\xi)}\\ &= \int_{[-T,T]^d} X_{t_1}e^{-i\xi\cdot t_1}\,dt_1 \int_{[-T,T]^d} \overline{X_{t_2}}e^{i\xi\cdot t_2}\,dt_2\\ &= \int_{[-T,T]^d}\int_{[-T,T]^d} X_{t_1}\overline{X_{t_2}}e^{i\xi\cdot(t_2-t_1)}\,dt_1dt_2.\end{aligned}$$

If we take the expectation and divide by $(4\pi T)^d$, we have by Fubini's Theorem

$$\begin{aligned}\frac{1}{(4T)^d}E\left[|\mathcal{F}_T[X_t](\xi)|^2\right] &= \frac{1}{(4\pi T)^d}E\left[\int_{[-T,T]^d}\int_{[-T,T]^d} X_{t_1}\overline{X_{t_2}}e^{i\xi\cdot(t_2-t_1)}\,dt_1dt_2\right]\\ &= \frac{1}{(4\pi T)^d}\int_{[-T,T]^d}\int_{[-T,T]^d} E\left[X_{t_1}\overline{X_{t_2}}\right]e^{i\xi\cdot(t_2-t_1)}\,dt_1dt_2\\ &= \frac{1}{(4\pi T)^d}\int_{[-T,T]^d}\int_{[-T,T]^d} R_X(t_2-t_1)e^{i\xi\cdot(t_2-t_1)}\,dt_1dt_2.\end{aligned}$$

For each $k = 1, 2, ..., d$, if we let $\tau^k = t_2^k - t_1^k$ and $u^k = t_1^k + t_2^k$,
then $dt_1^k\,dt_2^k = (J^k)^{-1}du\,d\tau$ where J^k is the Jacobian which is

$$|J^k| = \begin{vmatrix}\frac{\partial u^k}{\partial t_1^k} & \frac{\partial u^k}{\partial t_2^k}\\ \frac{\partial \tau^k}{\partial t_1^k} & \frac{\partial \tau^k}{\partial t_2^k}\end{vmatrix} = \begin{vmatrix}1 & 1\\ -1 & 1\end{vmatrix} = 2.$$

Then we have

$$
\begin{aligned}
&\frac{1}{(4\pi T)^d} E\left[|\mathcal{F}_T[X_t](\xi)|^2\right] \\
&\qquad = \frac{1}{(4\pi T)^d} \int_{[-2T,2T]^d} \left(\frac{1}{2}\right)^d \int_{-(2T-|\tau^1|)}^{2T-|\tau^1|} \ldots \int_{-(2T-|\tau^d|)}^{2T-|\tau^d|} R_X(\tau) e^{i\xi\cdot\tau}\, du\, d\tau \\
&\qquad = \frac{1}{(8\pi T)^d} \int_{[-2T,2T]^d} R_X(\tau) e^{i\xi\cdot\tau} \prod_{i=1}^{d} \int_{-(2T-|\tau^i|)}^{2T-|\tau^i|} 1\, du^i\, d\tau \\
&\qquad = \frac{1}{(8\pi T)^d} \int_{[-2T,2T]^d} R_X(\tau) e^{i\xi\cdot\tau} \prod_{i=1}^{d} (4T - 2|\tau^i|)\, d\tau \\
&\qquad = \frac{1}{(8\pi T)^d} \int_{[-2T,2T]^d} (4T)^d \prod_{i=1}^{d} (1 - \frac{|\tau^i|}{2T}) R_X(\tau) e^{i\xi\cdot\tau}\, d\tau \\
&\qquad = \frac{1}{(2\pi)^d} \int_{[-2T,2T]^d} \prod_{i=1}^{d} (1 - \frac{|\tau^i|}{2T}) R_X(\tau) e^{i\xi\cdot\tau}\, d\tau.
\end{aligned}
$$

If we take the limit as T goes to infinity, as for each i we have $|\tau^i| \leq 2T$ so that $0 \leq 1 - \dfrac{|\tau^i|}{2T} \leq 1$, then we have by the Dominated Convergence Theorem,

$$
\lim_{T\to\infty} \frac{1}{(4\pi T)^d} E\left[|\mathcal{F}_T[X_t](\xi)|^2\right] = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} R_X(\tau) e^{i\xi\cdot\tau}\, d\tau = S_X(\xi)
$$

provided that the auto-correlation function $R_X(\tau)$ is integrable and hence we are led to an alternative definition for the power spectral density function, namely:

Definition 3.17. (Alternative Definition of the Power Spectral Density Function) The power spectral density function can be defined as

$$
S_X(\lambda) = \lim_{T\to\infty} \frac{1}{(4\pi T)^d} E\left[|\mathcal{F}_T[X_t](\lambda)|^2\right]
$$

This method of calculating the power spectral density function is called the direct method and is equivalent to the correlation method.

Let $\{X_t\}_{t\in\mathbb{R}^d}$ and $\{Y_t\}_{t\in\mathbb{R}^d}$ be jointly weakly stationary random fields. We can now introduce the power cross spectral density Function in a similar way. For every value of the argument τ, similar to Remark 3.4, the cross-correlation function $R_{XY}(\tau)$ has the following properties:

(a) $R_{XY}(\tau) = \overline{R_{YX}(-\tau)}$

(b) $|R_{XY}(\tau)| \leq \frac{1}{2}[R_{XX}(0) + R_{YY}(0)]$

(c) $|R_{XY}(\tau)|^2 \leq R_{XX}(0)R_{YY}(0)$.

The power spectral density function for the auto-correlation has been defined. In a similar manner we define the cross-spectral density function.

Definition 3.18. (The Power Cross Spectral Density Function)

If X_t and Y_t are jointly weakly stationary random fields with respectively auto-correlation $R_X(\tau)$ and $R_Y(\tau)$ and cross-correlation $R_{XY}(\tau)$, then power cross spectral density function is defined as the generalized Fourier transform of the cross-correlation; that is

$$S_{XY}(\lambda) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\tau\cdot\lambda} R_{XY}(\tau)\, d\tau$$

and

$$R_{XY}(\tau) = \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda} S_{XY}(\lambda)\, d\lambda.$$

The term power cross-spectral density is often abbreviated to cross-spectral density.

Some properties of the cross-spectral density are:

(1) The cross-spectral density, $S_{XY}(\lambda)$, is not necessarily real, since $R_{XY}(\tau)$ is not necessarily even.

(2) $S_{XY}(\lambda) = S_{YX}(-\lambda)$, $\overline{S_{XY}(\lambda)} = S_{XY}(-\lambda)$.

Indeed, since $R_{XY}(\tau) = R_{YX}(-\tau)$ we have

$$S_{XY}(\lambda) = \int_{\mathbb{R}^d} R_{YX}(-\tau) e^{-i\lambda\cdot\tau}\, d\tau =^{\tau\mapsto-\tau} \int_{\mathbb{R}^d} R_{YX}(\tau) e^{-i(-\lambda)\cdot\tau}\, d\tau = S_{YX}(-\lambda).$$

It is also true that

$$\overline{S_{XY}(\lambda)} = \int_{\mathbb{R}^d} \overline{R_{XY}(\tau)} e^{i\lambda\cdot\tau}\, d\tau = S_{XY}(-\lambda)$$

and hence $S_{XY}(\lambda) = \overline{S_{XY}(-\lambda)}$ which is known as Hermitian symmetry.

3.5 Fractional Brownian Field and Fractional Brownian Sheet

In this section, we present the definition of fractional Brownian field. Further definitions, details and important properties on Brownian fields can be found in Appendix F, and on Brownian motion in Appendix E. Moreover, we review the definition of a fractional Brownian sheet, in particular, a Brownian sheet which we will use in our main results of Section 6.2 and 6.3.

Definition 3.19. (Fractional Brownian Field)
For a given real value $H \in (0,1)$, a fractional Brownian field $\{B_t^H\}_{t\in\mathbb{R}^d}$ with Hurst index H is a Gaussian random field (see Appendix D) with zero mean and covariance function given by

$$\mathrm{Cov}[B_s^H, B_t^H] = \frac{V_H}{2}\left(\|s\|^{2H} + \|t\|^{2H} - \|s-t\|^{2H}\right) \tag{3.14}$$

where $V_H = E[(B_1^H)^2]$, for $s, t \in \mathbb{R}^d$.
In case $d = 1$, $\{B_t^H\}_{t\in\mathbb{R}}$ is called fractional Brownian motion. In case $d = 1$ and $H = \frac{1}{2}$, $\{B_t^{\frac{1}{2}}\}_{t\in\mathbb{R}}$ is Brownian motion provided that sample paths are continuous, as in the next definition.

Definition 3.20. (Brownian Motion)
A stochastic process $\{B_t\}_{t\geq 0}$ is called Brownian motion, if

(B1) the sample paths are continuous, that is for each $\omega \in \Omega, t \longmapsto B_t(\omega)$ is continuous,

(B2) the increments of B_t are independent, that is for any finite set of times $0 < t_1 < t_2 < ... < t_n$ the random variables $B_{t_1}, B_{t_2} - B_{t_1}, ..., B_{t_n} - B_{t_{n-1}}$ are independent,

(B3) for any $0 \leq s < t < \infty$, the increment $B_t - B_s$ has Gaussian distribution with zero mean and variance $t-s$, that is $B_t - B_s$ is $N(0, t-s)$.

If in addition, $B_0 = 0$ then $\{B_t\}_{t\geq 0}$ is called standard Brownian motion.

Remark 3.6. If $\{B_t\}_{t\geq 0}$ is a standard Brownian motion, then it is a Gaussian process with zero mean and $\mathrm{Cov}[B_s, B_t] = \min(s,t)$; see further details in Theorem E.1 and for a converse, see Lemma E.2 in Appendix E.

Definition 3.21. (Fractional Brownian Sheet)

For a given vector $H = (H_1, H_2, ..., H_d) \in (0,1)^d$, a fractional Brownian sheet $\{B_t^H\}_{t\in\mathbb{R}^d}$ with Hurst index H is a real-valued, Gaussian field with zero mean and covariance function given by

$$\mathrm{Cov}[B_s^H, B_t^H] = \prod_{i=1}^{d} \frac{1}{2}\left(|s_i|^{2H_i} + |t_i|^{2H_i} - |s_i - t_i|^{2H_i}\right) \tag{3.15}$$

where $s = (s_1, ...s_d), t = (t_1, ..., t_d) \in \mathbb{R}^d$.

Remark 3.7. In case $H_i = \frac{1}{2}$ for all i, $\{B_t^H\}_{t\in\mathbb{R}^d}$ is called a Brownian sheet and we have

$$\mathrm{Cov}[B_s, B_t] = \prod_{i=1}^{d} \frac{1}{2}\left(|s_i| + |t_i| - |s_i - t_i|\right).$$

Moreover, for $\{B_t^H\}_{t\in[0,\infty)^d}$ in case $H_i = \frac{1}{2}$ for all i, we can see that

$$\mathrm{Cov}[B_s, B_t] = \prod_{i=1}^{d} \min(s_i, t_i).$$

3.6 Wavelet Transform of Random Fields

Let H be a matrix group and φ a real or complex valued mother wavelet function. For each $a \in H$ and $b \in \mathbb{R}^d$, following the Definition 2.10, the continuous wavelet transform of a second order random field $\{X_t(\omega)\}_{t\in\mathbb{R}^d}$ is defined by

$$CW_X^a(b,\omega) = |\det a|^{-\frac{1}{2}} \int_{\mathbb{R}^d} X_u(\omega)\overline{\varphi(a^{-1}(u-b))}\, du \tag{3.16}$$

provided this integral exists with probability one.

Remark 3.8. We can see that, if $\{X_t\}_{t\in\mathbb{R}^d}$ such that $E[|X_t|]$ is bounded, then the integral (3.16) exists with probability one. Indeed, formally

$$E[|CW_X^a(b,\omega)|] = E\left[\left|\int_{\mathbb{R}^d} |\det a|^{-\frac{1}{2}} X_u(\omega)\overline{\varphi(a^{-1}(u-b))}\, du\right|\right]$$
$$\leq |\det a|^{-\frac{1}{2}} E\left[\int_{\mathbb{R}^d} |X_u(\omega)| \left|\overline{\varphi(a^{-1}(u-b))}\right|\, du\right] \qquad (3.17)$$

By Fubini's Theorem we have

$$E[|CW_X^a(b,\omega)|] \leq |\det a|^{-\frac{1}{2}} \int_{\mathbb{R}^d} E\left[|X_u(\omega)|\right] \left|\overline{\varphi(a^{-1}(u-b))}\right|\, du.$$

Since $\{X_t\}_{t\in\mathbb{R}^d}$ has constant mean we get that

$$E[|CW_X^a(b,\omega)|] \leq |\det a|^{\frac{1}{2}} E\left[|X_u(\omega)|\right] \|\varphi\|_1 < \infty.$$

Next, consider

$$E\left[|CW_X^a(b,\omega)|^2\right] = E\left[|CW_X^a(b,\omega)\overline{CW_X^a(b,\omega)}|\right]$$
$$\leq E\left[\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |\det a|^{-1} |X_\xi(\omega)\overline{X_\eta(\omega)}||\overline{\varphi(a^{-1}(\xi-b))}\varphi(a^{-1}(\eta-b))|\, d\xi\, d\eta\right].$$

Since $E[|X_t|^2] < \infty$ (as $\{X_t\}_{t\in\mathbb{R}^d}$ is a second order random field), we have by Hölder's inequality,

$$E\left[|CW_X^a(b,\omega)|^2\right]$$
$$\leq |\det a|^{-1} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} E[|X_\xi(\omega)\overline{X_\eta(\omega)}|]|\overline{\varphi(a^{-1}(\xi-b))}\varphi(a^{-1}(\eta-b))|\, d\xi\, d\eta$$
$$\leq |\det a|^{-1} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \left(E[|X_\xi|^2]\right)^{\frac{1}{2}} \left(E[|X_\eta|^2]\right)^{\frac{1}{2}} |\overline{\varphi(a^{-1}(\xi-b))}\varphi(a^{-1}(\eta-b))|\, d\xi\, d\eta$$
$$= |\det a|^{-1} \int_{\mathbb{R}^d} \left(E[|X_\xi|^2]\right)^{\frac{1}{2}} |\overline{\varphi(a^{-1}(\xi-b))}|\, d\xi \int_{\mathbb{R}^d} \left(E[|X_\eta|^2]\right)^{\frac{1}{2}} |\varphi(a^{-1}(\eta-b))|\, d\eta$$
$$= |\det a|^{-1} \left|\int_{\mathbb{R}^d} \left(E[|X_\xi|^2]\right)^{\frac{1}{2}} |\varphi(a^{-1}(\xi-b))|\, d\xi\right|^2.$$

We have $E\left[|CW_X^a(b,\omega)|^2\right] < \infty$ under each of the following conditions:

(C1) If $E[\int_{\mathbb{R}^d} |X_t(\omega)|^2\, dt] < \infty$ and $\varphi \in L^2(\mathbb{R}^d)$, then by the Cauchy-

Schwartz inequality;

$$E\left[|CW_X^a(b,\omega)|^2\right] \leq |\det a|^{-1}\left[\int_{\mathbb{R}^d}\left(E[|X_\xi|^2]\right)^{\frac{1}{2}}|\varphi(a^{-1}(\xi-b))|\,d\xi\right]^2$$
$$\leq |\det a|^{-1}\int_{\mathbb{R}^d}E[|X_\xi|^2]\,d\xi\int_{\mathbb{R}^d}|\varphi(a^{-1}(\xi-b))|^2\,d\xi$$
$$= |\det a|^{-1}E\left[\int_{\mathbb{R}^d}|X_\xi|^2\,d\xi\right]\int_{\mathbb{R}^d}|\varphi(a^{-1}(\xi-b))|^2\,d\xi < \infty.$$

(C2) If $E[|X_t|^2] \leq M$ for some $M < \infty$ and all t, i.e. X_t is bounded in square mean, and $\varphi \in L^1(\mathbb{R}^d)$ we have

$$E\left[|CW_X^a(b,\omega)|^2\right] \leq |\det a|^{-1}\left[\int_{\mathbb{R}^d}\left(E[|X_\xi|^2]\right)^{\frac{1}{2}}|\varphi(a^{-1}(\xi-b))|\,d\xi\right]^2$$
$$\leq |\det a|^{-1}\left[\int_{\mathbb{R}^d}(M)^{\frac{1}{2}}|\varphi(a^{-1}(\xi-b))|\,d\xi\right]^2$$
$$= |\det a|M\|\varphi\|_1^2 < \infty.$$

(C3) If $E[|X_t|^2]$ is bounded on compact subsets of $\mathbb{R}^d$ and φ has compact support and is integrable, we have

$$E[|CW_X^a(b,\omega)|^2] \leq |\det a|^{-1}\left[\int_{\mathbb{R}^d}\left(E[|X_\xi|^2]\right)^{\frac{1}{2}}|\varphi(a^{-1}(\xi-b))|\,d\xi\right]^2 < \infty.$$

(C4) If $E[|X_t|^2] \leq M\left[+\|t\|^2\right]^\theta$ for some $\theta > 0$ and $\left[1+\|t\|^2\right]^{\frac{\theta}{2}}\varphi(x) \in L^1(\mathbb{R})$, then modify $(C2)$ we also have $E\left[|CW_X^a(b,\omega)|^2\right] < \infty$.
In all three cases, as 3.17 is finite, for all a and b, the wavelet transform 3.16 is defined.

Remark 3.9. The continuous wavelet transform of a random field $\{X_t(\omega)\}_{t\in\mathbb{R}^d}$ is a new random (position-scale) field $\{CW_X^a(b,\omega)\}_{a\in H,b\in\mathbb{R}^d}$ provided the path integral is defined with probability one. The continuous wavelet transform at scale $a \in H$ is the random (position) field $\{CW_X^a(b,\omega)\}_{b\in\mathbb{R}^d}$ which is the a-section of the wavelet transform $\{CW_X^a(b,\omega)\}_{a\in H,b\in\mathbb{R}^d}$. As such the output $\{CW_X^a(b)\}_{b\in\mathbb{R}^d}$ inherits certain features of the input X_t and here we focus primarily on how features of the input X_t may yield appropriate properties of the output, at fixed scale or at different scales.

In this thesis, we assume throughout that condition (C2), (C3) or (C4) is satisfied, so that $E[|CW_X^a(b,\omega)|^2] < M < \infty$ for all b. Then for fixed $a \in H$, the auto-correlation of $\{CW_X^a(b,\omega)\}_{b\in\mathbb{R}^d}$ exists and is given by

$$R_{CW_X^a}(\tau) = E\left[CW_X^a(b)\overline{CW_X^a(b+\tau)}\right] \tag{3.18}$$

and for fixed $a_1, a_2 \in H$, the cross-correlation of $\{CW_X^{a_1}(b)\}_{b\in\mathbb{R}^d}$ and $\{CW_X^{a_2}(b)\}_{b\in\mathbb{R}^d}$ exists, and is given by

$$R_{CW_X^{a_1}CW_X^{a_2}}(\tau) = E\left[CW_X^{a_1}(b)\overline{CW_X^{a_2}(b+\tau)}\right]. \tag{3.19}$$

This assumption will allow us to apply Fubini's Theorem when computing $E[CW_X^a(b)]$, as

$$\begin{aligned} E[|CW_X^a(b)|] &\le E\left[\int_{\mathbb{R}^d} |\det a|^{-\frac{1}{2}}|X_u||\varphi(a^{-1}(u-b))|\,du\right] \\ &\le M|\det a|^{\frac{1}{2}}\|\varphi\|_1 \quad \text{for some constant } M. \end{aligned} \tag{3.20}$$

CHAPTER IV

POWER SPECTRAL REPRESENTATION OF THE WAVELET TRANSFORM OF A RANDOM FIELD

Let a second order random field with desired properties, such as square integrable sample paths, and bounded and continuous auto-correlation function be given. The power spectral density function is usually defined for weakly (also strongly) stationary random fields. It turns out that the continuous wavelet transforms of weakly stationary, strongly stationary increments and weakly stationary increments random field are weakly stationary random fields. It is thus natural to ask how the power spectral density functions of these wavelet transform random fields look like. These wavelet transform deal with an integrable mother wavelet function φ via scaling parameter a in a matrix group H and translation parameter $b \in \mathbb{R}^d$.

In this chapter, the spectral density function of the continuous wavelet transform of a random field is discussed. The first section deals with weakly stationary random fields, random fields with strongly stationary increments are discussed in the second section, and random fields with weakly stationary increments are discussed in the fourth section. Moreover, the third section gives an example of the power spectral density function of the wavelet transform of a particular strongly stationary increments random field, namely of a fractional Brownian field.

4.1 The Spectral Representation of the Wavelet Transform of a Weakly Stationary Random Field

Let $\{X_t\}_{t\in\mathbb{R}^d}$ be a weakly stationary random field. Then the auto-correlation function $R_X(\tau)$ is a positive definite function, hence by Bochner's Theorem it has the spectral representation

$$R_X(\tau) = \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}\, dF_X(\lambda) \quad \text{a.e.} \tag{4.1}$$

where $F_X(\lambda)$ is a finite Borel measure on $\mathbb{R}^d$.

The spectral representation for the wavelet transform of a weakly stationary random field with scalar scaling parameter was determined by Elias Masry in 1998. That is, the matrix group is of the simple form $H = \{aI_d : a > 0\}$. It was found that for fixed scaling parameter $a > 0$, the spectral representation of the weakly stationary random field $\{CW_X^a(b)\}_{b\in\mathbb{R}^d}$ is

$$R_{CW_X^a}(\tau) = a^d \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}|\hat{\varphi}(a\lambda)|^2\, dF_X(\lambda) \tag{4.2}$$

and for $a_1, a_2 > 0$, the cross-spectral representation is

$$R_{CW_X^{a_1}CW_X^{a_2}}(\tau) = (a_1a_2)^{\frac{d}{2}} \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}\overline{\hat{\varphi}(a_1\lambda)}\hat{\varphi}(a_2\lambda)\, dF_X(\lambda). \tag{4.3}$$

We now determine the spectral density function of the wavelet transform of a weakly stationary random field with arbitrary dilation matrix.

Theorem 4.1. *Let H be a matrix group and $a, a_1, a_2 \in H$. If $\{X_t\}_{t\in\mathbb{R}^d}$ is a weakly stationary random field, then the random fields $\{CW_X^{a_1}(t)\}_{t\in\mathbb{R}^d}$ and $\{CW_X^{a_2}(t)\}_{t\in\mathbb{R}^d}$ are jointly weakly stationary with constant mean $|\det a|^{\frac{1}{2}} E[X_0]\,\overline{\hat{\varphi}}(0)$. Moreover, the cross-correlation function has the spectral representation*

$$R_{CW_X^{a_1}CW_X^{a_2}}(\tau) = |\det a_1a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}\overline{\hat{\varphi}(a_1^T\lambda)}\hat{\varphi}(a_2^T\lambda)\, dF_X(\lambda). \tag{4.4}$$

In particular, the auto-correlation function has the spectral representation

$$R_{CW_X^a}(\tau) = |\det a| \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda} |\hat{\varphi}(a^T\lambda)|^2 \, dF_X(\lambda). \tag{4.5}$$

Proof. Consider for each $a \in H$, we have by Fubini' theorem (see inequality (3.20))

$$\begin{aligned} E[CW_X^a(t)] &= E\left[\int_{\mathbb{R}^d} |\det a|^{-\frac{1}{2}} X_u \overline{\varphi\left(a^{-1}(u-t)\right)}\, du\right] \\ &= |\det a|^{-\frac{1}{2}} \int_{\mathbb{R}^d} E\left[X_u\right] \overline{\varphi\left(a^{-1}(u-t)\right)}\, du \\ &= |\det a|^{\frac{1}{2}} E\left[X_0\right] \int_{\mathbb{R}^d} \overline{\varphi\left(u\right)}\, du \\ &= |\det a|^{\frac{1}{2}} E\left[X_0\right] \overline{\hat{\varphi}}(0) \end{aligned}$$

where we have used the fact that $\{X_t\}_{t\in\mathbb{R}^d}$ has constant mean, $E[X_u] = E[X_0]$ for all $u \in \mathbb{R}^d$, and $\varphi \in L^1(\mathbb{R}^d)$. This shows that $\{CW_X^a(t)\}_{t\in\mathbb{R}^d}$ has constant mean. Note that if $\{X_t\}_{t\in\mathbb{R}^d}$ is a zero mean random field or the mother wavelet function φ satisfies the condition $\hat{\varphi}(0) = 0$, then we obtain a zero mean random field $\{CW_X^a(t)\}_{t\in\mathbb{R}^d}$.

Furthermore, for each $a_1, a_2 \in H$, we have by Fubini' theorem (see inequality (3.20))

$$\begin{aligned} &R_{CW_X^{a_1} CW_X^{a_2}}(t, t+\tau) \\ &= E\left[|\det a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d} X_\xi \overline{\varphi\left(a_1^{-1}(\xi-t)\right)}\, d\xi \int_{\mathbb{R}^d} \overline{X_\eta} \varphi\left(a_2^{-1}(\eta-t-\tau)\right)\, d\eta\right] \\ &= |\det a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} E\left[X_\xi \overline{X_\eta}\right] \overline{\varphi\left(a_1^{-1}(\xi-t)\right)} \varphi\left(a_2^{-1}(\eta-t-\tau)\right)\, d\xi\, d\eta. \end{aligned}$$

Since $E[X_\xi \overline{X_\eta}] = R_X(\eta - \xi)$, using equation (4.1) we have

$$\begin{aligned} &R_{CW_X^{a_1} CW_X^{a_2}}(t, t+\tau) \\ &= |\det a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{-i(\eta-\xi)\cdot\lambda}\, dF_X(\lambda)\; \overline{\varphi\left(a_1^{-1}(\xi-t)\right)} \varphi\left(a_2^{-1}(\eta-t-\tau)\right)\, d\xi\, d\eta \\ &= |\det a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{i\xi\cdot\lambda}\; \overline{\varphi\left(a_1^{-1}(\xi-t)\right)}\, d\xi \int_{\mathbb{R}^d} e^{-i\eta\cdot\lambda} \varphi\left(a_2^{-1}(\eta-t-\tau)\right)\, d\eta \\ &\quad dF_X(\lambda). \end{aligned}$$

Changing variables, $\xi \mapsto a_1\xi + t, \quad \eta \mapsto a_2\eta + (t+\tau)$ it follows that

$$\begin{aligned}
&R_{CW_X^{a_1} CW_X^{a_2}}(t, t+\tau) \\
&= |\,\mathrm{det} a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{i(a_1\xi)\cdot\lambda} e^{it\cdot\lambda}\overline{\varphi\left(\xi\right)}\, d\xi \int_{\mathbb{R}^d} e^{-i(a_2\eta)\cdot\lambda} e^{-it\cdot\lambda} e^{-i\tau\cdot\lambda}\varphi\left(\eta\right)\, d\eta\, dF_X(\lambda) \\
&= |\,\mathrm{det} a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda} \int_{\mathbb{R}^d} e^{i\xi\cdot(a_1^T\lambda)}\overline{\varphi\left(\xi\right)}\, d\xi \int_{\mathbb{R}^d} e^{-i\eta\cdot(a_2^T\lambda)}\varphi\left(\eta\right)\, d\eta\, dF_X(\lambda) \\
&= |\,\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}\overline{\hat{\varphi}\left(a_1^T\lambda\right)}\hat{\varphi}\left(a_2^T\lambda\right)\, dF_X(\lambda).
\end{aligned}$$

If $a_1 = a_2 = a$, we have

$$R_{CW_X^a}(t, t+\tau) = |\,\det a| \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}|\hat{\varphi}(a^T\lambda)|^2\, dF_X(\lambda).$$

We can see that the cross-correlation function $R_{CW_X^{a_1} CW_X^{a_2}}(t, t+\tau)$, and the auto-correlation function $R_{CW_X^a}(t, t+\tau)$ depend on position translation τ only, hence we denote them by $R_{CW_X^{a_1} CW_X^{a_2}}(\tau)$ and $R_{CW_X^a}(\tau)$, respectively. It follows that the random fields $\{CW_X^{a_1}(t)\}_{t\in\mathbb{R}^d}$ and $\{CW_X^{a_2}(t)\}_{t\in\mathbb{R}^d}$ are jointly weakly stationary, and the cross power spectral representation of $R_{CW_X^{a_1} CW_X^{a_2}}(\tau)$ and the power spectral representation of $R_{CW_X^a}(\tau)$ are given as above. □

Remark 4.1. By Definition 3.18, the cross power spectral density function is arrived at as follows

$$\begin{aligned}
\int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda} S_{CW_X^{a_1} CW_X^{a_2}}(\lambda)\, d\lambda &= |\,\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}\overline{\hat{\varphi}\left(a_1^T\lambda\right)}\hat{\varphi}\left(a_2^T\lambda\right)\, dF_X(\lambda) \\
&= |\,\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}\overline{\hat{\varphi}\left(a_1^T\lambda\right)}\hat{\varphi}\left(a_2^T\lambda\right) S_X(\lambda)\, d\lambda,
\end{aligned}$$

and the power spectral density function is arrived as

$$\begin{aligned}
\int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda} S_{CW_X^a}(\lambda)\, d\lambda &= |\,\det a| \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}|\hat{\varphi}(a^T\lambda)|^2\, dF_X(\lambda) \\
&= |\,\det a| \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}|\hat{\varphi}(a^T\lambda)|^2\, S_X(\lambda)\, d\lambda.
\end{aligned}$$

4.2 The Spectral Representation of the Wavelet Transform of a Strongly Stationary Increments Random Field

Let $\{X_t\}_{t\in\mathbb{R}^d}$ be a stationary increments random field with zero mean. For $d = 1$, its auto-correlation function has the spectral representation, as outlined in Dzhaparide (2005),

$$R_X(t,s) = \int_{\mathbb{R}} (e^{i\lambda t} - 1)(e^{-i\lambda s} - 1)\, dF_X(\lambda) \quad s,t \in \mathbb{R} \tag{4.6}$$

where $dF_X(\lambda)$ is a Borel measure on $\mathbb{R}$ which satisfies $\displaystyle\int_{\mathbb{R}} \frac{|\lambda|^2}{1+|\lambda|^2}\, dF_X(\lambda)$.
In the special case of fractional Brownian motion $\{B_t^H\}_{t\in\mathbb{R}}$, the representation of its auto-correlation function is

$$R_{B^H}(t,s) = C_H^2 \int_{\mathbb{R}} (e^{i\lambda t} - 1)(e^{-i\lambda s} - 1)\, \frac{d\lambda}{|\lambda|^{2H+1}}, \quad s,t \in \mathbb{R}, \tag{4.7}$$

for some positive constant C_H^2.
Malyarenko (2005), see Dzhaparide (2005), treats the multidimensional case; the auto-correlation of a stationary increments random field has the spectral representation

$$R_X(t,s) = \int_{\mathbb{R}^d} (e^{i\lambda\cdot t} - 1)(e^{-i\lambda\cdot s} - 1)\, dF_X(\lambda), \quad s,t \in \mathbb{R}^d \tag{4.8}$$

where $dF_X(\lambda)$ is a Borel measure on $\mathbb{R}^d$ which satisfies $\displaystyle\int_{\mathbb{R}^d} \frac{\|\lambda\|^2}{1+\|\lambda\|^2}\, dF_X(\lambda)$.
For a fractional Brownian field $\{B_t^H\}_{t\in\mathbb{R}^d}$ the representation of its auto-correlation function is

$$R_{B^H}(t,s) = C_H^2 \int_{\mathbb{R}^d} (e^{i\lambda\cdot t} - 1)(e^{-i\lambda\cdot s} - 1)\, \frac{d\lambda}{\|\lambda\|^{2H+d}}, \quad s,t \in \mathbb{R}^d, \tag{4.9}$$

for some positive constant C_H^2.
We have the following theorem for the continuous wavelet transform of a strongly stationary increments random field, assuming that the integrable mother wavelet function φ satisfies $\hat{\varphi}(0) = 0$ and condition (C3) in Remark 3.8 holds.

Theorem 4.2. *Let $\{X_t\}_{t\in\mathbb{R}^d}$ be a stationary increments random field with zero mean. Let H be a matrix group and $a, a_1, a_2 \in H$. Then $\{CW_X^{a_1}(t)\}_{t\in\mathbb{R}^d}$ and $\{CW_X^{a_2}(t)\}_{t\in\mathbb{R}^d}$ are jointly weakly stationary random fields with zero mean. Moreover, the cross-correlation has the spectral representations as*

$$R_{CW_X^{a_1} CW_X^{a_2}}(\tau) = |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}\overline{\hat{\varphi}(a_1^T\lambda)}\hat{\varphi}(a_2^T\lambda)\, dF_X(\lambda)$$

where $dF_X(\lambda)$ is a Borel measure on $\mathbb{R}^d$ which satisfies $\displaystyle\int_{\mathbb{R}^d} \frac{\|\lambda\|^2}{1+\|\lambda\|^2}\, dF_X(\lambda)$. In particular, the auto-correlation function has the spectral representations as

$$R_{CW_X^{a}}(\tau) = |\det a| \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}|\hat{\varphi}(a^T\lambda)|^2\, dF_X(\lambda).$$

Proof. For each scaling parameter $a \in H$, we have by Fubini' theorem (see 3.20)

$$\begin{aligned} E[CW_X^a(t)] &= E\left[\int_{\mathbb{R}^d} |\det a|^{-\frac{1}{2}} X_u \overline{\varphi\left(a^{-1}(u-t)\right)}\, du\right] \\ &= |\det a|^{-\frac{1}{2}} \int_{\mathbb{R}^d} E\left[X_u\right] \overline{\varphi\left(a^{-1}(u-t)\right)}\, du. \end{aligned}$$

Since $E[X_t] = 0$ for all $t \in \mathbb{R}^d$ it follows that $E[CW_X^a(t)] = 0$.

Furthermore, for each $a_1, a_2 \in H$,

$$\begin{aligned} &R_{CW_X^{a_1} CW_X^{a_2}}(t, t+\tau) \\ &\quad= E\left[|\det a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d} X_\xi \overline{\varphi\left(a_1^{-1}(\xi-t)\right)}\, d\xi \overline{\int_{\mathbb{R}^d} X_\eta \overline{\varphi\left(a_2^{-1}(\eta-t-\tau)\right)}\, d\eta}\right] \\ &\quad= |\det a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} E\left[X_\xi \overline{X_\eta}\right] \overline{\varphi\left(a_1^{-1}(\xi-t)\right)}\varphi\left(a_2^{-1}(\eta-t-\tau)\right)\, d\xi\, d\eta. \end{aligned}$$

Applying the representation of $R_X(\xi,\eta)$ by equation (4.8) we have

$$\begin{aligned} R_{CW_X^{a_1} CW_X^{a_2}}(t, t+\tau) &= |\det a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} (e^{i\lambda\cdot\xi}-1)(e^{-i\lambda\cdot\eta}-1)\, dF_X(\lambda)\right) \\ &\qquad \overline{\varphi\left(a_1^{-1}(\xi-t)\right)}\varphi\left(a_2^{-1}(\eta-t-\tau)\right)\, d\xi\, d\eta \\ &= |\det a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} (e^{i\lambda\cdot\xi}-1)\overline{\varphi\left(a_1^{-1}(\xi-t)\right)}\, d\xi \\ &\qquad \int_{\mathbb{R}^d} (e^{-i\lambda\cdot\eta}-1)\varphi\left(a_2^{-1}(\eta-t-\tau)\right)\, d\eta\, dF_X(\lambda). \end{aligned}$$

Since $\hat{\varphi}(0) = 0$ we have $\int_{\mathbb{R}^d} \varphi(a_1^{-1}(\xi - t))\, d\xi = 0$ and $\int_{\mathbb{R}^d} \varphi(a_2^{-1}(\eta - t - \tau))\, d\eta = 0$ so that

$$R_{CW_X^{a_1} CW_X^{a_2}}(t, t+\tau) = |\det a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{i\xi\cdot\lambda}\, \overline{\varphi\left(a_1^{-1}(\xi - t)\right)}\, d\xi \int_{\mathbb{R}^d} e^{-i\eta\cdot\lambda}$$
$$\varphi\left(a_2^{-1}(\eta - t - \tau)\right)\, d\eta\, dF_X(\lambda).$$

Continuing as in the proof of Theorem 4.1, we arrive at the assertion. □

In the special case of fractional Brownian field, the spectral density can be computed, details are given in the next section.

4.3 The Spectral Representation of the Wavelet Transform of a Fractional Brownian Field

Let $\{B_t^H\}_{t\in\mathbb{R}}$ be fractional Brownian motion with Hurst index $0 < H < 1$. In 1998, Takeshi Kato and Elias Masry computed the formula of the power spectral density function of the wavelet transform of fractional Brownian motion as

$$S_{CW_{BH}^a}(\lambda) = \frac{aV_H \Gamma(2H+1)\sin(\pi H)|\hat{\varphi}(-a\lambda)|^2}{2H|\lambda|^{2H+1}} \tag{4.10}$$

where $V_H = E[(B_1^H)^2]$. In this thesis we present the simple computation of the spectral density function of a fractional Brownian field by the representation (4.9) for the d-dimensional case, in cases $d = 1$ this has been done in Masry by a different method.

Let $\{B_t^H\}_{t\in\mathbb{R}^d}$ be fractional Brownian motion with Hurst index $0 < H < 1$. Computing the spectral density function by representation (4.9), we obtain the following theorem.

Theorem 4.3. *Let φ be an integrable mother wavelet function with $\sup_{\|\lambda\|<\epsilon} \frac{|\hat{\varphi}(\lambda)|}{\|\lambda\|^{\frac{d}{2}+H}} \leq M$ for some $\epsilon > 0$ and $0 < M < \infty$ and condition (C3) in Remark 3.8 holds. Let*

$\mathcal{H}$ be a matrix group and $a, a_1, a_2 \in \mathcal{H}$. Then the random fields $\{CW^{a_1}_{B^H}(t)\}_{t\in\mathbb{R}^d}$ and $\{CW^{a_2}_{B^H}(t)\}_{t\in\mathbb{R}^d}$ are jointly weakly stationary with zero mean. Moreover, the cross-correlation function has the power spectral representation

$$R_{CW^{a_1}_X CW^{a_2}_X}(\tau) = |\,\mathrm{det} a_1 a_2|^{\frac{1}{2}} C_H^2 \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}\overline{\hat{\varphi}(a_1^T\lambda)}\hat{\varphi}(a_2^T\lambda)\,\frac{d\lambda}{\|\lambda\|^{2H+d}}. \tag{4.11}$$

where $dF_X(\lambda)$ is a Borel measure on $\mathbb{R}^d$ which satisfies $\int_{\mathbb{R}^d} \frac{\|\lambda\|^2}{1+\|\lambda\|^2}\, dF_X(\lambda)$. In particular, the auto-correlation function has the power spectral representation

$$R_{CW^a_{B^H}}(\tau) = |\,\mathrm{det} a| C_H^2 \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}|\hat{\varphi}(a^T\lambda)|^2\,\frac{d\lambda}{\|\lambda\|^{2H+d}}. \tag{4.12}$$

The cross power spectral density function is

$$S_{CW^{a_1}_X CW^{a_2}_X}(\lambda) = \frac{|\,\mathrm{det} a_1 a_2|^{\frac{1}{2}} C_H^2\overline{\hat{\varphi}(a_1^T\lambda)}\hat{\varphi}(a_2^T\lambda)}{\|\lambda\|^{2H+d}} \tag{4.13}$$

and the power spectral density function is

$$S_{CW^a_X}(\lambda) = \frac{|\,\mathrm{det}\, a| C_H^2 |\hat{\varphi}(a^T\lambda)|^2}{\|\lambda\|^{2H+d}}. \tag{4.14}$$

Proof. Since $E[B^H_t] = 0$ for all $t \in \mathbb{R}^d$ we have

$$E[CW^a_{B^H}(t)] = |\,\mathrm{det} a|^{-\frac{1}{2}} \int_{\mathbb{R}^d} E\left[B^H_s\right]\overline{\varphi(a^{-1}(s-t))}\, ds = 0.$$

The quantity of interest are the auto-correlation and cross-correlation of the random field $\{CW^a_{B^H}(t)\}_{t\in\mathbb{R}^d}$ for fixed scaling parameter $a \in H$. Now we will consider the cross-correlation

$$\begin{aligned} R_{CW^{a_1}_{B^H} CW^{a_2}_{B^H}}(t, t+\tau) &= E\left[CW^{a_1}_{B^H}(t)\overline{CW^{a_2}_{B^H}(t+\tau)}\right] \\ &= E\left[|\,\mathrm{det} a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d} B^H_\xi \overline{\varphi\left(a_1^{-1}(\xi - t)\right)}\, d\xi \int_{\mathbb{R}^d} \overline{B^H_\eta}\varphi\left(a_2^{-1}(\eta - t - \tau)\right)\, d\eta\right]. \end{aligned}$$

Applying equation (4.9),

$$\begin{aligned} R_{CW^{a_1}_{B^H} CW^{a_2}_{B^H}}(t, t+\tau) = C_H^2 |\,\mathrm{det} a_1 a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} (e^{i\lambda\cdot\xi} - 1)(e^{-i\lambda\cdot\eta} - 1)\frac{d\lambda}{\|\lambda\|^{2H+d}} \\ \overline{\varphi(a_1^{-1}(\xi - t))}\varphi(a_2^{-1}(\eta - t - \tau))\, d\xi\, d\eta \end{aligned}$$

and using Fubini's Theorem,

$$R_{CW^{a_1}_{BH}CW^{a_2}_{BH}}(t,t+\tau) = C_H^2|\,\mathrm{det}a_1a_2|^{-\frac{1}{2}}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(e^{i\lambda\cdot\xi}-1)\overline{\varphi(a_1^{-1}(\xi-t))}\,d\xi$$
$$\int_{\mathbb{R}^d}(e^{-i\lambda\cdot\eta}-1)\varphi(a_2^{-1}(\eta-t-\tau))\,d\eta\,\frac{d\lambda}{\|\lambda\|^{2H+d}}.$$

Since $\hat{\varphi}(0)=0$ we have $\int_{\mathbb{R}^d}\overline{\varphi(a_1^{-1}(\xi-t))}\,d\xi=0$ and $\int_{\mathbb{R}^d}\varphi(a_2^{-1}(\eta-t-\tau))\,d\eta=0$, and the above simplifies to

$$R_{CW^{a_1}_{BH}CW^{a_2}_{BH}}(\tau) = C_H^2|\,\mathrm{det}a_1a_2|^{-\frac{1}{2}}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}e^{i\lambda\cdot\xi}\overline{\varphi(a_1^{-1}(\xi-t))}\,d\xi$$
$$\int_{\mathbb{R}^d}e^{-i\lambda\cdot\eta}\varphi(a_2^{-1}(\eta-t-\tau))\,d\eta\,\frac{d\lambda}{\|\lambda\|^{2H+d}}.$$

Changing variables $\xi\longmapsto a_1\xi+t$ and $\eta\longmapsto a_2\eta+t+\tau$ we have by definition of the Fourier transform

$$\begin{aligned}
&R_{CW^{a_1}_{BH}CW^{a_2}_{BH}}(\tau)\\
&= C_H^2|\,\mathrm{det}a_1a_2|^{\frac{1}{2}}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}e^{i\lambda\cdot(a_1\xi+t)}\overline{\varphi(\xi)}\,d\xi\int_{\mathbb{R}^d}e^{-i\lambda\cdot(a_2\eta+t+\tau)}\varphi(\eta)\,d\eta\,\frac{d\lambda}{\|\lambda\|^{2H+d}}\\
&= C_H^2|\,\mathrm{det}a_1a_2|^{\frac{1}{2}}\int_{\mathbb{R}^d}e^{-i\lambda\cdot\tau}\int_{\mathbb{R}^d}e^{i\xi\cdot a_1^T\lambda}\overline{\varphi(\xi)}\,d\xi\int_{\mathbb{R}^d}e^{-i\eta\cdot a_2^T\lambda}\varphi(\eta)\,d\eta\,\frac{d\lambda}{\|\lambda\|^{2H+d}}\\
&= C_H^2|\,\mathrm{det}a_1a_2|^{\frac{1}{2}}\int_{\mathbb{R}^d}e^{-i\lambda\cdot\tau}\overline{\hat{\varphi}(a_1^T\lambda)}\hat{\varphi}(a_2^T\lambda)\,\frac{d\lambda}{\|\lambda\|^{2H+d}}.
\end{aligned}$$

We now claim that the function $g(\lambda)=\frac{|\hat{\varphi}(\lambda)|^2}{\|\lambda\|^{2H+d}}$ is integrable for $0<H<1$. By assumption, there exist $\epsilon>0$ and $M>0$ so that and $\sup_{\|\lambda\|<\epsilon}\frac{|\hat{\varphi}(\lambda)|}{\|\lambda\|^{\frac{d}{2}+H}}\le M$. We split

$$\int_{\mathbb{R}^d}|g(\lambda)|\,d\lambda=\int_{\|\lambda\|\le\epsilon}|g(\lambda)|\,d\lambda\quad+\quad\int_{\|\lambda\|>\epsilon}|g(\lambda)|\,d\lambda.$$

Then we have

$$\int_{\|\lambda\|\le\epsilon}|g(\lambda)|\,d\lambda=\int_{\|\lambda\|\le\epsilon}\frac{|\hat{\varphi}(\lambda)|^2}{\|\lambda\|^{2H+d}}\,d\lambda\le\int_{\|\lambda\|\le\epsilon}M^2\,d\lambda<\infty.$$

For the second integral, we have as $\varphi\in L^1(\mathbb{R})$ and $|\hat{\varphi}(\lambda)|\le\|\varphi\|_{L^1}$ by Remark 2.1 (ii),

$$\int_{\|\lambda\|>\epsilon} |g(\lambda)|\, d\lambda = \int_{\|\lambda\|>\epsilon} \frac{|\hat{\varphi}(\lambda)|^2}{\|\lambda\|^{2H+d}}\, d\lambda \leq \int_{\|\lambda\|>\epsilon} \frac{\|\varphi\|_{L^1}^2}{\|\lambda\|^{2H+d}}\, d\lambda$$
$$= \|\varphi\|_{L^1}^2 \int_{\|\lambda\|>\epsilon} \frac{1}{\|\lambda\|^{2H+d}}\, d\lambda < \infty.$$

Hence $g(\lambda) = \dfrac{|\hat{\varphi}(\lambda)|^2}{\|\lambda\|^{2H+d}}$ is integrable for $0 < H < 1$. We can see that, if a is any invertible matrix, then $b\|\lambda\| \leq \|a\lambda\| \leq c\|\lambda\|$ for all $\lambda \neq 0$ and some constants b, c, hence $\dfrac{|\hat{\varphi}(a\lambda)|^2}{\|\lambda\|^{2H+d}}$ is also integrable. By Cauchy-Schwartz, $h(\lambda) = |\det a_1 a_2|^{\frac{1}{2}} C_H^2 \dfrac{\overline{\hat{\varphi}(a_1^T\lambda)}\hat{\varphi}(a_2^T\lambda)}{\|\lambda\|^{2H+d}}$ is integrable for $0 < H < 1$. It follows that $R_{CW^{a_1}_{B^H} CW^{a_2}_{B^H}}$ is the Fourier transform of $h(\lambda)$, hence $h(\lambda)$ is the cross-power spectral density function of the random field $\{CW^{a_1}_{B^H}(t)\}_{t\in\mathbb{R}^d}$ and $\{CW^{a_2}_{B^H}(t)\}_{t\in\mathbb{R}^d}$ and is given by

$$S_{CW^{a_1}_{B^H} CW^{a_2}_{B^H}}(\lambda) = C_H^2 |\operatorname{det} a_1 \operatorname{det} a_2|^{\frac{1}{2}} \frac{\overline{\hat{\varphi}(a_1^T\lambda)}\hat{\varphi}(a_2^T\lambda)}{\|\lambda\|^{2H+d}}.$$

If $a_1 = a_2 = a$, we have

$$R_{CW^{a}_{B^H}}(\tau) = |\operatorname{det} a| C_H^2 \int_{\mathbb{R}^d} e^{-i\lambda\cdot\tau} |\hat{\varphi}(a^T\lambda)|^2 \frac{d\lambda}{\|\lambda\|^{2H+d}}.$$

thus the spectral density function, $S_{CW^{a}_{B^H}}(\lambda)$ equals

$$S_{CW^{a}_{B^H}}(\lambda) = |\operatorname{det} a| C_H^2 \frac{|\hat{\varphi}(a^T\lambda)|^2}{\|\lambda\|^{2H+d}}.$$

□

4.4 The Spectral Representation of the Wavelet Transform of Weakly Stationary Increments Random Field

Let $\{X_t\}_{t\in\mathbb{R}^d}$ be a weakly stationary increments random field with zero mean. The auto-correlation function of its increments $R_{\Delta X(t,\tau_1)\Delta X(s,\tau_2)} = E[(X_{t+\tau_1} - X_t)\overline{(X_{s+\tau_2} - X_s)}]$ admits the spectral representation (Yaglom, 1962)

$$R_{\Delta X(t,\tau_1)\Delta X(s,\tau_2)} = \int_{\mathbb{R}^d/\{0\}} e^{i(t-s)\lambda}(1 - e^{i\tau_1\cdot\lambda})(1 - e^{-i\tau_2\cdot\lambda})\, dF_X(\lambda) + (A\tau_1)\cdot\tau_2 \quad (4.15)$$

where $dF_X(\lambda)$ is a measure on $\mathbb{R}^d/\{0\}$ satisfying $\int_{\mathbb{R}^d/\{0\}} \frac{\|\lambda\|^2}{1+\|\lambda\|^2}\, dF_X(\lambda) < \infty$ and A is a positive definite Hermitian matrix. The term $(A\tau_1)\cdot\tau_2$ represents the contribution to the integral at $\lambda = 0$. Assume that condition (C4) in Remark 3.8 holds, we have the following theorem.

Theorem 4.4. *Let H be a matrix group and $a, a_1, a_2 \in H$. Let φ be an integrable mother wavelet function which has zero first moments, that is $\int_{\mathbb{R}^d} u_i\varphi(u)\, du = 0$ for all $i = 1, 2, ..., d$, and $\hat{\varphi}(0) = 0$. If $\{X_t\}_{t\in\mathbb{R}^d}$ has weakly stationary increments and zero mean, then the random fields $\{CW_X^{a_1}(t)\}_{t\in\mathbb{R}^d}$ and $\{CW_X^{a_2}(t)\}_{t\in\mathbb{R}^d}$ are jointly weakly stationary with zero mean, and the auto-correlation and cross-correlation have the power spectral representations and cross power spectral representations*

$$R_{CW_X^a}(t, t+\tau) = |\det a| \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda} |\hat{\varphi}(a^T\lambda)|^2\, dF_X(\lambda) \tag{4.16}$$

$$R_{CW_X^{a_1}CW_X^{a_2}}(t, t+\tau) = |\det a_1a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda} \overline{\hat{\varphi}(a_1^T\lambda)}\hat{\varphi}(a_2^T\lambda)\, dF_X(\lambda), \tag{4.17}$$

respectively, where $dF_X(\lambda)$ is a measure on $\mathbb{R}^d/\{0\}$ satisfying $\int_{\mathbb{R}^d/\{0\}} \frac{\|\lambda\|^2}{1+\|\lambda\|^2}\, dF_X(\lambda) < \infty$.

Proof. As in the previous proofs, $E[X_u] = 0$ for all $u \in \mathbb{R}^d$ yields $E[CW_X^a(t)] = 0$ and for $a_1, a_2 \in H$ we have

$$R_{CW_X^{a_1}CW_X^{a_2}}(t, t+\tau) = |\det a_1a_2|^{-\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} E\left[X_\xi\overline{X_\eta}\right] \overline{\varphi\left(a_1^{-1}(\xi-t)\right)}\varphi\left(a_2^{-1}(\eta-t-\tau)\right) d\xi\, d\eta.$$

Changing variables $\xi \mapsto a_1\xi + t$ and $\eta \mapsto a_2\eta + t + \tau$,

$$R_{CW_X^{a_1}CW_X^{a_2}}(t, t+\tau) = |\det a_1a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} E\left[X_{a_1\xi+t}\overline{X_{a_2\eta+t+\tau}}\right] \overline{\varphi(\xi)}\varphi(\eta)\, d\xi\, d\eta.$$

Since $\hat{\varphi}(0) = 0$, we add zero value terms as follows,

$$
\begin{aligned}
R_{CW_X^{a_1}CW_X^{a_2}}(t, t+\tau) = |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} E\left[X_{a_1\xi+t}\overline{X_{a_2\eta+t+\tau}}\right] \overline{\varphi(\xi)}\varphi(\eta)\, d\xi\, d\eta \\
- |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d} E\left[X_t\overline{X_{a_2\eta+t+\tau}}\right] \int_{\mathbb{R}^d} \overline{\varphi(\xi)}\, d\xi\, \varphi(\eta)\, d\eta \\
- |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d} E\left[X_{a_1\xi+t}\overline{X_{t+\tau}}\right] \int_{\mathbb{R}^d} \varphi(\eta)\, d\eta\, \overline{\varphi(\xi)}\, d\xi
\end{aligned}
$$

$$
\begin{aligned}
+ |\det a_1 a_2|^{\frac{1}{2}} E\left[X_t\overline{X_{t+\tau}}\right] \int_{\mathbb{R}^d} \overline{\varphi(\xi)}\, d\xi \int_{\mathbb{R}^d} \varphi(\eta)\, d\eta \\
= |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} E\left[(X_{a_1\xi+t} - X_t)\overline{(X_{a_2\eta+t+\tau} - X_{t+\tau})}\right] \overline{\varphi(\xi)}\varphi(\eta)\, d\xi\, d\eta.
\end{aligned}
$$

Applying the power spectral representation of $R_{\Delta X(t,\tau_1)\Delta X(s,\tau_2)}$ as in equation (4.15), $t = t, s = t+\tau, \tau_1 = a_1\xi, \tau_2 = a_2\eta$.

$$
\begin{aligned}
& R_{CW_X^{a_1}CW_X^{a_2}}(t, t+\tau) \\
& = |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d/\{0\}} e^{-i\tau\cdot\lambda}(1 - e^{ia_1\xi\cdot\lambda})(1 - e^{-ia_2\eta\cdot\lambda})dF_X(\lambda)\overline{\varphi(\xi)}\varphi(\eta)\, d\xi\, d\eta \\
& \quad + |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} (Aa_1\xi)\cdot(a_2\eta)\overline{\varphi(\xi)}\varphi(\eta)\, d\xi\, d\eta \\
& = |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d/\{0\}} e^{-i\tau\cdot\lambda} \int_{\mathbb{R}^d} (1 - e^{ia_1\xi\cdot\lambda})\overline{\varphi(\xi)}\, d\xi \int_{\mathbb{R}^d} (1 - e^{-ia_2\eta\cdot\lambda})\varphi(\eta)\, d\eta\, dF_X(\lambda) \\
& \quad + |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \left(\sum_{i=1}^{d}\sum_{j=1}^{d} \tilde{a}_{ij}\xi_j\eta_i\right) \overline{\varphi(\xi)}\varphi(\eta)\, d\xi\, d\eta \quad \text{where } [\tilde{a}_{ij}] = a_2^T A a_1.
\end{aligned}
$$

Note that the use of Fubini's theorem is justified by the estimation of integral and $\int_{\mathbb{R}^d/\{0\}} \frac{\|\lambda\|^2}{1+\|\lambda\|^2}\, dF_X(\lambda) < \infty$. Since $\hat{\varphi}(0) = 0$ and φ has zero first moments, we can reduce the above equation

$$
\begin{aligned}
& R_{CW_X^{a_1}CW_X^{a_2}}(t, t+\tau) \\
& = |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d/\{0\}} e^{-i\tau\cdot\lambda} \int_{\mathbb{R}^d} e^{i\xi\cdot a_1^T\lambda}\overline{\varphi(\xi)}\, d\xi \int_{\mathbb{R}^d} e^{-i\eta\cdot a_2^T\lambda}\varphi(\eta)\, d\eta\, dF_X(\lambda) \\
& = |\det a_1 a_2|^{\frac{1}{2}} \int_{\mathbb{R}^d/\{0\}} e^{-i\tau\cdot\lambda}\overline{\hat{\varphi}(a_1^T\lambda)}\, \hat{\varphi}(a_2^T\lambda)\, dF_X(\lambda).
\end{aligned}
$$

If $a_1 = a_2 = a$, we have

$$
R_{CW_X^a}(\tau) = |\det a| \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda}|\hat{\varphi}(a^T\lambda)|^2\, dF_X(\lambda).
$$

We can see that the cross-correlation function $R_{CW_X^{a_1} CW_X^{a_2}}(t, t+\tau)$, and the auto-correlation function $R_{CW_X^a}(t, t+\tau)$ depend on position translation τ only, hence we denote them by $R_{CW_X^{a_1} CW_X^{a_2}}(\tau)$ and $R_{CW_X^a}(\tau)$, respectively. It follows that random fields $\{CW_X^{a_1}(t)\}_{t\in\mathbb{R}^d}$ and $\{CW_X^{a_2}(t)\}_{t\in\mathbb{R}^d}$ are jointly weakly stationary, and the cross power spectral representation of $R_{CW_X^{a_1} CW_X^{a_2}}(\tau)$ and the power spectral representation of $R_{CW_X^a}(\tau)$ are given as above. □

CHAPTER V

ERGODICITY PROPERTIES

Ergodicity is very useful and widely used. However, ergodic theorems were stated in the past under a variety of conditions regarding the random field to which they applied; earlier versions were preoccupied with stationary or weakly stationary random fields only, as it was thought for a while that stationarity was needed for ergodicity. The first section gives the basic concept of the ergodicity. Details can be found in Childers (1997), Grimmett and Stirzaker (1998), Papoulis and Unnikrishno (2002) and Yannis (1998). The ergodic theorem for weakly stationary random fields is introduced and proven in the second section. Finally, we will show how ergodicity properties are connected to the continuous wavelet transform.

5.1 Mean Ergodic Random Fields

Let $\{X_t\}_{t\in\mathbb{R}^d}$ be a random field. There are several conventions for denoting position (time) averages. One convention uses the one-sided average, which is expressed for continuous random processes as

$$\langle X_t \rangle = \lim_{T\to\infty} \frac{1}{T^d} \int_{[0,T]^d} X_t \, dt.$$

Another convention is called the two-sided average, which is denoted as

$$\langle X_t \rangle = \lim_{T\to\infty} \frac{1}{(2T)^d} \int_{[-T,T]^d} X_t \, dt.$$

The two-sided convention appears more common, especially for theoretical definitions. Ergodic in the mean or in short mean ergodic says that the estimate for the mean converges to the true mean in the mean square sense as per following definition.

Definition 5.1. (Mean Ergodic Field)

A random field $\{X_t\}_{t\in\mathbb{R}^d}$ with constant mean m is mean ergodic if

$$\lim_{T\to\infty}\frac{1}{(2T)^d}\int_{[-T,T]^d} X_t\,dt = m \tag{5.1}$$

in the mean square sense.

The following theorems provide conditions for mean ergodicity and suggests an alternative definition for mean ergodicity of weakly stationary random fields.

Theorem 5.1. *Let $\{X_t\}_{t\in\mathbb{R}^d}$ be a weakly stationary random field with constant mean m and auto-covariance function $C_X(\tau)$. A necessary and sufficient condition for $\{X_t\}_{t\in\mathbb{R}^d}$ to be mean ergodic is*

$$\lim_{T\to\infty}\frac{1}{(2T)^d}\int_{[-2T,2T]^d}\prod_{i=1}^{d}\left(1-\frac{|\tau^i|}{2T}\right)C_X(\tau)\,d\tau = 0. \tag{5.2}$$

Proof. Let $T>0$. For simplicity, let as define

$$\langle X\rangle_T = \frac{1}{(2T)^d}\int_{[-T,T]^d} X_t\,dt.$$

Then, as $E[|X_t|^2]$ is a constant and $E[|X_t|]\le M$ we can use Fubini's theorem,

$$E[\langle X\rangle_T] = \frac{1}{(2T)^d}\int_{[-T,T]^d} E[X_t]\,dt = \frac{m}{(2T)^d}\int_{[-T,T]^d} 1\,dt = m$$

and

$$\begin{aligned}
\operatorname{var}[\langle X\rangle_T] &= E[|\langle X\rangle_T - m|^2]\\
&= \frac{1}{(2T)^{2d}}E\left[\left(\int_{[-T,T]^d}(X_t - E[X_t])\,dt\right)\overline{\left(\int_{[-T,T]^d}(X_s - E[X_s])\,ds\right)}\right]\\
&= \frac{1}{(2T)^{2d}}\int_{[-T,T]^d}\int_{[-T,T]^d} E[(X_t - E[X_t])(\overline{X_s - E[X_s]})]\,dt\,ds\\
&= \frac{1}{(2T)^{2d}}\int_{[-T,T]^d}\int_{[-T,T]^d} C_X(s-t)\,dt\,ds.
\end{aligned}$$

For each $i = 1, 2, ..., d$, if we let $\tau_i = s_i - t_i$ and $u_i = s_i + t_i$,

then $dt_i\, ds_i = |J_i|^{-1} du_i\, d\tau_i$ where $|J_i|$ is the Jacobian which is

$$|J^i| = \begin{vmatrix} \frac{\partial u_i}{\partial t_i} & \frac{\partial u_i}{\partial s_i} \\ \frac{\partial \tau_i}{\partial t_i} & \frac{\partial \tau_i}{\partial s_i} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = 2.$$

Then we have

$$\begin{aligned} \mathrm{var}[\langle X \rangle_T] &= \frac{1}{(2T)^d} \int_{[-2T,2T]^d} \int_{-(2T-|\tau_1|)}^{2T-|\tau_1|} \int_{-(2T-|\tau_2|)}^{2T-|\tau_2|} \cdots \int_{-(2T-|\tau_d|)}^{2T-|\tau_d|} C_X(\tau) (\frac{1}{2})^d \, du \, d\tau \\ &= \frac{1}{2^d (2T)^{2d}} \int_{[-2T,2T]^d} C_X(\tau) \prod_{i=1}^{d} \int_{-(2T-|\tau_i|)}^{2T-|\tau_i|} 1 \, du_i \, d\tau \\ &= \frac{1}{2^d (2T)^{2d}} \int_{[-2T,2T]^d} C_X(\tau) \prod_{i=1}^{d} (4T - 2|\tau_i|) \, d\tau \\ &= \frac{1}{(2T)^d} \int_{[-2T,2T]^d} \prod_{i=1}^{d} (1 - \frac{|\tau_i|}{2T}) C_X(\tau) \, d\tau. \end{aligned}$$

The condition 5.2 of the theorem is thus equivalent to $\mathrm{var}[\langle X \rangle_T] \to 0$, that is, to mean ergodicity. □

Because of this theorem, an alternative definition of mean ergodicity for a weakly stationary random field is then given.

Definition 5.2. (Alternative Definition of Mean Ergodicity)
A weakly stationary random field $\{X_t\}_{t \in \mathbb{R}^d}$ with constant mean is ergodic in the mean if and only if

$$\lim_{T \to \infty} \frac{1}{(2T)^d} \int_{[-2T,2T]^d} \prod_{i=1}^{d} (1 - \frac{|\tau^i|}{2T}) C_X(\tau) \, d\tau = 0. \tag{5.3}$$

Remark 5.1. If $\{X_t\}_{t \in \mathbb{R}^d}$ is a weakly stationary random field with zero mean, then the auto-covariance $C_X(\tau)$ coincides with the auto-correlation function $R_X(\tau)$ and hence $\{X_t\}_{t \in \mathbb{R}^d}$ is mean ergodic if and only if

$$\lim_{T \to \infty} \frac{1}{(2T)^d} \int_{[-2T,2T]^d} \prod_{i=1}^{d} (1 - \frac{|\tau^i|}{2T}) R_X(\tau) \, d\tau = 0. \tag{5.4}$$

5.2 Ergodic Theorems for Weakly Stationary Random Fields

Ergodic theorems relate functionals calculated along individual sample paths, say the position (time) average, to functionals calculated over the whole distribution, say the expectation. The basic idea is that the two should be close and they should get closer the longer the trajectory we use, because in some sense any one sample path, carried far enough, is representative of the whole distribution.

Since there are many different kinds of functionals, and many different modes of convergence, there are many different kinds of ergodic theorems. The classical ergodic theorems say that position (time) averages converge either in the p-th mean, or almost surely, both implying convergence in distribution or in probability. The traditional centrepiece of the ergodic theorems is Birkhoff's "individual" ergodic theorem, asserting a.s. convergence and in mean convergence, details of this theorem and its proof can be found in Grimmett and Stirzaker (1998). By contrast, the L^2 or mean square ergodic theorem, attributed to Von Neumann is already in our grasp, and holds for weakly stationary random fields. We will see its proof in the next theorem.

Remark 5.2. Recall that, the norm $\| \, . \, \|_2$ of a complex-valued random variable Z is defined by

$$\|Z\|_2 = \left(E[|Z|^2]\right)^{\frac{1}{2}} . \tag{5.5}$$

For a second order random field $\{Y_t\}_{t\in\mathbb{R}^d}$,

$$\begin{aligned}
\| \int_{[-n,n]^d} Y_t \, dt\|_2 &\leq \left(E\left[[\int_{[-n,n]^d} |Y_t| \, dt]^2\right]\right)^{\frac{1}{2}} \\
&= \left(E\left[\int_{[-n,n]^d} |Y_t| \, dt \int_{[-n,n]^d} |Y_s| \, ds\right]\right)^{\frac{1}{2}} \\
&\leq \left(E\left[\int_{[-n,n]^d} \int_{[-n,n]^d} |Y_t||Y_s| \, dt \, ds\right]\right)^{\frac{1}{2}}
\end{aligned}$$

$$\begin{aligned}
\| \int_{[-n,n]^d} Y_t \, dt \|_2 &\le \left(\int_{[-n,n]^d} \int_{[-n,n]^d} E\left[|Y_t Y_s| \right] \, dt \, ds \right)^{\frac{1}{2}} \\
&\le \left(\int_{[-n,n]^d} \int_{[-n,n]^d} \left(E\left[|Y_t|^2 \right] E\left[|Y_s|^2 \right] \right)^{\frac{1}{2}} \, dt \, ds \right)^{\frac{1}{2}} \\
&\qquad \text{as } \; E[|XY|] \le \sqrt{E[|X|^2] E[|Y|^2]} \\
&= \left(\int_{[-n,n]^d} \int_{[-n,n]^d} (E\left[|Y_t|^2 \right])^{\frac{1}{2}} (E\left[|Y_s|^2 \right])^{\frac{1}{2}} \, dt \, ds \right)^{\frac{1}{2}} \\
&= \left(\int_{[-n,n]^d} \|Y_t\|_2 \, dt \int_{[-n,n]^d} \|Y_s\|_2 \, ds \right)^{\frac{1}{2}} \\
&= \int_{[-n,n]^d} \|Y_t\|_2 \, dt.
\end{aligned}$$

Theorem 5.2. *(Ergodic Theorem for Weakly Stationary Random Fields)* *If $\{X_t\}_{t \in \mathbb{R}^d}$ is a weakly stationary random field, then there exists a random variable Y such that $E[Y] = E[X_0]$ and*

$$\lim_{T \to \infty} \frac{1}{(2T)^d} \int_{[-T,T]^d} X_t \, dt = Y \tag{5.6}$$

in the square mean.

Proof. We wish to show that for $\langle X \rangle_n = \frac{1}{(2n)^d} \int_{[-n,n]^d} X_t \, dt$,

$$\| \langle X \rangle_n - \langle X \rangle_m \|_2 \longrightarrow 0 \quad \text{as } n, m \to \infty.$$

Set

$$\mu_n = \inf_{\lambda} \| \int_{[-n,n]^d} \lambda(t) X_t \, dt \|_2 \tag{5.7}$$

where the infinitum is calculated over all functions $\lambda(t) \ge 0$ with $\int_{[-n,n]^d} \lambda(t) \, dt = 1$. For $0 < n_1 \le n_2$,

$$\mu_{n_1} = \inf_{\lambda(t)} \| \int_{[-n_1,n_1]^d} \lambda(t) X_t \, dt \|_2 \;\; \ge \;\; \inf_{\lambda(t)} \| \int_{[-n_2,n_2]^d} \lambda(t) X_t \, dt \|_2 = \mu_{n_2}$$

so that

$$\mu := \lim_{n \to \infty} \mu_n = \inf_n \mu_n \tag{5.8}$$

exists as μ_n is decreasing and $\mu_n \geq 0$ for all n. In particular, $\|\langle X\rangle_n\|_2 \geq \mu$ for all $n > 0$. If $m < n$ then

$$\begin{aligned}
&\|\langle X\rangle_n + \langle X\rangle_m\|_2 \\
&= \|\frac{1}{(2n)^d}\int_{[-n,n]^d} X_t\,dt + \frac{1}{(2m)^d}\int_{[-m,m]^d} X_t\,dt\|_2 \\
&= \|\frac{1}{(2n)^d}\int_{[-n,n]^d\setminus[-m,m]^d} X_t\,dt + \frac{1}{(2n)^d}\int_{[-m,m]^d} X_t\,dt + \frac{1}{(2m)^d}\int_{[-m,m]^d} X_t\,dt\|_2 \\
&= 2\|\int_{[-n,n]^d\setminus[-m,m]^d} \frac{1}{2(2n)^d} X_t\,dt + \int_{[-m,m]^d} \frac{1}{2}\left(\frac{1}{(2n)^d} + \frac{1}{(2m)^d}\right) X_t\,dt\|_2 \\
&= 2\|\int_{[-n,n]^d} \lambda(t) X_t\,dt\|_2
\end{aligned}$$

where

$$\lambda(t) = \begin{cases} \frac{1}{2}\left(\frac{1}{(2m)^d} + \frac{1}{(2n)^d}\right) & \text{if } t \in [-m,m]^d \\ \frac{1}{2(2n)^d} & \text{if } t \in [-n,n]^d \setminus [-m,m]^d, \end{cases}$$

and

$$\begin{aligned}
\int_{[-n,n]^d} \lambda(t)\,dt &= \int_{[-m,m]^d} \frac{1}{2}\left(\frac{1}{(2n)^d} + \frac{1}{(2m)^d}\right) dt + \int_{[-n,n]^d\setminus[-m,m]^d} \frac{1}{2(2n)^d}\,dt \\
&= \frac{(2m)^d}{2}\left(\frac{1}{(2n)^d} + \frac{1}{(2m)^d}\right) + \frac{1}{2(2n)^d}\left[(2n)^d - (2m)^d\right] = 1.
\end{aligned}$$

Thus, as $\mu_n \geq \inf_n \mu_n = \mu$ we have

$$\|\langle X\rangle_n + \langle X\rangle_m\|_2 \geq 2\inf_{\lambda(t)} \|\int_{[-n,n]^d} \lambda(t) X_t\,dt\|_2 = 2\mu_n \geq 2\inf_n \mu_n = 2\mu. \tag{5.9}$$

Consider

$$\begin{aligned}
\|\langle X\rangle_n - \langle X\rangle_m\|_2^2 + \|\langle X\rangle_n + \langle X\rangle_m\|_2^2 &= E[|\langle X\rangle_n - \langle X\rangle_m|^2] + E[|\langle X\rangle_n + \langle X\rangle_m|^2] \\
&= 2E[|\langle X\rangle_n|^2] + 2E[|\langle X\rangle_m|^2] \\
&= 2\|\langle X\rangle_n\|_2^2 + 2\|\langle X\rangle_m\|_2^2.
\end{aligned}$$

Then

$$\|\langle X\rangle_n - \langle X\rangle_m\|_2^2 = 2\|\langle X\rangle_n\|_2^2 + 2\|\langle X\rangle_m\|_2^2 - \|\langle X\rangle_n + \langle X\rangle_m\|_2^2.$$

By inequality (5.9) we have

$$\|\langle X\rangle_n - \langle X\rangle_m\|_2^2 \le 2\|\langle X\rangle_n\|_2^2 + 2\|\langle X\rangle_m\|_2^2 - 4\mu^2$$
$$\le 2\left|\|\langle X\rangle_n\|_2^2 - \mu^2\right| + 2\left|\|\langle X\rangle_m\|_2^2 - \mu^2\right|. \quad (5.10)$$

Now we claim that $\|\langle X\rangle_n\|_2 \longrightarrow \mu$ as $n \to \infty$.

Choose any $\epsilon > 0$ and pick T and λ such that

$$\| \int_{[-T,T]^d} \lambda(t) X_t \, dt\|_2 \le \mu + \epsilon \quad (5.11)$$

where $\lambda(t) \ge 0$ and $\int_{[-T,T]^d} \lambda(t)\, dt = 1$.
Define the moving average

$$Y_k = \int_{[-T,T]^d} \lambda(t) X_{t+k}\, dt. \quad (5.12)$$

It is not difficult to see that $\{Y_k\}_{k\in\mathbb{R}^d}$ has constant mean as $\{X_t\}_{t\in\mathbb{R}^d}$ has constant mean. And for all $k \in \mathbb{R}^d$,

$$\begin{aligned}
\|Y_k\|_2^2 &= E\left[\left|\int_{[-T,T]^d} \lambda(t) X_{t+k}\, dt\right|^2\right] \\
&= E\left[\int_{[-T,T]^d} \lambda(t) X_{t+k}\, dt \overline{\int_{[-T,T]^d} \lambda(s) X_{s+k}\, ds}\right] \\
&= \int_{[-T,T]^d}\int_{[-T,T]^d} \lambda(t)\lambda(s) E[X_{t+k}\overline{X_{s+k}}]\, dt\, ds \\
&= \int_{[-T,T]^d}\int_{[-T,T]^d} \lambda(t)\lambda(s) R_X(t+k, s+k)\, dt\, ds \\
&= \int_{[-T,T]^d}\int_{[-T,T]^d} \lambda(t)\lambda(s) R_X(t, s)\, dt\, ds \text{ as weakly stationary of } \{X_t\} \\
&= \int_{[-T,T]^d}\int_{[-T,T]^d} \lambda(t)\lambda(s) E[X_t\overline{X_s}]\, dt\, ds \\
&= E\left[\left|\int_{[-T,T]^d} \lambda(t) X_{t+k}\, dt\right|^2\right] \\
&= \|Y_0\|_2^2
\end{aligned}$$

We shall show that

$$\|\langle Y\rangle_n - \langle X\rangle_n\|_2 \longrightarrow 0 \ \text{ as } n \to \infty \quad (5.13)$$

where $\langle Y\rangle_n = \dfrac{1}{(2n)^d}\displaystyle\int_{[-n,n]^d} Y_t\, dt.$

Note first that by Remark 5.2 and $\|Y_t\|_2 = \|Y_0\|_2$ for all t, and then use Equations (5.11) and (5.12) we have

$$\begin{aligned}
\|\langle Y\rangle_n\|_2 &\le \frac{1}{(2n)^d}\int_{[-n,n]^d} \|Y_t\|_2\, dt \\
&= \frac{1}{(2n)^d}\|Y_0\|_2 \int_{[-n,n]^d} 1\, dt = \|Y_0\|_2 \\
&= \|\int_{[-T,T]^d} \lambda(t) X_t\, dt\|_2 \le \mu + \epsilon \quad \text{ for all } n.
\end{aligned}$$

Now, by definition of $\langle . \rangle_n$ and Y_t we have

$$\langle Y\rangle_n = \frac{1}{(2n)^d}\int_{[-n,n]^d} Y_t\, dt = \frac{1}{(2n)^d}\int_{[-n,n]^d}\int_{[-T,T]^d} \lambda(s) X_{s+t}\, ds\, dt.$$

By Fubini's Theorem and the change variable of t to $t-s$ we have

$$\begin{aligned}
\langle Y\rangle_n &= \frac{1}{(2n)^d}\int_{[-T,T]^d} \lambda(s)\int_{[-n,n]^d} X_{s+t}\, dt\, ds \\
&= \frac{1}{(2n)^d}\int_{[-T,T]^d} \lambda(s)\int_{\prod_{i=1}^d [-n+s_i, n+s_i]} X_t\, dt\, ds = \int_{[-T,T]^d} \lambda(s)\langle X\rangle_{s,n}\, ds
\end{aligned}$$

where $\langle X\rangle_{s,n} = \dfrac{1}{(2n)^d}\displaystyle\int_{\prod_{i=1}^d [-n+s_i, n+s_i]} X_t\, dt.$

Now use the fact that $\langle X\rangle_{0,n} = \dfrac{1}{(2n)^d}\displaystyle\int_{[-n,n]^d} X_t\, dt = \langle X\rangle_n$, and the triangle inequality to deduce that

$$\begin{aligned}
\|\langle Y\rangle_n - \langle X\rangle_n\|_2 &= \|\int_{[-T,T]^d} \lambda(s)\langle X\rangle_{s,n}\, ds - \langle X\rangle_{0,n}\|_2 \\
&= \|\int_{[-T,T]^d} \lambda(s)\langle X\rangle_{s,n}\, ds - \int_{[-T,T]^d} \lambda(s)\, ds \langle X\rangle_{0,n}\|_2 \\
&= \|\int_{[-T,T]^d} \lambda(s)\left(\langle X\rangle_{s,n} - \langle X\rangle_{0,n}\right) ds\|_2 \\
&\le \int_{[-T,T]^d} \lambda(s)\|\langle X\rangle_{s,n} - \langle X\rangle_{0,n}\|_2\, ds
\end{aligned}$$

computing as in Remark 5.2. Consider

$$\|\langle X\rangle_{s,n} - \langle X\rangle_{0,n}\|_2 = \frac{1}{(2n)^d}\|\int_{\prod_{i=1}^d [-n+s_i, n+s_i]} X_t\, dt - \int_{[-n,n]^d} X_t\, dt\|_2$$

$$= \frac{1}{(2n)^d} \| \int_{\prod_{i=1}^d [-n+s_i, n+s_i] \setminus [-n,n]^d} X_t\,dt + \int_{(\prod_{i=1}^d [-n+s_i, n+s_i]) \cap [-n,n]^d} X_t\,dt$$
$$- \int_{[-n,n]^d \cap (\prod_{i=1}^d [-n+s_i, n+s_i])} X_t\,dt - \int_{[-n,n]^d \setminus (\prod_{i=1}^d [-n+s_i, n+s_i])} X_t\,dt \|_2$$
$$= \frac{1}{(2n)^d} \| \int_{(\prod_{i=1}^d [-n+s_i, n+s_i]) \setminus [-n,n]^d} X_t\,dt - \int_{[-n,n]^d \setminus (\prod_{i=1}^d [-n+s_i, n+s_i])} X_t\,dt \|_2$$
$$\leq \frac{1}{(2n)^d} \left(\| \int_{(\prod_{i=1}^d [-n+s_i, n+s_i]) \setminus [-n,n]^d} X_t\,dt \|_2 + \| \int_{[-n,n]^d \setminus (\prod_{i=1}^d [-n+s_i, n+s_i])} X_t\,dt \|_2 \right)$$
$$\leq \frac{1}{(2n)^d} \left(\int_{(\prod_{i=1}^d [-n+s_i, n+s_i]) \setminus [-n,n]^d} \|X_t\|_2\,dt + \int_{[-n,n]^d \setminus (\prod_{i=1}^d [-n+s_i, n+s_i])} \|X_t\|_2\,dt \right)$$

computing as in remark 5.2. As $\{X_t\}_{t \in \mathbb{R}^d}$ is a weakly stationary random field, $\|X_t\|_2 = \|X_0\|_2$ for all $t \in \mathbb{R}^d$, and we get that

$$\|\langle X \rangle_{s,n} - \langle X \rangle_{0,n}\|_2$$
$$\leq \frac{\|X_0\|_2}{(2n)^d} \left[\int_{(\prod_{i=1}^d [-n+s_i, n+s_i]) \setminus [-n,n]^d} 1\,dt + \int_{[-n,n]^d \setminus (\prod_{i=1}^d [-n+s_i, n+s_i])} 1\,dt \right].$$

Directly computing the integrals we obtain that

$$\|\langle X \rangle_{s,n} - \langle X \rangle_{0,n}\|_2 \leq \frac{\|X_0\|_2}{(2n)^d} \left[2 \prod_{i=1}^d |s_i| + 2 \sum_{i=1}^d |s_i| \prod_{j \neq i} |2n - |s_j|| \right].$$

Thus

$$\|\langle Y \rangle_n - \langle X \rangle_n\|_2 \leq \int_{[-T,T]^d} \lambda(s) \frac{\|X_0\|}{(2n)^d} \left[2 \prod_{i=1}^d |s_i| + 2 \sum_{i=1}^d |s_i| \prod_{j \neq i} |2n - |s_j|| \right] ds$$
$$= \frac{\|X_0\|}{(2n)^d} \int_{[-T,T]^d} \lambda(s) \left[2 \prod_{i=1}^d |s_i| + 2 \sum_{i=1}^d |s_i| \prod_{j \neq i} |2n - |s_j|| \right] ds$$
$$\leq \frac{\|X_0\|}{(2n)^d} \int_{[-T,T]^d} \lambda(s) \left[2 \prod_{i=1}^d T + 2 \sum_{i=1}^d T \prod_{j \neq i} |2n + T| \right] ds$$
$$= \frac{\|X_0\|}{(2n)^d} \left[2T^d + 2T \sum_{i=1}^d |2n + T|^{d-1} \right] \int_{[-T,T]^d} \lambda(s)\,ds$$
$$= \frac{2\|X_0\|_2 T^d}{(2n)^d} + \frac{2dT\|X_0\|_2 |2n+T|^{d-1}}{(2n)^d}.$$

Let $n \to \infty$ to deduce that $\|\langle Y \rangle_n - \langle X \rangle_n\|_2 \longrightarrow 0$ as $n \to \infty$ holds. Use

$\|\langle Y\rangle_n\|_2 \le \mu + \epsilon$ to obtain

$$\mu \le \|\langle X\rangle_n\|_2 \le \|\langle X\rangle_n - \langle Y\rangle_n\|_2 + \|\langle Y\rangle_n\|_2 \le \|\langle X\rangle_n - \langle Y\rangle_n\|_2 + \mu + \epsilon$$
$$\to \mu + \epsilon \quad \text{as } n \to \infty.$$

But ϵ was arbitrary, hence the claim holds. Thus by inequality (5.10), $\{\langle X\rangle_n\}_{n\ge 0}$ is a Cauchy in the square mean. Since $L^2(\Omega)$ is complete, there exists a square integrable random variable Y such that $\langle X\rangle_n \xrightarrow{2} Y$. Moreover, by the same argument from inequality (5.10) we have $\|\langle X\rangle_{T_1} - \langle X\rangle_{T_2}\|_2^2 \le 2\big|\|\langle X\rangle_{T_1}\|_2^2 - \mu^2\big| + 2\big|\|\langle X\rangle_{T_2}\|_2^2 - \mu^2\big|$ for all positive real number T_1 and T_2 so that $\|\langle X\rangle_{T_1} - \langle X\rangle_{T_2}\|_2^2 \to 0$ as $T_1, T_2 \to \infty$. Now for $T \in \mathbb{R}^+$, there exist $n \in \mathbb{N}$ such that $T \le n < T+1$ and then we have

$$\|\langle X\rangle_T - Y\|_2 \le \|\langle X\rangle_T - \langle X\rangle_n\|_2 + \|\langle X\rangle_n - Y\|_2.$$

Thus $\|\langle X\rangle_T - Y\|_2 \to 0$ as $T \to \infty$. Hence $\langle X\rangle_T \xrightarrow{2} Y$, we have that $\langle X\rangle_T \xrightarrow{1} Y$ which implies that $E[\langle X\rangle_T] \to E[Y]$. However, $E[\langle X\rangle_T] = E[X_0]$, whence $E[Y] = E[X_0]$. □

Corollary 5.3. *If $\{X_t\}_{t\in\mathbb{R}^d}$ is a weakly stationary random field with zero mean and auto-correlation function $R_X(\tau)$, then the limit variable*

$$Y = \lim_{T\to\infty} \frac{1}{(2T)^d} \int_{[-T,T]^d} X_t \, dt \tag{5.14}$$

satisfies

$$E[Y] = 0, \quad E[|Y|^2] \le 2^d R_X(0) \tag{5.15}$$

Proof. Consider $\langle X\rangle_T = \dfrac{1}{(2T)^d} \displaystyle\int_{[-T,T]^d} X_t \, dt$. By the proof of Theorem 5.2, there exists a random variable Y such that $\langle X\rangle_T \to Y$ in mean square, so that $E[\langle X\rangle_T] \to E[Y]$. But $E[\langle X\rangle_T] = E[X_1] = 0$ for all T, thus it follows that $E[Y] = 0$. Since

$\langle X\rangle_T \to Y$ in square mean, then $E[|\langle X\rangle_T|^2] \to E[|Y|^2]$. Now

$$
\begin{aligned}
E[|\langle X\rangle_T]|^2] = E\left[\langle X\rangle_T\overline{\langle X\rangle_T}\right] &= E\left[\left(\frac{1}{(2T)^d}\int_{[-T,T]^d} X_t\,dt\right)\left(\overline{\frac{1}{(2T)^d}\int_{[-T,T]} X_s\,ds}\right)\right] \\
&= \frac{1}{(2T)^{2d}}\int_{[-T,T]^d}\int_{[-T,T]^d} E[X_t\overline{X_s}]\,dt\,ds \\
&= \frac{1}{(2T)^{2d}}\int_{[-T,T]^d}\int_{[-T,T]^d} R_X(s-t)\,dt\,ds
\end{aligned}
$$

For each $i = 1, 2, ..., d$, if we let $\tau_i = s_i - t_i$ and $u_i = s_i + t_i$, then $dt_i\,ds_i = |J_i|^{-1} du\,d\tau$ where $|J_i|$ is the Jacobian which is

$$
|J_i| = \begin{vmatrix} \frac{\partial u_i}{\partial t_i} & \frac{\partial u_i}{\partial s_i} \\ \frac{\partial \tau_i}{\partial t_i} & \frac{\partial \tau_i}{\partial s_i} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = 2.
$$

Then we have

$$
\begin{aligned}
E[|\langle X\rangle_T]|^2] &= \frac{1}{2^d(2T)^{2d}}\int_{[-2T,2T]^d}\int_{\prod_{i=1}^d[-2T+|\tau_i|,2T-|\tau_i|]} R_X(\tau)\,du\,d\tau \\
&= \frac{1}{2^d(2T)^{2d}}\int_{[-2T,2T]^d}\prod_{i=1}^d(4T-2|\tau_i|)\,R_X(\tau)\,d\tau \\
&= \frac{1}{(2T)^d}\int_{[-2T,2T]^d}\prod_{i=1}^d\left(1-\frac{|\tau_i|}{2T}\right)R_X(\tau)\,d\tau.
\end{aligned}
$$

Since $-2T \le \tau_i \le 2T$ then $|1-\frac{\tau_i}{2T}| \le 1$ so that we obtain that

$$
\begin{aligned}
E[|Y|^2] = \lim_{T\to\infty} E[|\langle X\rangle_T|^2] &= \lim_{T\to\infty}\frac{1}{(2T)^d}\int_{[-2T,2T]^d}\prod_{i=1}^d\left(1-\frac{|\tau_i|}{2T}\right)R_X(\tau)\,d\tau \\
&\le \lim_{T\to\infty}\frac{1}{(2T)^d}\int_{[-2T,2T]^d} R_X(\tau)\,d\tau \\
&\le \lim_{T\to\infty}\frac{1}{(2T)^d}\int_{[-2T,2T]^d} R_X(0)\,d\tau \\
&= \lim_{T\to\infty}\frac{1}{(2T)^d}[R_X(0)(4T)^d] \\
&= 2^d R_X(0).
\end{aligned}
$$

□

5.3 Ergodic Properties of the Wavelet Transform

By the previous theorem, if $\{X_t\}_{t\in\mathbb{R}^d}$ is a weakly stationary random field it follows immediately that there exists a random variable Y such that

$$\lim_{T\to\infty} \frac{1}{(2T)^d} \int_{[-T,T]^d} X_t \, dt = Y. \tag{5.16}$$

In this section will show that if $\{X_t\}_{t\in\mathbb{R}^d}$ is a random field with either stationary increments or weakly stationary increments, then we have an ergodic theorem and hence ergodic properties for the wavelet transform. Now we consider the following assumption:

Assumption A

1) The random field $\{X_t\}_{t\in\mathbb{R}^d}$ is weakly stationary with zero mean.

2) The mother wavelet function φ is in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.

or

1') The random field $\{X_t\}_{t\in\mathbb{R}^d}$ is weakly stationary.

2') The mother wavelet function φ is in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ and $\hat{\varphi}(0) = 0$.

Assumption B

1) The random field $\{X_t\}_{t\in\mathbb{R}^d}$ has strongly stationary increments and zero mean.

2) The mother wavelet function φ is in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ and $\hat{\varphi}(0) = 0$.

Assumption C

1) The random field $\{X_t\}_{t\in\mathbb{R}^d}$ has weakly stationary increments and zero mean.

2) The mother wavelet function φ is in $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ such that $\hat{\varphi}(0) = 0$ and all first moment are zero.

Theorem 5.4. *Let H be a matrix group and $a \in H$. If either of assumption A, Assumption B or Assumption C is satisfied then there exists a random variable Y*

depending on a such that

$$\lim_{T\to\infty} \frac{1}{(2T)^d} \int_{[-T,T]^d} CW_X^a(t)\, dt = Y \tag{5.17}$$

in the mean square sense with $E[Y] = 0$ *and*

$$E[|Y|^2] \leq 2^d R_{CW_X^a}(0). \tag{5.18}$$

Proof. By Theorems 4.1, 4.2 and 4.4 for fixed $a \in H$, the random field $\{CW_X^a(t)\}_{t\in\mathbb{R}^d}$ is weakly stationary with zero mean and auto-correlation $R_{CW_X^a}(\tau)$. Hence by Corollary 5.3, there exists a random variable Y such that

$$\lim_{T\to\infty} \frac{1}{(2T)^d} \int_{[-T,T]^d} CW_X^a(t)\, dt = Y \tag{5.19}$$

satisfying $E[Y] = 0$ and

$$E[|Y|^2] = 2^d R_{CW_X^a}(0). \tag{5.20}$$

□

Theorem 5.5. *Let H be a matrix group, $a \in H$. If assumption A is satisfied and random field $\{X_t\}_{t\in\mathbb{R}^d}$, with nonnegative real value auto-correlation function, is ergodic in mean, then $\{CW_X(t,a)\}_{t\in\mathbb{R}^d}$ is also ergodic in mean.*

Proof. Since $\{X_t\}_{t\in\mathbb{R}^d}$ is a weakly stationary random field, the auto-correlation function $R_X(\tau)$ is a positive definite function and hence it has a power spectral representation, by Bochner's Theorem, which is given by equation (4.1) in Section 4.1. Since $\{X_t\}_{t\in\mathbb{R}^d}$ is a weakly stationary random field, then by Theorem 4.1, $\{CW_X^a(t)\}_{t\in\mathbb{R}^d}$ is weakly stationary with zero mean and the auto-correlation function $R_{CW_X^a}(\tau)$ has the power spectral representation given by equation (4.5) in Section 4.1. As $\{X_t\}_{t\in\mathbb{R}^d}$ is a mean ergodic and weakly stationary random field with zero mean, by Remark 5.1

$$\lim_{T\to\infty} \frac{1}{(2T)^d} \int_{[-2T,2T]^d} \prod_{i=1}^{d} \left(1 - \frac{|\tau^i|}{2T}\right) R_X(\tau)\, d\tau = 0. \tag{5.21}$$

Now, using equation (4.5) in Section 4.1, remark 2.1 (ii) and then equation (4.1) in Section 4.1 consecutively, we have

$$\begin{aligned}
&\left| \frac{1}{(2T)^d} \int_{[-2T,2T]^d} \prod_{i=1}^{d} \left(1 - \frac{|\tau^i|}{2T}\right) R_{CW_X^a}(\tau)\, d\tau \right| \\
&\quad \le \frac{1}{(2T)^d} \int_{[-2T,2T]^d} \prod_{i=1}^{d} \left(1 - \frac{|\tau^i|}{2T}\right) |\det a| \left| \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda} |\hat{\varphi}(a^T\lambda)|^2 \, dF_X(\lambda) \right| d\tau \\
&\quad \le \frac{|\det a| \|\varphi\|_{L^1}^2}{(2T)^d} \int_{[-2T,2T]^d} \prod_{i=1}^{d} \left(1 - \frac{|\tau^i|}{2T}\right) \left| \int_{\mathbb{R}^d} e^{-i\tau\cdot\lambda} \, dF_X(\lambda) \right| d\tau \\
&\quad = |\det a| \|\varphi\|_{L^1}^2 \frac{1}{(2T)^d} \int_{[-2T,2T]^d} \prod_{i=1}^{d} \left(1 - \frac{|\tau^i|}{2T}\right) |R_X(\tau)| \, d\tau.
\end{aligned}$$

Thus by equation (5.21) we have

$$\lim_{T\to\infty} \frac{1}{(2T)^d} \int_{[-2T,2T]^d} \prod_{i=1}^{d} \left(1 - \frac{|\tau^i|}{2T}\right) R_{CW_X^a}(\tau)\, d\tau = 0$$

so that by Remark 5.1 $\{CW_X^a(t)\}_{t\in\mathbb{R}^d}$ is mean ergodic. □

CHAPTER VI

WAVELET REPRESENTATION OF RANDOM FIELDS

We will establish the existence of Brownian motion and Brownian sheets by providing an explicit series expansion. The calculations we make with this series are quite basic, but still require some facts about function spaces. In this chapter we begin in the first section with reviewing the construction of the Haar wavelet representation of Brownian motion, and then in the following section we develop a framework for constructing the Haar wavelet representation of a Brownian sheet. Finally, we construct the wavelet representation of a Brownian sheet in finite dimension via arbitrary compactly supported wavelet functions. In the particular one dimensional case, we construct a wavelet representation of Brownian motion via a compactly supported wavelet basis of $L^2[0,1]$.

Remark 6.1. We first review the important properties of a complete orthonormal basis of a Hilbert space, as well as Parseval's identity.

(1) If $\{\phi_n \in L^2[0,1] : n \in I\}$ is a complete orthonormal basis of $L^2[0,1]$, then for all $f \in L^2[0,1]$ we have the representation

$$f = \sum_{n \in I} \langle f, \phi_n \rangle \phi_n.$$

Furthermore, by Parseval's identity we have for all $f, g \in L^2[0,1]$,

$$\int_0^1 f(x)\overline{g(x)}\,dx = \langle f, g \rangle = \sum_{n \in I} \langle f, \phi_n \rangle \overline{\langle g, \phi_n \rangle}.$$

Next, if $f(x) = \chi_{[0,s]}(x)$ and $g(x) = \chi_{[0,t]}(x)$ for $s, t \in [0,1]$ we get that

$$\sum_{n\in I} \int_0^s \phi_n(x)\,dx \int_0^t \overline{\phi_n(y)}\,dy = \sum_{n\in I} \langle \chi_{[0,s]}, \phi_n \rangle \overline{\langle \chi_{[0,t]}, \phi_n \rangle} = \langle \chi_{[0,s]}, \chi_{[0,t]} \rangle,$$

that is

$$\begin{aligned} \sum_{n\in I} \int_0^s \phi_n(x)\,dx \int_0^t \overline{\phi_n(y)}\,dy &= \int_0^1 \chi_{[0,s]}(x)\chi_{[0,t]}(x)\,dx \\ &= \begin{cases} \int_0^s 1\,dx & \text{if } s \le t \\ \int_0^t 1\,dx & \text{if } t \le s \end{cases} = \begin{cases} s & \text{if } s \le t \\ t & \text{if } t \le s \end{cases} = \min(s,t). \end{aligned}$$

(2) We apply this idea to the d-dimensional case. If $\{\phi_n \in L^2([0,1]^d) : n \in I\}$ is an orthonormal basis of $L^2([0,1]^d)$, then for $f(x) = \chi_{\prod_{i=1}^d [0,s_i]}(x)$ and $g(x) = \chi_{\prod_{i=1}^d [0,t_i]}(x)$, $x = (x_1, ..., x_d) \in [0,1]^d$, we get that

$$\begin{aligned} \sum_{n\in I} \int_{\prod_{i=1}^d [0,s_i]} \phi_n(x)\,dx \int_{\prod_{i=1}^d [0,t_i]} \overline{\phi_n(x)}\,dx &= \langle \chi_{\prod_{i=1}^d [0,s_i]}, \chi_{\prod_{i=1}^d [0,t_i]} \rangle_{L^2([0,1]^d)} \\ &= \int_{[0,1]^d} \chi_{\prod_{i=1}^d [0,s_i]}(x)\chi_{\prod_{i=1}^d [0,t_i]}(x)\,dx \\ &= \int_{[0,1]^d} \prod_{i=1}^d \chi_{[0,s_i]}(x_i)\chi_{[0,t_i]}(x_i)\,dx_i \\ &= \prod_{i=1}^d \int_{[0,1]} \chi_{[0,s_i]}(x_i)\chi_{[0,t_i]}(x_i)\,dx_i \\ &= \prod_{i=1}^d \min(s_i, t_i). \end{aligned}$$

6.1 Haar Wavelet Representation of Brownian Motion

Consider the Haar function $H : \mathbb{R} \to \mathbb{R}$ given by $H(t) = \chi_{[0,\frac{1}{2})}(t) - \chi_{[\frac{1}{2},1)}(t)$. Let $n \in \mathbb{N}$ be arbitrary. Then n can be written uniquely in the form $n = 2^j + k$ for $j = 0, 1, 2, ...$ and $k = 0, 1, ..., 2^j - 1$. We define

$H_n(t) = 2^{\frac{j}{2}} H(2^j t - k)$ for $n = 2^j + k$, where $j = 0, 1, 2, ...$ and $k = 0, 1, ..., 2^j - 1$, and $H_0(t) = 1$. One can show that $\{H_n : n \in \mathbb{N} \cup \{0\}\}$ is a complete orthonormal sequence for $L^2[0,1]$. Details of the proof can be found in Daubechies (1992), pp.10-13, and Walnut (2001), pp.115-123.

Remark 6.2. Since $n = 2^j + k$ for $j = 0, 1, 2, ...$ and $k = 0, 1, ..., 2^j - 1$, we can represent $\{H_n\}$ as a doubly indexed sequence, indexed by j and k. For each j, the function H_{2^j+k} is simply a translation of H_{2^j} by k, and the functions H_{2^j+k} have disjoint supports for different k, so that for all $x \in [0,1], H_{2^j+k}(x) \neq 0$ for at most one k.

Next, consider the triangle function $T : \mathbb{R} \to \mathbb{R}$ given by

$$T(t) = 2t\chi_{[0,\frac{1}{2})} + 2(1-t)\chi_{[\frac{1}{2},1)}.$$

Then for $n = 2^j + k$ with $j = 0, 1, 2, ...,$ and $k = 0, 1, ..., 2^j - 1$ we set $T_n(t) = T(2^j t - k)$ and we also set $T_0(t) = t$.

Remark 6.3.

1) $0 \leq T_n(t) \leq 1$ for all $t \in [0,1]$ and all n.

2) $\int_0^t H(x)\, dx = \frac{1}{2}T(t)$ for all $t \in [0,1]$.

3) For $n = 2^j + k$ with $j = 0, 1, 2, ...,$ and $k = 0, 1, ..., 2^j - 1$ by note 2) we have

$$\int_0^t H_n(x)\, dx = \lambda_n T_n(t) \text{ where } \lambda_0 = 1, \lambda_n = 2^{-\frac{j}{2}-1}.$$

Proof. 1) Since $0 \leq T(t) \leq 1$, we have $0 \leq T_n(t) = T(2^j t - k) \leq 1$ where $n = 2^j + k$ for $j = 0, 1, 2, ...,$ and $k = 0, 1, ..., 2^j - 1$.

2) Let $t \in \mathbb{R}$. Obviously we have

$$\int_0^t H(x)\, dx = t\chi_{[0,\frac{1}{2})} + (1-t)\chi_{[\frac{1}{2},1)} = \frac{1}{2}T(t).$$

3) Let $n = 2^j + k$ with $j = 0, 1, 2, ...,$ and $k = 0, 1, ..., 2^j - 1$. Then

$$\begin{aligned}\int_0^t H_n(x)\,dx &= 2^{\frac{j}{2}} \int_0^t H(2^j x - k)\,dx \\ &= 2^{\frac{j}{2}} \int_{-k}^{2^j t - k} H(x) 2^{-j}\,dx \qquad \text{as } x \mapsto 2^{-j}(x+k) \\ &= 2^{-\frac{j}{2}} \left[\int_0^{2^j t - k} H(x)\,dx + \int_{-k}^0 H(x)\,dx \right] \\ &= 2^{-\frac{j}{2}-1} T(2^j t - k) \\ &= 2^{-\frac{j}{2}-1} T_n(t).\end{aligned}$$

□

Remark 6.4. Since $n = 2^j + k$ for $j = 0, 1, 2, ...,$ and $k = 0, 1, ..., 2^j - 1$ we can represent $\{T_n\}$ as a doubly indexed sequence indexed by j and k. For each j, the T_{2^j+k} are simply translations of T_{2^j} by k, also the T_{2^j+k} have disjoint supports, so that for all $x \in [0, 1], T_{2^j+k}(x) \neq 0$ for at most one k, and hence by remark 6.3 1), $0 \leq \sum_{k=0}^{2^j-1} T(2^j x - k) \leq 1$ for all $x \in [0, 1]$.

Theorem 6.1. *(Steele, 2000) If $\{Z_n\}$ is a sequence of independent Gaussian variables with mean 0 and variance 1, then the series defined by*

$$X_t(\omega) = \sum_{n=0}^{\infty} \lambda_n Z_n(\omega) T_n(t) \tag{6.1}$$

converges uniformly on $[0, 1]$ with probability one. Moreover, the process $\{X_t\}_{t \in [0,1]}$ defined by the limit is a standard Brownian motion.

Proof. (I) First we verify Uniform convergence with probability one. For each $n \in \mathbb{N}$ there exists a unique $j \geq 0$ such that $n \in [2^j, 2^{j+1})$. Then $\ln n < \ln 2^{j+1} = (j+1)\ln 2 < j + 1$. Then by Lemma G.2 in Appendix G, there exist a random variable C such that $|Z_n| \leq C\sqrt{\ln n} < C\sqrt{j+1}$ a.e. ω for all $n \geq 2$.

Let $J \in \mathbb{N}$ and $M = 2^J$. We have

$$\begin{aligned}
\sum_{n=M}^{\infty} \lambda_n |Z_n(\omega)| T_n(t) &\le C \sum_{n=M}^{\infty} \lambda_n \sqrt{\ln n}\, T_n(t) \quad \text{a.e. } \omega \\
&\le C \sum_{j=J}^{\infty} \sum_{k=0}^{2^j-1} 2^{-\frac{j}{2}-1} \sqrt{j+1}\, T_{2^j+k}(t) \quad \text{a.e. } \omega \\
&= C \sum_{j=J}^{\infty} 2^{-\frac{j}{2}-1} \sqrt{j+1} \sum_{k=0}^{2^j-1} T_{2^j+k}(t) \quad \text{a.e. } \omega.
\end{aligned}$$

By Remark 6.4 we have $\sum_{k=0}^{2^j-1} T_{2^j+k}(t) \le 1$, so that

$$\sum_{n=M}^{\infty} \lambda_n |Z_n(\omega)| T_n(t) \le C \sum_{j=J}^{\infty} 2^{-\frac{j}{2}-1} \sqrt{j+1} < \infty \quad \text{a.e. } \omega. \tag{6.2}$$

We can see that, if $J \to \infty$ then by Equation (6.2), $\sum_{n=M}^{\infty} \lambda_n |Z_n(\omega)| T_n(t) \to 0$ uniformly on $[0, 1]$ a.e. ω, so that $\sum_{n=0}^{N} \lambda_n |Z_n(\omega)| T_n(t)$ is a uniformly Cauchy sequence, and hence it is uniformly convergent. Thus $\sum_{n=0}^{\infty} \lambda_n Z_n(\omega) T_n(t)$ is uniformly and absolutely convergent a.e. ω. It follows that the sample paths of $\{X_t\}_{t\in[0,1]}$ are continuous with probability one.

(II) Next we calculate the covariance functions, consider

$$\begin{aligned}
E[X_s X_t] &= E\left[\sum_{n=0}^{\infty} \lambda_n Z_n(\omega) T_n(s) \sum_{m=0}^{\infty} \lambda_m Z_m(\omega) T_m(t)\right] \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \lambda_n \lambda_m E\left[Z_n(\omega) Z_m(\omega)\right] T_n(s) T_m(t).
\end{aligned}$$

Since $\{Z_n\}_{n\ge 0}$ are independent, have mean 0 and variance 1, we have by corollary 3.2,

$$\begin{aligned}
E[X_s X_t] = \sum_{n=0}^{\infty} \lambda_n^2 T_n(s) T_n(t) &= \sum_{n=0}^{\infty} [\lambda_n T_n(s)] \cdot [\lambda_n T_n(t)] \\
&= \sum_{n=0}^{\infty} \int_0^s H_n(u)\, du \int_0^t H_n(v)\, dv.
\end{aligned}$$

By Remark 6.1 (1) and as $\{H_n\}_{n\geq 0}$ is a complete orthonormal basis of $L^2([0,1])$ we obtain that

$$E[X_sX_t] = \min(s,t).$$

Since $E[X_t] = E[\sum_{n=0}^{\infty} \lambda_n Z_n(\omega)T_n(t)] = \sum_{n=0}^{\infty} \lambda_n E[Z_n(\omega)]T_n(t) = 0$, we have

$$\mathrm{Cov}(X_s, X_t) = E[X_sX_t] - E[X_s]E[X_t] = E[X_sX_t] = \min(s,t).$$

(III) Now we verify that $\{X_t\}_{t\in[0,1]}$ is a Gaussian process.

Let $t = \{t_1, t_2, ..., t_m : t_i < t_j, i < j\}$ be any choice of finite sequence with coresponding vector $X = \{X_{t_1}, ..., X_{t_m}\}$. Consider for $s = (s_1, ..., s_m) \in [0,1]^m$

$$\begin{aligned} E[e^{is^TX}] &= E\left[\exp\left(i\sum_{j=1}^{m} s_jX_{t_j}\right)\right] = E[\exp(i\sum_{j=1}^{m} s_j \sum_{n=0}^{\infty} \lambda_n Z_n T_n(t_j))] \\ &= E[\exp(i\sum_{n=0}^{\infty} \lambda_n Z_n \sum_{j=1}^{m} s_j T_n(t_j))]. \end{aligned}$$

Since $\{Z_n\}$ are independent processes we have by Theorem 3.1,

$$E[e^{is^TX}] = \prod_{n=0}^{\infty} E[\exp(i\lambda_n Z_n \sum_{j=1}^{m} s_j T_n(t_j))].$$

As each Z_n is Gaussian, we have by remark D.3,

$$E[e^{is^TX}] = \prod_{n=0}^{\infty} \exp\left(i(\sum_{j=1}^{m} s_j T_n(t_j))E[\lambda_n Z_n] - \frac{1}{2}(\sum_{j=1}^{m} s_j T_n(t_j))^2(\mathrm{Var}(\lambda_n Z_n))\right).$$

As $E[cX] = cE[X], \mathrm{Var}(cX) = c^2\mathrm{Var}(X)$ and $\{Z_n\}$ has zero mean and variance 1 we have

$$\begin{aligned} E[e^{is^TX}] &= \prod_{n=0}^{\infty} \exp\left(-\frac{1}{2}\lambda_n^2(\sum_{j=1}^{m} s_j T_n(t_j))^2\right) = \exp\left(-\frac{1}{2}\sum_{n=0}^{\infty} \lambda_n^2(\sum_{j=1}^{m} s_j T_n(t_j))^2\right) \\ &= \exp\left(-\frac{1}{2}\sum_{n=0}^{\infty} \lambda_n^2 \sum_{j=1}^{m}\sum_{k=1}^{m} s_j s_k T_n(t_j)T_n(t_k)\right) \\ &= \exp\left(-\frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{m} s_j s_k \sum_{n=0}^{\infty} \lambda_n T_n(t_j)\lambda_n T_n(t_k)\right) \\ &= \exp\left(-\frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{m} s_j s_k \sum_{n=0}^{\infty} \int_0^{t_j} H_n(u)\,du \int_0^{t_k} H_n(v)\,dv\right). \end{aligned}$$

By the calculation in part (II) we have

$$E[e^{is^TX}] = \exp\left(-\frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{m} s_j s_k \mathrm{Cov}(t_j, t_k)\right)$$
$$= \exp\left(is^T E[X] - \frac{1}{2}s^T\sigma s\right) \quad \text{as } E[X_t] = 0 \text{ for all } t$$

where $\sigma = [\mathrm{Cov}(t_j, t_k)]_{m\times m}$. Hence, by Remark D.3 in Appendix D , $X = \{X_{t_1}, ..., X_{t_m}\}$ has the multivariate Gaussian distribution. This shows that $\{X_t\}_{t\in[0,1]}$ is a Gaussian process. (The definition of multivariate Gassian distribution and Gaussian process can be found in Appendix D.)

Note that $X_0(\omega) = \sum_{n=0}^{\infty} \lambda_n Z_n(\omega) T_n(0) = 0$ as $T_n(0) = 0$ for all n.

Therefore, by (I), (I), (II) and Lemma E.2 in Appendix E, $\{X_t\}_{t\in[0,1]}$ is standard Brownian motion on $[0, 1]$. $\square$

6.2 Haar Wavelet Representation of a Brownian Sheet

Let $x = (x_1, x_2, ..., x_d) \in \mathbb{R}^d$. For each $i = 1, 2, ..., d$ we define a doubly-indexed family of Haar functions by dilating and translating as

$$H_{j_i,k_i}(x_i) = 2^{\frac{j_i}{2}} H(2^{j_i}x_i - k_i) \quad \text{for } j_i = 0, 1, 2, ... \text{ and } k_i = 0, ..., 2^{j_i} - 1 \tag{6.3}$$

$$\text{and} \qquad H_{-1,0}(x_i) \equiv 1. \tag{6.4}$$

Using the notation $J = \{j = (j_1, j_2, ..., j_d) : j_i = -1, 0, 1, 2, ...\}$ and $\mathcal{K}^j = \{k = (k_1, k_2, ..., k_d)$: if $j_i = -1$ then $k_i = 0$, else $k_i = 0, 1, ..., 2^{j_i} - 1\}$, then for $j \in J, k \in \mathcal{K}^j$ we define

$$H^d_{j,k}(x) = \prod_{i=1}^{d} H_{j_i,k_i}(x_i) \ \text{ for all } x = (x_1, ..., x_d) \in \mathbb{R}^d.$$

We know that the family $\{H_{-1,0},\ \ H_{j,k}\ ; j = 0, 1, 2, ...$ and $k = 0, 1, ..., 2^j - 1\}$ is a complete orthonormal basis of $L^2[0, 1]$, hence by Remark 2.9 we obtain that $\{H^d_{j,k} : j \in J, k \in \mathcal{K}^j\}$ is a complete orthonormal basis for $L^2([0, 1]^d)$.

Remark 6.5. For each $i = 1, 2, ..., d$ if $j_i = 0, 1, 2, ...$ and $k_i = 0, ..., 2^{j_i} - 1$ we have

$$H(2^{j_i}x_i - k_i) = \chi_{[2^{-j_i}k_i\ ,\ \ 2^{-j_i-1}+2^{-j_i}k_i)} - \chi_{[2^{-j_i-1}+2^{-j_i}k_i\ ,\ \ 2^{-j_i}+2^{-j_i}k_i)}.$$

We can see that for $t \in [0, 1]^d$ and for each $j \in J$,

1) $H^d_{j,k}(x)$ are simply translations of $H^d_{j,\bar{0}}(x)$.

2) $H^d_{j,k}(x)$ have disjoint support, that is $H^d_{j,k}(x) \neq 0$ for at most one k at each level j, and hence $-1 \leq \sum_{k \in \mathcal{K}^j} H^d_{j,k}(x) \leq 1$.

Next, consider the triangle function $T : \mathbb{R} \to \mathbb{R}$ such that

$$T(t) = 2t\chi_{[0,\frac{1}{2})}(t) + 2(1 - t)\chi_{[\frac{1}{2},1)}(t).$$

Now for $j_i = 0, 1, 2, ...$ and $k = 0, ..., 2^{j_i} - 1$, we define a doubly -indexed family of triangle functions as

$$T_{j_i,k_i}(t_i) = T(2^{j_i}t_i - k_i)$$

$$\text{and } T_{-1,0}(t_i) = t_i\chi_{[0,1]}(t_i).$$

For $j \in J, k \in \mathcal{K}^j$ we define

$$T^d_{j,k}(t) = \prod_{i=1}^{d} T_{j_i,k_i}(t_i)$$

Remark 6.6. Let $j \in J, k \in \mathcal{K}^j$. We have the following properties.

1) As $0 \leq T(t) \leq 1$ then also $0 \leq T^d_{j,k}(t) \leq 1$ for all $t \in [0, 1]^d$. Moreover, $T^d_{j,k}(x)$ are simply translations of $H^d_{j,\bar{0}}(x)$ and have disjoint support, that is $T^d_{j,k}(x) \neq 0$ for at most one k at each level j, and hence $0 \leq \sum_{k \in \mathcal{K}^j} T^d_{j,k}(x) \leq 1$.

2) As $T(t) = 2\int_0^t H(x)\,dx$ for all $t \in [0, 1]$, then for each i,

$$T_{j_i,k_i}(t_i) = T(2^{j_i}t_i - k_i) = 2\int_0^{2^{j_i}t_i-k_i} H(x_i)\,dx_i.$$

3) $\int_{\prod_{i=1}^d [0,t_i]} H^d_{j,k}(x)\,dx = \lambda_j T^d_{j,k}(t)$ for $\lambda_j = \prod_{j=1}^{d} \lambda_{j_i}$

where $\lambda_{j_i} = \begin{cases} 2^{-\frac{j_i}{2}-1} & \text{if } j_i \neq -1 \\ 1 & \text{if } j_i = -1. \end{cases}$

In fact, $\displaystyle\int_0^t H(x)\,dx = t\chi_{[0,\frac{1}{2})} + (1-t)\chi_{[\frac{1}{2},1)} = \frac{1}{2}T(t)$, gives

$$\int_{\prod_{i=1}^d [0,t_i]} H_{j,k}^d(x)\,dx = \int_{\prod_{i=1}^d [0,t_i]} \prod_{i=1}^d H_{j_i,k_i}(x_i)\,d(x_1,...,x_d)$$

$$= \prod_{i=1}^d \begin{cases} 2^{\frac{j_i}{2}} \displaystyle\int_{[0,t_i]} H(2^{j_i}x_i - k_i)\,dx_i & \text{if } j_i \neq -1 \\ \displaystyle\int_{[0,t_i]} 1\,dx_i & \text{if } j_i = -1. \end{cases}$$

By changing variables, $x_i \mapsto 2^{-j_i}(x_i + k_i)$, we have

$$\int_{\prod_{i=1}^d [0,t_i]} H_{j,k}^d(x)\,dx = \prod_{i=1}^d \begin{cases} 2^{-\frac{j_i}{2}} \displaystyle\int_{-k_1}^{2^{j_i}t_i - k_i} H(x_i)\,dx_i & \text{if } j_i \neq -1 \\ t_i & \text{if } j_i = -1. \end{cases}$$

Since $\displaystyle\int_{-k_i}^0 H(x_i)\,dx_i = 0$ we have

$$\int_{\prod_{i=1}^d [0,t_i]} H_{j,k}^d(x)\,dx = \prod_{i=1}^d \begin{cases} 2^{-\frac{j_i}{2}} \displaystyle\int_{0}^{2^{j_i}t_i - k_i} H(x_i)\,dx_i & \text{if } j_i \neq -1 \\ t_i & \text{if } j_i = -1. \end{cases}$$

By 2) we obtain that

$$\int_{\prod_{i=1}^d [0,t_i]} H_{j,k}^d(x)\,dx = \prod_{i=1}^d \begin{cases} 2^{-\frac{j_i}{2}-1} T_{j_i,k_i}(t_i) & \text{if } j_i \neq -1 \\ t_i & \text{if } j_i = -1 \end{cases}$$

$$= \left(\prod_{i=1}^d \lambda_{j_i}\right)\left(\prod_{i=1}^d T_{j_i,k_i}(t_i)\right) \text{ where } \lambda_{j_i} = \begin{cases} 2^{-\frac{j_i}{2}-1} & \text{if } j_i \neq -1 \\ 1 & \text{if } j_i = -1 \end{cases}$$

$$= \lambda_j \cdot T_{j,k}^d(t) \quad \text{where } \lambda_j = \prod_{i=1}^d \lambda_{j_i} \leq 1.$$

Remark 6.7. Fix an integer j_0 (in particular, $j_0 = -1$), let $J = \{(j_1, j_2, ..., j_d) : j_i \geq j_0 \text{ for all } i\}$ and define $|j| = \max\limits_{i=1,\ldots,d}\{|j_i|\}$ for $j = (j_1, ..., j_d) \in J$. Then $\sum\limits_{\substack{j\in J\\ |j|\geq N}} \prod\limits_{i=1}^{d} (j_i+1)^{\frac{1}{2}} 2^{-\frac{j_i}{2}} \to 0$ as $N \to \infty$.

Proof. For each $n \in \mathbb{N} \cup \{0, -1\}$, let $S_n = \{j \in J : |j| = n\}$.
Then $\bigcup\limits_{n=-1}^{\infty} S_n = J$, $|S_n| \leq n^d$, and $S_m \cap S_n = \phi$ if $m \neq n$.
Thus

$$\sum_{\substack{j\in J\\ |j|\geq N}} \prod_{i=1}^{d} (j_i+1)^{\frac{1}{2}} 2^{-\frac{j_i}{2}} \leq \sum_{n=N}^{\infty} \sum_{j\in S_n} \prod_{i=1}^{d} (j_i+1)^{\frac{1}{2}} 2^{-\frac{j_i}{2}}.$$

Now if $j \in S_n$, there exists $k = k(n, j)$ such that $j_k = n$. Separating the corresponding factor out,

$$\begin{aligned}
\sum_{\substack{j\in J\\ |j|\geq N}} \prod_{i=1}^{d} (j_i+1)^{\frac{1}{2}} 2^{-\frac{j_i}{2}} &\leq \sum_{n=N}^{\infty} \sum_{j\in S_n} (n+1)^{\frac{1}{2}} 2^{-\frac{n}{2}} \prod_{\substack{i=1,\\ i\neq k}}^{d} (j_i+1)^{\frac{1}{2}} 2^{-\frac{j_i}{2}} \\
&= \sum_{n=N}^{\infty} (n+1)^{\frac{1}{2}} 2^{-\frac{n}{2}} \sum_{j\in S_n} \prod_{\substack{i=1,\\ i\neq k}}^{d} \left(\frac{j_i+1}{2^{j_i}}\right)^{\frac{1}{2}} \\
&\leq \sum_{n=N}^{\infty} (1+n)^{\frac{1}{2}} 2^{-\frac{n}{2}} \sum_{j\in S_n} 1 = \sum_{n=N}^{\infty} (1+n)^{\frac{1}{2}} 2^{-\frac{n}{2}} |S_n| \\
&\leq \sum_{n=N}^{\infty} (1+n)^{\frac{1}{2}} 2^{-\frac{n}{2}} n^d \to 0 \quad \text{as } N \to \infty.
\end{aligned}$$

This proves the remark. □

Theorem 6.2. *If $\{Z_{j,k} : j \in J \text{ and } k \in \mathcal{K}^j\}$ is a collection of independent Gaussian variables with mean 0 and variance 1, then the series defined by*

$$X_t(\omega) = \sum_{j\in J} \sum_{k\in\mathcal{K}^j} Z_{j,k}(\omega) \lambda_j T^d_{j,k}(t) \tag{6.5}$$

converges uniformly and absolutely on $[0,1]^d$ with probability one. Moreover, the random field $\{X_t\}_{t\in[0,1]^d}$ defined by the limit is a Brownian sheet with zero mean.

Proof. (I) Verify uniform convergence with probability 1. For each $i = 1, ..., d$, we have $\ln(2^{j_i} + k_i) < \ln(2^{j_i+1})$ as $k < 2^{j_i}$, and hence $\ln(2^{j_i} + |k_i|) < j_i + 1$. Then by Lemma G.3 in Appendix G, there exist a random variable C

$$|Z_{j,k}| \leq C \left(\sum_{i=1}^{d} \ln(2^{j_i} + |k_i|) \right)^{\frac{1}{2}} \leq C \left(\sum_{i=1}^{d} (j_i + 1) \right)^{\frac{1}{2}} \quad \text{a.e. } \omega.$$

It is easy to see that, if $j_i \geq 1$ for all i then $\left(\sum_{i=1}^{d} (j_i + 1) \right)^{\frac{1}{2}} \leq \prod_{i=1}^{d} \sqrt{j_i + 1}$, so that the above inequality becomes $|Z_{j,k}| \leq C \prod_{i=1}^{d} \sqrt{j_i + 1}$. Now, consider

$$\begin{aligned} \sum_{\substack{j \in J \\ |j| \geq N}} \sum_{k \in \mathcal{K}^j} |Z_{j,k}| \lambda_j T^d_{j,k}(t) &\leq \sum_{\substack{j \in J \\ |j| \geq N}} \sum_{k \in \mathcal{K}^j} C \prod_{i=1}^{d} \sqrt{j_i + 1} \lambda_j T^d_{j,k}(t) \text{ a.e. } \omega \\ &= C \sum_{\substack{j \in J \\ |j| \geq N}} \prod_{i=1}^{d} \lambda_j \sqrt{j_i + 1} \sum_{k \in \mathcal{K}^j} T^d_{j,k}(t) \quad \text{a.e. } \omega. \end{aligned}$$

By definition of λ and as by Remark 6.6 (1), $0 \leq \sum_{k \in \mathcal{K}^j} T^d_{j,k}(t) \leq 1$, we have

$$\sum_{\substack{j \in J \\ |j| \geq N}} \sum_{k \in \mathcal{K}^j} |Z_{j,k}| \lambda_j T^d_{j,k}(t) \leq C \sum_{\substack{j \in J \\ |j| \geq N}} \prod_{i=1}^{d} 2^{-\frac{j_i}{2} - 1} \sqrt{j_i + 1} \text{ a.e. } \omega$$

Then by Remark 6.7,

$$\sum_{\substack{j \in J \\ |j| \geq N}} \sum_{k \in \mathcal{K}^j} |Z_{j,k}| \lambda_j T^d_{j,k}(t) \leq \prod_{i=1}^{d} C \sum_{\substack{j \in J \\ |j| \geq N}} 2^{-\frac{j_i}{2} - 1} \sqrt{j_i + 1} \longrightarrow 0 \text{ a.e. } \omega \text{ as } N \to \infty.$$

This show that $\sum_{\substack{j \in J \\ |j| \leq N}} \sum_{k \in \mathcal{K}^j} |Z_{j,k}| \lambda_j T^d_{j,k}(t)$ is a uniformly Cauchy sequence, and hence it converges uniformly. Thus, $\sum_{j \in J} \sum_{k \in \mathcal{K}^j} Z_{j,k} \lambda_j T^d_{j,k}(t)$ is absolutely and uniformly convergent a.e. ω. In particular, the paths of the random field $\{X_t\}_{t \in [0,1]^d}$ are continuous with probability one.

(II) Next we calculate the auto-covariance function. Consider

$$\begin{aligned}\sum_{j\in J}\sum_{k\in\mathcal{K}^j} E[|\lambda_j Z_{j,k} T^d_{j,k}(t)|^2] &= \sum_{j\in J}\sum_{k\in\mathcal{K}^j} |\lambda_j|^2 E[|Z_{j,k}|^2]|T^d_{j,k}(t)|^2 \\ &= \sum_{j\in J}|\lambda_j|^2 \sum_{k\in\mathcal{K}^j}|T^d_{j,k}(t)|^2 \le \sum_{j\in J}|\lambda_j|^2 \\ &= \prod_{i=}^{d}\sum_{j\in J} 2^{-j_i-2} < \infty,\end{aligned}$$

then we have by Beppo Levi theorem

$$\begin{aligned}E[X_sX_t] &= E\left[\left(\sum_{j\in J}\sum_{k\in\mathcal{K}^j}\lambda_j Z_{j,k}T^d_{j,k}(s)\right)\cdot\left(\sum_{\tilde{j}\in J}\sum_{\tilde{k}\in\mathcal{K}^{\tilde{j}}}\lambda_{\tilde{j}} Z_{\tilde{j},\tilde{k}}T^d_{\tilde{j},\tilde{k}}(t)\right)\right] \\ &= \sum_{j\in J}\sum_{\tilde{j}\in J}\sum_{k\in\mathcal{K}^j}\sum_{\tilde{k}\in\mathcal{K}^{\tilde{j}}}\lambda_j\lambda_{\tilde{j}}E[Z_{j,k}Z_{\tilde{j},\tilde{k}}]T^d_{j,k}(s)T^d_{\tilde{j},\tilde{k}}(t).\end{aligned}$$

Since the random variables $Z_{j,k}$ are independent with mean 0 and variance 1 we have by corollary 3.2 and Remark 6.6,

$$E[X_sX_t] = \sum_{j\in J}\sum_{k\in\mathcal{K}^j}\lambda_j^2 T^d_{j,k}(s)T^d_{j,k}(t) = \sum_{j\in J}\sum_{k\in\mathcal{K}^j}\int_{\prod_{i=1}^d[0,t_i]} H^d_{j,k}\,dx\int_{\prod_{i=1}^d[0,s_i]} H^d_{j,k}\,dx.$$

By Remark 6.1 (2) and since $\{H^d_{j,k} : j \in J, k \in \mathcal{K}^j\}$ is a complete orthonormal basis for $L^2([0,1]^d)$ we have

$$E[X_sX_t] = \prod_{i=1}^{d}\min\{s_i, t_i\}.$$

For $t \in [0,1]^d$, we have as $E[Z_{j,k}] = 0$ for all j, k

$$E[X_t] = E[\sum_{j\in J}\sum_{k\in\mathcal{K}^j}\lambda_j Z_{j,k}T^d_{j,k}(t)] = \sum_{j\in J}\sum_{k\in\mathcal{K}^j}\lambda_j E[Z_{j,k}]T^d_{j,k}(t) = 0.$$

Hence

$$\mathrm{Cov}[X_t, X_s] = E[X_tX_s] - E[X_t]E[X_s] = \prod_{i=1}^{d}\min(s_i, t_i).$$

(III) Finally we verify that $\{X_t\}_{t\in\mathbb{R}^d}$ is a Gaussian random field. Let $t = \{t^1, t^2, ..., t^m\}, t^i \in [0,1]^d$ be any finite sequence, and consider the vector

$X = (X_{t^1}, ..., X_{t^m})$. We claim that X has the multivariate Gaussian distribution. For each $l = 1, ..., m$, t^l is a vector $(t^l_1, ..., t^l_d)$. Furthermore let $s = (s_1, s_2, ..., s_m) \in \mathbb{R}^d$ be arbitrary. Consider

$$\begin{aligned} E[e^{is^TX}] = E[\exp(i\sum_{l=1}^{m} s_l X_{t^l})] &= E\left[\exp\left(i\sum_{l=1}^{m} s_l\left[\sum_{j\in J}\sum_{k\in\mathcal{K}^j}\lambda_j Z_{j,k}T_{j,k}(t^l)\right]\right)\right] \\ &= E\left[\exp\left(i\sum_{j\in J}\sum_{k\in\mathcal{K}^j}\lambda_j Z_{j,k}\sum_{l=1}^{m} s_l T_{j,k}(t^l)\right)\right] \end{aligned}$$

Since $\{Z_{j,k}\}$ is independent, we have by Theorem 3.1,

$$E[e^{is^TX}] = E[\exp(i\sum_{l=1}^{m} s_l X_{t^l})] = \prod_{j\in J}\prod_{k\in\mathcal{K}^j} E\left[\exp\left(i\lambda_j Z_{j,k}\sum_{l=1}^{m} s_l T_{j,k}(t^l)\right)\right].$$

Since each $Z_{j,k}$ has Gaussian distribution with zero mean and variance 1 we have

$$\begin{aligned} E[e^{is^TX}] &= \prod_{j\in J}\prod_{k\in\mathcal{K}^j}\exp\left(i\sum_{l=1}^{m} s_l T_{j,k}(t^l)\lambda_j E[Z_{j,k}] - \frac{1}{2}\left[\sum_{l=1}^{m} s_l T^d_{j,k}(t_l)\right]^2 \lambda_j^2 \mathrm{Var}(Z_{j,k})\right) \\ &= \exp\left(-\frac{1}{2}\sum_{j\in J}\sum_{k\in\mathcal{K}^j}\lambda_j^2\left[\sum_{l=1}^{m} s_l T^d_{j,k}(t^l)\right]^2\right) \\ &= \exp\left[-\frac{1}{2}\sum_{j\in J}\sum_{k\in\mathcal{K}^j}\lambda_j^2\sum_{l=1}^{m}\sum_{q=1}^{m} s_l s_q T^d_{j,k}(t^l)T^d_{j,k}(t^q)\right] \\ &= \exp\left[-\frac{1}{2}\sum_{l=1}^{m}\sum_{q=1}^{m} s_l s_q\left(\sum_{j\in J}\sum_{k\in\mathcal{K}^j}\lambda_j^2 T^d_{j,k}(t^l)\lambda_j T^d_{j,k}(t^q)\right)\right]. \end{aligned}$$

By the calculation of (II) we have

$$\begin{aligned} E[e^{is^TX}] = \exp\left[-\frac{1}{2}\sum_{l=1}^{m}\sum_{q=1}^{m} s_l s_q \mathrm{Cov}(X_{t^l}, X_{t^q})\right] &= \exp[-\frac{1}{2}s^T\sigma s] \\ &= \exp[is^T E[X] - \frac{1}{2}s^T\sigma s], \end{aligned}$$

where $\sigma = [\sigma_{l,q}]_{m\times m}, \quad \sigma_{l,q} = \mathrm{Cov}[X_{t^l}, X_{t^q}]$.

Hence by Remark D.3 in Appendix D, $\{X_{t^1}, ..., X_{t^m}\}$ has the multivariate Gaussian distribution. This shows that, $\{X_t\}_{t\in[0,1]^d}$ is a Gaussian process. Therefore, $\{X_t\}_{t\in[0,1]^d}$ is a Brownian sheet (see the definition of Brownian sheet in Section 3.5). □

6.3 Compactly Supported Wavelet Representation of Brownian Sheet

Let $\{V_j(\mathbb{R})\}_{j\in\mathbb{Z}}$ be a multiresolution analysis on $L^2(\mathbb{R})$, φ the real valued scaling function of $V_j(\mathbb{R})$ satisfying $\hat{\varphi}(0)=0$, and ψ the associated wavelet. Suppose that suppφ, supp$\psi \subset [0,2N-1]$. Let $j_0\in\mathbb{Z}$ be such that $2^{j_0}\geq 4N-4$. By Theorem 2.14, the following collection

$$\begin{aligned}
&2^{\frac{j_0}{2}}\varphi^{\alpha}_{-2N+2}(2^{j_0}x),...,2^{\frac{j_0}{2}}\varphi^{\alpha}_{-1}(2^{j_0}x),\\
&\varphi_{j_0,k}(x)\big|_{[0,1]}, 0\leq k\leq 2^{j_0}-2N+1,\\
&2^{\frac{j_0}{2}}\varphi^{\beta}_{2^{j_0}-2N+2}(2^{j_0}(1-x)),...,2^{\frac{j_0}{2}}\varphi^{\beta}_{2^{j_0}-1}(2^{j_0}(1-x))\\
&2^{\frac{j}{2}}\psi^{\alpha}_{-N+1}(2^{j}x),...,2^{\frac{j}{2}}\psi^{\alpha}_{-1}(2^{j}x),\\
&\psi_{j,k}\big|_{[0,1]}, 0\leq k\leq 2^{j}-2N+1,\\
&2^{\frac{j}{2}}\psi^{\beta}_{2^{j}-2N+2}(2^{j}(1-x)),...,2^{\frac{j}{2}}\psi^{\beta}_{2^{j}-N}(2^{j}(1-x)), j\geq j_0
\end{aligned}$$

is a complete orthonormal basis of $L^2[0,1]$.

Now let $x=(x_1,...,x_d)\in[0,1]^d, j=(j_1,...,j_d)\in\mathbb{Z}^d, j_i\geq j_0, k=(k_1,...,k_d)\in\mathbb{Z}^d$. Introduce a separate dilation index $\tilde{j}_0$ for the scaling function, $\tilde{j}_0=j_0$, and for each $i\in\{1,...,d\}$ we define $\varphi^{(\tilde{j}_0,k_i)}$ for $-2N+2\leq k_i\leq 2^{\tilde{j}_0}-1$ by

$$\varphi^{(\tilde{j}_0,k_i)}(x_i)=\begin{cases}2^{\frac{\tilde{j}_0}{2}}\varphi^{\alpha}_{k_i}(2^{\tilde{j}_0}x_i), & \text{if } -2N+2\leq k_i\leq -1\\ \varphi_{\tilde{j}_0,k_i}(x_i)\big|_{[0,1]}, & \text{if } 0\leq k_i\leq 2^{\tilde{j}_0}-2N+1\\ 2^{\frac{\tilde{j}_0}{2}}\varphi^{\beta}_{k_i}(2^{\tilde{j}_0}(1-x_i)), & \text{if } 2^{\tilde{j}_0}-2N+2\leq k_i\leq 2^{\tilde{j}_0}-1,\end{cases}$$

and for each $j_i\geq j_0$, define $\psi^{(j_i,k_i)}$ for $-N+1\leq k_i\leq 2^{j_i}-N$ by

$$\psi^{(j_i,k_i)}(x_i)=\begin{cases}2^{\frac{j_i}{2}}\psi^{\alpha}_{k_i}(2^{j}x_i), & \text{if } -N+1\leq k_i\leq -1\\ \psi_{j_i,k_i}(x_i)\big|_{[0,1]}, & \text{if } 0\leq k_i\leq 2^{j_i}-2N+1\\ 2^{\frac{j_i}{2}}\psi^{\beta}_{k_i}(2^{j_i}(1-x_i)), & \text{if } 2^{j_i}-2N+2\leq k_i\leq 2^{j_i}-N,\end{cases}$$

and define the collections

$$\mathcal{C}_i = \{\phi_{j_i,k_i} : \phi_{j_i,k_i} = \varphi^{(\tilde{j}_0,k_i)}, -2N+2 \le k_i \le 2^{\tilde{j}_0} - 1,$$
$$\text{or } \phi_{j_i,k_i} = \psi^{(j_i,k_i)}, -N+1 \le k_i \le 2^{i_j} - N, j_i \ge j_0\big\} \quad \text{for } i = 1, ..., d.$$

We will use the notation $J = \{j = (j_1, ..., j_d) : j_i = \tilde{j}_0, j_0, j_0 + 1, ...\}$

and

$$\mathcal{K}^j = \{k = (k_1, ..., k_d) : -2N+2 \le k_i \le 2^{\tilde{j}_0} - 1 \text{ if } j_i = \tilde{j}_0$$
$$\text{or } -N+1 \le k_i \le 2^{j_i} - N \text{ if } j_i \ne \tilde{j}_0\big\}.$$

For each $j \in J$ and $k \in \mathcal{K}^j$, we define

$$\Phi_{j,k}(x) = \prod_{i=1}^{d} \phi_{j_i,k_i}(x_i) \quad \text{for all } x \in [0,1]^d \text{ and } \phi_{j_i,k_i} \in \mathcal{C}_i. \tag{6.6}$$

By Remark 2.9 we obtain that $\{\Phi_{j,k} : j \in J, k \in \mathcal{K}^j\}$ is a complete orthonormal basis for $L^2([0,1]^d)$.

Lemma 6.3. *For $j = \tilde{j}_0, j_0, j_0+1, ...,$ and $t \in \mathbb{R}$, $\displaystyle\sum_{k=-N+1}^{2^j-N} \left| \int_{-k}^{2^j t - k} \psi(x)\, dx \right| \le M$ for some constant M which does not depend on j.*

Proof. Since ψ has compact support, $\|\psi\|_1 < \infty$ and then we have

$$\begin{aligned}
\sum_{k=-N+1}^{2^j-N} \left| \int_{-k}^{2^j t-k} \psi(x)\, dx \right| &= \sum_{k=-N+1}^{-1} \left| \int_{-k}^{2^j t-k} \psi(x)\, dx \right| + \sum_{k=0}^{2^j-N} \left| \int_{-k}^{2^j t-k} \psi(x)\, dx \right| \\
&\le \sum_{k=-N+1}^{-1} \int_{-k}^{2^j t-k} |\psi(x)|\, dx + \sum_{k=0}^{2^j-N} \left| \int_{-k}^{2^j t-k} \psi(x)\, dx \right| \\
&\le \sum_{k=-N+1}^{-1} \int_{-\infty}^{\infty} |\psi(x)|\, dx + \sum_{k=0}^{2^j-N} \left| \int_{0}^{2^j t-k} \psi(x)\, dx \right| \\
&= (N-1)\|\psi\|_1 + \sum_{k=0}^{2^j-N} \left| \int_{0}^{2^j t-k} \psi(x)\, dx \right|.
\end{aligned}$$

It remains to show that $\displaystyle\sum_{k=0}^{2^j-N} \left| \int_{0}^{2^j t-k} \psi(x)\, dx \right| < C$ for some constant C which does not depend on j. Since $\text{supp}\psi \subset [0, 2N-1]$ and $\hat{\psi}(0) = 0$ (as $\hat{\varphi}(0) = 0$ and

$\hat{\psi}(\xi) = h_1(\frac{\xi}{2})\hat{\varphi}(\frac{\xi}{2})$, details can be found in Walnut (2001) pp.185) then obviously,

$$\int_0^t \psi(x)\,dx = \left(\int_0^t \psi(x)\,dx\right)\chi_{[0,2N-1]}(t).$$

We define $F(t) = \left|\int_0^t \psi(x)\,dx\right|\chi_{[0,2N-1]}(t)$ and $F_{j,k}(t) = F(2^j t - k)$ for all $t \in \mathbb{R}$. We can see that $\text{supp}(F) \subset [0, 2N-1]$, and $\text{supp}F_{j,k} \subset \left[\frac{k}{2^j}, \frac{2N-1+k}{2^j}\right]$. Now consider the following cases of $k \in \mathbb{N} \cup \{0\}$.

If $k = l(2N-1)$ where $l \geq 0$ we have

$$\text{supp}F_{j,k} \subset \left[\frac{l(2N-1)}{2^j}, \frac{l(2N-1)+2N-1}{2^j}\right] = \left[\frac{l(2N-1)}{2^j}, \frac{(l+1)(2N-1)}{2^j}\right].$$

We can see that if $l \neq \tilde{l}$ and $k = l(2N-1)$ and $\tilde{k} = \tilde{l}(2N-1)$ then

$$\text{supp}F_{j,k} \cap \text{supp}F_{j,\tilde{k}} = \phi.$$

If $k = l(2N-1)+1$ where $l \geq 0$ we have

$$\begin{aligned}\text{supp}F_{j,k} &\subset \left[\frac{l(2N-1)+1}{2^j}, \frac{l(2N-1)+1+2N-1}{2^j}\right] \\ &= \left[\frac{l(2N-1)+1}{2^j}, \frac{(l+1)(2N-1)+1}{2^j}\right].\end{aligned}$$

We can see that if $l \neq \tilde{l}$ and $k = l(2N-1)+1$ and $\tilde{k} = \tilde{l}(2N-1)+1$ then

$$\text{supp}F_{j,k} \cap \text{supp}F_{j,\tilde{k}} = \phi.$$

Continuing until $k = l(2N-1)+2N-2$ where $l \geq 0$ we have

$$\begin{aligned}\text{supp}F_{j,k} &\subset \left[\frac{l(2N-1)+2N-2}{2^j}, \frac{l(2N-1)+2N-2+2N-1}{2^j}\right] \\ &= \left[\frac{l(2N-1)+2N-2}{2^j}, \frac{(l+1)(2N-1)+2N-2}{2^j}\right].\end{aligned}$$

We can see that if $l \neq \tilde{l}$ and $k = l(2N-1)$ and $\tilde{k} = \tilde{l}(2N-1)$ then

$$\text{supp}F_{j,k} \cap \text{supp}F_{j,\tilde{k}} = \phi.$$

Hence

$$\sum_{k=0}^{\infty} F_{j,k}(t) = \sum_{l=0}^{\infty}\sum_{n=0}^{2N-2} F_{j,l(2N-1)+n}(t) = \sum_{n=0}^{2N-2}\sum_{l=0}^{\infty} F_{j,l(2N-1)+n}(t)$$
$$\leq \sum_{n=0}^{2N-2} \tilde{C} = \tilde{C}(2N-1)$$

for some constant $\tilde{C}$, as for each t, only one term in $\sum_{l=1}^{\infty} F_{j,l(2N-1)+n}(t)$ is nonzero, and hence $\| \sum_{l=1}^{\infty} F_{j,l(2N-1)+n}(t)| \leq \sup_{l} \|F_{j,l(2N-1)+n}\|_{\infty} \leq \|\psi\|_1$.

Thus,

$$\sum_{k=0}^{2^j-N} \left| \int_0^{2^j t-k} \psi(x)\,dx \right| = \sum_{k=0}^{\infty} F_{j,k}(t) \leq (2N-1)\tilde{C}$$

so that

$$\sum_{k=-N+1}^{2^j-N} \left| \int_{-k}^{2^j t-k} \psi(x)\,dx \right| \leq (N-1)\left\|\psi\right\|_1 + (2N-1)\tilde{C} = M$$

for some constant M. □

By a similar argument, we also obtain that $\sum_{k=-2N+2}^{2^{\tilde{j}_0}-1} \left| \int_{-k}^{2^{\tilde{j}_0} t-k} \varphi(x)\,dx \right| \leq \tilde{M}$ for some constant $\tilde{M}$.

Lemma 6.4. *For $t \in [0,1]$, $\sum_{k=-N+1}^{2^j-N} \left| \int_0^t \phi_{j,k}(x)\,dx \right| \leq C2^{-\frac{j}{2}}$ for some constant C, for all $j = \tilde{j}_0, j_0, j_0+1, ...,$ where $\mathcal{K}^j$ and $\phi_{j,k}$ are as defined at the beginning of this section.*

Proof. Recall to remark 2.8 the construction of the boundary wavelets using the Gram-Schmidt process. Thus, for each $k = -N+1, ..., -1$ we have

$$2^{\frac{j}{2}} \psi_k^{\alpha}(2^j x) = \sum_{l=-N+1}^{-1} C_l^k \psi_{j,l}(x), \tag{6.7}$$

then

$$\sum_{k=-N+1}^{-1} \left| \int_0^t 2^{\frac{j}{2}} \psi_l^{\alpha}(2^j x)\,dx \right| = \sum_{k=-N+1}^{-1} \left| \int_0^t \sum_{l=-N+1}^{-1} C_l^k \psi_{j,l}(x)\,dx \right|$$

$$\leq \sum_{l=-N+1}^{-1} \sum_{k=-N+1}^{-1} \left| \int_0^t C_l^k \psi_{j,l}(x)\, dx \right|$$

$$\leq \sum_{l=-N+1}^{-1} \left| \int_0^t C_l \psi_{j,l}(x)\, dx \right|$$

$$\text{where } C_l = \max\{|C_l^k|\}, \tag{6.8}$$

similarly

$$\sum_{k=2^j-2N+2}^{2^j-N} \left| \int_0^t 2^{\frac{j}{2}} \psi_k^\beta(2^j(1-x))\, dx \right| = \sum_{l=2^j-2N+2}^{2^j-N} \left| \int_0^t \tilde{C}_l \psi_{j,l}(x)\, dx \right|, \tag{6.9}$$

for some scalars $C_l, \tilde{C}_l$.

Now consider, for each $t \in [0,1]$,

$$\sum_{k=-N+1}^{2^j-N} \left| \int_0^t \psi^{(j,k)}(x)\, dx \right| = \sum_{k=-N+1}^{-1} \left| \int_0^t \psi^{(j,k)}(x)\, dx \right| + \sum_{k=0}^{2^j-2N+1} \left| \int_0^t \psi^{(j,k)}(x)\, dx \right|$$
$$+ \sum_{k=2^j-2N+2}^{2^j-N} \left| \int_0^t \psi^{(j,k)}(x)\, dx \right|.$$

By the definition of $\psi^{(j,k)}$,

$$\sum_{k=-N+1}^{2^j-N} \left| \int_0^t \psi^{(j,k)}(x)\, dx \right| = \sum_{k=-N+1}^{-1} \left| \int_0^t 2^{\frac{j}{2}} \psi_k^\alpha(2^j x)\, dx \right| + \sum_{k=0}^{2^j-2N+1} \left| \int_0^t \psi_{j,k}(x)\, dx \right|$$
$$+ \sum_{k=2^j-2N+2}^{2^j-N} \left| \int_0^t 2^{\frac{j}{2}} \psi_k^\beta(2^j(1-x))\, dx \right|.$$

By (6.8) and (6.9),

$$\sum_{k=-N+1}^{2^j-N} \left| \int_0^t \psi^{(j,k)}(x)\, dx \right| \leq \sum_{k=-N+1}^{-1} \left| \int_0^t C_k \psi_{j,k}(x)\, dx \right| + \sum_{k=0}^{2^j-2N+1} \left| \int_0^t \psi_{j,k}(x)\, dx \right|$$
$$+ \sum_{k=2^j-2N+2}^{2^j-N} \left| \int_0^t \tilde{C}_k \psi_{j,k}(x)\, dx \right|$$
$$\leq M \sum_{k=-N+1}^{2^j-N} \left| \int_0^t \psi_{j,k}(x)\, dx \right|$$

$$\text{where } \; M = \max\{|C_k|, 1, |\tilde{C}_k|\}$$

By changing variable $x \mapsto 2^{-j}(x+k)$ we have

$$\sum_{k=-N+1}^{2^j-N}\left|\int_0^t \psi^{(j,k)}(x)\,dx\right| \leq M\sum_{k=-N+1}^{2^j-N}2^{-\frac{j}{2}}\left|\int_{-k}^{2^jt-k}\psi(x)\,dx\right| = M2^{-\frac{j}{2}}\sum_{k=-N+1}^{2^j-N}\left|\int_{-k}^{2^jt-k}\psi(x)\,dx\right|.$$

By Lemma 6.3, we obtain that, for some a constant M',

$$\sum_{k=-N+1}^{2^j-N}\left|\int_0^t \psi^{(j,k)}(x)\,dx\right| \leq MM'2^{-\frac{j}{2}}. \tag{6.10}$$

By the same argument we obtain that,

$$\sum_{k=-2N+2}^{2^{\tilde{j}_0}-1}\left|\int_0^t \varphi^{(\tilde{j}_0,k)}(x)\,dx\right| \leq \tilde{M}2^{-\frac{\tilde{j}_0}{2}}\sum_{k=-2N+2}^{2^{\tilde{j}_0}-1}\left|\int_{-k}^{2^{j_0}t-k}\varphi(x)\,dx\right| \leq \tilde{M}\tilde{M}'2^{-\frac{\tilde{j}_0}{2}}. \tag{6.11}$$

for some constant $\tilde{M}$. Hence,

$$\begin{aligned}\sum_{k\in\mathcal{K}^j}\left|\int_0^t \phi_{j,k}(x)\,dx\right| &= \begin{cases}\displaystyle\sum_{k=-2N+2}^{2^{\tilde{j}_0}-1}\left|\int_0^t \varphi^{(\tilde{j}_0,k)}(x)\,dx\right| & \text{if } j=\tilde{j}_0\\ \displaystyle\sum_{k=-N+1}^{2^j-N}\left|\int_0^t \psi^{j,k}(x)\,dx\right| & \text{if } j\neq\tilde{j}_0\end{cases}\\ &\leq \begin{cases}\tilde{M}\tilde{M}'2^{-\frac{\tilde{j}_0}{2}} & \text{if } j=\tilde{j}_0\\ MM'2^{-\frac{j}{2}} & \text{if } j\neq\tilde{j}_0\end{cases}\\ &\leq \tilde{C}2^{-\frac{j}{2}} \quad \text{for } \tilde{C}=\max\{\tilde{M}\tilde{M}', MM'\}.\end{aligned}$$

□

Remark 6.8. For each $j\in J$ we have

$$\begin{aligned}\sum_{k\in\mathcal{K}^j}\left|\int_{\prod_{i=1}^d[0,t_i]}\Phi_{j,k}(x)\,dx\right| &= \sum_{k\in\mathcal{K}^j}\left|\int_{\prod_{i=1}^d[0,t_i]}\prod_{i=1}^d\phi_{j_i,k_i}(x_i)\,dx_1...dx_d\right|\\ &= \sum_{k\in\mathcal{K}^j}\prod_{i=1}^d\left|\int_0^{t_i}\phi_{j_i,k_i}(x_i)\,dx_i\right| = \prod_{i=1}^d\sum_{k\in\mathcal{K}^j}\left|\int_0^{t_i}\phi_{j_i,k_i}(x_i)\,dx_i\right|\\ &\leq \tilde{C}\prod_{i=1}^d 2^{-\frac{j_i}{2}} \quad \text{by Lemma 6.4.}\end{aligned}$$

Theorem 6.5. *Let* $\{Z_{j,k} : j \in J$ *and* $k \in \mathcal{K}^j\}$ *be a sequence of independent Gaussian variables with mean* 0 *and variance* 1*. Then the series defined by*

$$X_t(\omega) = \sum_{j\in J}\sum_{k\in\mathcal{K}^j} Z_{j,k}(\omega)\int_{\prod_{i=1}^d[0,t_i]} \Phi_{j,k}(x)\,dx, \quad \text{for } t=(t_1,...,t_d)\in[0,1]^d, \tag{6.12}$$

converges uniformly and absolutely on $[0,1]^d$ *with probability one. Moreover,* $\{X_t\}_{t\in[0,1]^d}$ *is a Brownian sheet with zero mean.*

Proof. (i) Verify uniform convergence with probability one.

For each $i = 1,...,d$, we have $\ln(2^{j_i}+|k_i|) < \ln(2^{j_i+1})$ as $k_i < 2^{j_i}$, and hence $\ln(2^{j_i}+|k_i|) < j_i+1$. Then by Lemma G.3 in Appendix G, there exist a random variable C such that

$$|Z_{j,k}| \le C\left(\sum_{i=1}^d \ln(2^{j_i}+|k_i|)\right)^{\frac{1}{2}} \le C\left(\sum_{i=1}^d (j_i+1)\right)^{\frac{1}{2}} \quad \text{a.e. } \omega$$

It is easy to see that, if $j_i \ge j_0, \tilde{j}_0 \ge 1$ for all i then $\left(\sum_{i=1}^d (j_i+1)\right)^{\frac{1}{2}} \le \prod_{i=1}^d \sqrt{j_i+1}$, so that the above inequality become $|Z_{j,k}| \le \prod_{i=1}^d \sqrt{j_i+1}$. Now, we have for each $N\in\mathbb{N}$

$$\begin{aligned}\sum_{\substack{j\in J\\|j|\ge N}}\sum_{k\in\mathcal{K}^j} |Z_{j,k}|\left|\int_{\prod_{i=1}^d[0,t_i]}\Phi_{j,k}(x)\,dx\right| &\le C\sum_{\substack{j\in J\\|j|\ge N}}\sum_{k\in\mathcal{K}^j}\left[\prod_{i=1}^d (j_i+1)^{\frac{1}{2}}\right]\left|\int_{\prod_{i=1}^d[0,t_i]}\Phi_{j,k}(x)\,dx\right| \\ &= C\sum_{\substack{j\in J\\|j|\ge N}}\left[\prod_{i=1}^d (j_i+1)^{\frac{1}{2}}\right]\sum_{k\in\mathcal{K}^j}\left|\int_{\prod_{i=1}^d[0,t_i]}\Phi_{j,k}(x)\,dx\right|.\end{aligned}$$

By Remark 6.8, we obtain that

$$\begin{aligned}\sum_{\substack{j\in J\\|j|\ge N}}\sum_{k\in\mathcal{K}^j} |Z_{j,k}|\left|\int_{\prod_{i=1}^d[0,t_i]}\Phi_{j,k}(x)\,dx\right| &\le C\tilde{C}\sum_{\substack{j\in J\\|j|\ge N}}\prod_{i=1}^d (j_i+1)^{\frac{1}{2}}\prod_{i=1}^d 2^{-\frac{j_i}{2}} \\ &= \tilde{\tilde{C}}\sum_{\substack{j\in J\\|j|\ge N}}\prod_{i=1}^d (j_i+1)^{\frac{1}{2}}2^{-\frac{j_i}{2}} \to 0 \ \text{ as } N\to\infty\end{aligned}$$

arguing as in the proof of Theorem 6.2. Thus $\sum_{\substack{j\in J\\|j|\leq N}}\sum_{k\in\mathcal{K}^j}|Z_{j,k}|\left|\int_{\prod_{i=1}^d[0,t_i]}\Phi_{j,k}(x)\,dx\right|$ is a uniformly Cauchy sequence and thus converges uniformly. It follows that $\sum_{j\in J}\sum_{k\in\mathcal{K}^j}Z_{j,k}\int_{\prod_{i=1}^d[0,t_i]}\Phi_{j,k}(x)\,dx$ converges absolutely and uniformly a.e. ω on $[0,1]^d$.

(ii) Next, we calculate the covariance function. Consider

$$\begin{aligned}E[X_sX_t] &= E\left[\left(\sum_{j\in J}\sum_{k\in\mathcal{K}^j}Z_{j,k}\int_{\prod_{i=1}^d[0,s_i]}\Phi_{j,k}(x)\,dx\right)\left(\sum_{\tilde{j}\in J}\sum_{\tilde{k}\in\mathcal{K}^j}Z_{\tilde{j},\tilde{k}}\int_{\prod_{i=1}^d[0,t_i]}\Phi_{\tilde{j},\tilde{k}}(x)\,dx\right)\right]\\ &= \sum_{j\in J}\sum_{\tilde{j}\in J}\sum_{k\in\mathcal{K}^j}\sum_{\tilde{k}\in\mathcal{K}^j}E\left[Z_{j,k}Z_{\tilde{j},\tilde{k}}\right]\int_{\prod_{i=1}^d[0,s_i]}\Phi_{j,k}(x)\,dx\int_{\prod_{i=1}^d[0,t_i]}\Phi_{\tilde{j},\tilde{k}}(x)\,dx.\end{aligned}$$

Since $\{Z_{j,k}\}$ is independent with mean 0 and variance 1 we have by Corollary 3.2,

$$E[X_sX_t] = \sum_{j\in J}\sum_{k\in\mathcal{K}^j}\int_{\prod_{i=1}^d[0,s_i]}\Phi_{j,k}(x)\,dx\int_{\prod_{i=1}^d[0,t_i]}\Phi_{j,k}(x)\,dx.$$

By Remark 6.1 (2) and as $\{\Phi_{j,k} : j\in J, k\in\mathcal{K}^j\}$ is a complete orthonormal basis for $L^2([0,1]^d)$ we have

$$E[X_s, X_t] = \prod_{i=1}^d \min(s_i, t_i).$$

For $t\in[0,1]^d$, since $E[Z_{j,k}]=0$ for all j,k we obtain that

$$E[X_t] = \sum_{j\in J}\sum_{k\in\mathcal{K}^j}E\left[Z_{j,k}\right]\int_{\prod_{i=1}^d[0,t_i]}\Phi_{j,k}(x)\,dx = 0.$$

Hence

$$\mathrm{Cov}[X_t, X_s] = E[X_tX_s] - E[X_t]E[X_s] = \prod_{i=1}^d \min(s_i, t_i).$$

(iii) Verify that the random field is a Gaussian random field.

Let $t=\{t^1, t^2, ..., t^m\}, t^l\in[0,1]^d$ be any choice of finite sequence with corresponding vector $X=(X_{t^1}, ..., X_{t^m})$. Claim that X has the multivariate Gaussian distribution.

For each $l = 1, ..., m$ we set $t^l = (t^l_1, ..., t^l_d)$ and also $s = (s_1, ..., s_m)$.

Consider

$$\begin{aligned} E[e^{is^TX}] &= E[\exp(i\sum_{l=1}^{m} s_l X_{t^l})] \\ &= E\left[\exp\left(i\sum_{l=1}^{m} s_l \left[\sum_{j\in J}\sum_{k\in\mathcal{K}^j} Z_{j,k}\int_{\prod_{i=1}^{d}[0,t_i^l]} \Phi_{j,k}(x)\,dx\right]\right)\right] \\ &= E\left[\exp\left(i\sum_{j\in J}\sum_{k\in\mathcal{K}^j} Z_{j,k}\sum_{l=1}^{m} s_l\int_{\prod_{i=1}^{d}[0,t_i^l]} \Phi_{j,k}(x)\,dx\right)\right]. \end{aligned}$$

Since $\{Z_{j,k}\}$ is independent , we have by Theorem 3.1

$$E[e^{is^TX}] = \prod_{j\in J}\prod_{k\in\mathcal{K}^j} E\left[\exp\left(iZ_{j,k}\sum_{l=1}^{m} s_l\int_{\prod_{i=1}^{d}[0,t_i^l]} \Phi_{j,k}(x)\,dx\right)\right].$$

Since $\{Z_{j,k}\}$ is a Gaussian random field we have

$$\begin{aligned} E[e^{is^TX}] = \prod_{j\in J}\prod_{k\in\mathcal{K}^j} \exp\Bigg(& i\sum_{l=1}^{m} s_l\int_{\prod_{i=1}^{d}[0,t_i^l]} \Phi_{j,k}(x)\,dx E\left[Z_{j,k}\right] \\ & -\frac{1}{2}\left(\sum_{l=1}^{m} s_l\int_{\prod_{i=1}^{d}[0,t_i^l]} \Phi_{j,k}(x)\,dx\right)^2 \mathrm{Var}[Z_{j,k}]\Bigg). \end{aligned}$$

Since $\{Z_{j,k}\}$ has zero mean and variance 1 we obtain that

$$\begin{aligned} E[e^{is^TX}] &= \prod_{j\in J}\prod_{k\in\mathcal{K}^j} \exp\left(-\frac{1}{2}\left(\sum_{l=1}^{m} s_l\int_{\prod_{i=1}^{d}[0,t_i^l]} \Phi_{j,k}(x)\,dx\right)^2\right) \\ &= \exp\left(-\frac{1}{2}\sum_{j\in J}\sum_{k\in\mathcal{K}^j}\left(\sum_{l=1}^{m} s_l\int_{\prod_{i=1}^{d}[0,t_i^l]} \Phi_{j,k}(x)\,dx\right)^2\right) \\ &= \exp\left(-\frac{1}{2}\sum_{j\in J}\sum_{k\in\mathcal{K}^j}\sum_{l=1}^{m}\sum_{q=1}^{m} s_l s_q\int_{\prod_{i=1}^{d}[0,t_i^l]} \Phi_{j,k}(x)\,dx\int_{\prod_{i=1}^{d}[0,t_i^q]} \Phi_{j,k}(x)\,dx\right) \\ &= \exp\left(-\frac{1}{2}\sum_{l=1}^{m}\sum_{q=1}^{m} s_l s_q\left[\sum_{j\in J}\sum_{k\in\mathcal{K}^j}\int_{\prod_{i=1}^{d}[0,t_i^l]} \Phi_{j,k}(x)\,dx\int_{\prod_{i=1}^{d}[0,t_i^q]} \Phi_{j,k}(x)\,dx\right]\right). \end{aligned}$$

By the calculation of covariance function in step (ii) we get that

$$\begin{aligned} E[e^{is^TX}] &= \exp\left(-\frac{1}{2}\sum_{l=1}^{m}\sum_{q=1}^{m} s_l s_q \mathrm{Cov}[X_{t^l}, X_{t^q}]\right) \\ &= \exp(-\frac{1}{2}s^T\sigma s) \quad \text{where } \sigma = [\sigma_{l,q}]_{m\times m}, \sigma_{l,q} = \mathrm{Cov}[X_{t^l}, X_{t^q}] \\ &= \exp(is^TE[X] - \frac{1}{2}s^T\sigma s) \text{ as } E[X_t] = 0 \text{ for all } t \in [0,1]^d. \end{aligned}$$

Hence by Remark D.3 in Appendix D , $\{X_{t_l},...,X_{t^m}\}$ has the multivariate Gaussian distribution. Thus $\{X_t\}_{t\in[0,1]^d}$ is a Gaussian process. Therefore by (ii) and (iii), $\{X_t\}_{t\in[0,1]^d}$ is a Brownian sheet (see the definition of Brownian sheet in Section 3.5). □

Remark 6.9. In the one dimensional case ($d=1$), we have the collection

$$\mathcal{C}=\{\phi_{j,k}:\phi_{j,k}=\varphi^{(\tilde{j}_0,k)},-2N+2\le k\le 2^{\tilde{j}_0}-1,$$
$$\text{or } \phi_{j,k}=\psi^{(j,k)},-N+1\le k\le 2^j-N,j\ge j_0\Big\}$$

In notation $J=\{\tilde{j}_0,j_0,j_0+1,...\}$ and

$$\mathcal{K}^j=\{k\in\mathbb{Z}:-2N+2\le k\le 2^{\tilde{j}_0}-1 \text{ if } \phi_{j,k}=\varphi^{(\tilde{j}_0,k)}\in\mathcal{C},$$
$$\text{or } -N+1\le k\le 2^j-N \text{ if } \phi_{j,k}=\psi^{(j,k)}\in\mathcal{C}\Big\}$$

Then the collection $\{\phi_{j,k}:j\in J,k\in\mathcal{K}^j\}$ is a complete orthonormal basis for $L^2([0,1])$. Following Theorem 6.5, the series written as

$$X_t(\omega)=\sum_{j\in J}\sum_{k\in\mathcal{K}^j}Z_{j,k}(\omega)\int_0^t\phi_{j,k}(x)\,dx\quad\text{for } t\in[0,1]$$

converges uniformly on $[0,1]$ with probability one. We therefore find that the sample paths of $\{X_t\}_{t\in[0,1]}$ are continuous with probability one. As in step (ii) of the proof we have $E[X_sX_t]=\min(s,t)$, it follows that $\text{Cov}[X_t,X_s]=\min(s,t)$ for all $s,t\in[0,1]$. Hence, by Lemma E.2 in Appendix E, the process $\{X_t\}_{t\in[0,1]}$ is a Brownian motion.

CHAPTER VII

CONCLUSION

In this thesis, we have discussed two main topics; how to obtain the spectral density function of a random field which is the continuous wavelet transform of some random field with arbitrary dilation matrix, and then use this spectral density function to obtain the ergodic properties, and how to construct Brownian motion and a Brownian sheet from the Haar wavelet function and more generally, from arbitrary compactly supported wavelet functions.

In Chapter IV, we discussed the continuous wavelet transform of three types of random fields, a weakly stationary random field, a random field with stationary increments and a random field with weakly stationary increments, via arbitrary dilation matrix. The wavelet transform gives new random fields, which are weakly stationary, as well as jointly weakly stationary for different dilation matrices. Moreover, we calculated the power spectral and cross-power spectral density function of those continuous wavelet transforms. Starting from a weakly stationary random field (Section 4.1), we calculated the cross-power spectral density function of the wavelet transform of such a field, by a formula involving the product of two of Fourier transforms of the mother wavelet function, each dilated by one of the dilation matrices as well as the power spectral function of the original weakly stationary random field, in Theorem 4.1. We then obtained the cross-power spectral density function of the continuous wavelet transform of a stationary increment random field (Section 4.2) by the formula of Theorem 4.2. We gave some examples of a random field with stationary increments in the one dimensional and d-dimensional cases respectively, namely of fractional Brownian motion and the fractional Brownian field; the finite Borel measures on $\mathbb{R}$ and $\mathbb{R}^d$ are found in equation (4.7) and (4.9), respectively. We also determined the cross-power spectral density function for the wavelet transform of Brownian field in Theorem 4.3, and of fractional Brownian motion in the one dimensional case. In the final section of Chapter IV we showed that the cross-power spectral density function of the continuous wavelet transform of a random field with weakly stationary increments involves a formula of products of two Fourier transform of the mother wavelet each dilated by one of the dilation matrix, and a finite Borel measure which derives from the spectral density function of the increments of the original random field.

In Chapter V, we then showed that the continuous wavelet transform of a weakly stationary, strongly stationary increments or weakly stationary increments random field satisfies the ergodic property, that is there exists a random variable with zero mean such that the estimate for the mean of the wavelet transform converges to this random variable (Equation (5.17)) in the mean square sense, and the estimate for the mean of the auto-correlation function of wavelet transform converges to the square mean of this random variable (Equation (5.18)). In addition, for a weakly stationary random field with zero mean, if it is ergodic in mean then its wavelet transform is also ergodic in mean.

In Chapter VI, the discrete wavelet method was used to construct Brownian motion and Brownian sheets. The main contribution of the present work is the construction of a Brownian motion from a wavelet basis. By the tensor product construction, the Haar wavelet of $L^2[0,1]$ basis gave a basis of $L^2[0,1]^d$. We then constructed the Brownian sheet from this Haar wavelet basis of $L^2([0,1]^d)$ in Theorem 6.2. Secondly, recently some mathematicians constructed complete orthonormal bases of $L^2([0,1]^d)$ via a compactly supported wavelet function, as explained in Section 2.6 and also the proofs in Appendix C. We then used this basis to construct a Brownian sheet in Theorem 6.5, which reduces to Brownian motion in the one dimensional case, as shown in Remark 6.9.

REFERENCES

REFERENCES

Averkamp, R. and Houdre, C. (1998). Some distributional properties of the continuous wavelet transform of random processes. **IEEE Transactions on Information Theory**. 44(3): 1111-1124.

Benedetto, J.J. and Frazier, M.W. (1994). **Wavelets: mathematics and applications**. New York: CRC Press.

Bibi, H., Jouini, A. and Kratou, M. (2008). More general constructions of wavelets on the interval. **Communications in Mathematical Analysis**. 4(1): 45-57.

Bogachev, V.I. (2007). **Measure theory**. Berlin Heidelberg: Springer-Verlag.

Boggess, A. and Narcowich, F.J. (2001). **A first couse in waveletes with Fourier analysis**. New Jersey: Prentice-Hall, Inc.

Cambanis, S. and Houdre, C. (1995). On the continuous wavelet transform of second-order random processes. **IEEE Transactions on Information Theory**. 41(3): 628-642.

Childers, D.G. (1998). **Probability and random processes**. New York: The McGraw-Hill Companies, Inc.

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. **Comm. Pure Applied Mathematics** 41: 909-996.

Daubechies, I. (1992). **Ten lectures on wavelets society for industrial and applied mathematics**. Philadelphia: SIAM.

Donals, C.L. (1997). **Measure theory**. Birkhäuser.

Dzhaparidze, K. (2005). Representation of isotropic random fields with homogeneous increments. **Journal of Applied Mathematics and Stochastic Analysis**. 2006(72731): 25-60.

Flandrin, P. (1989). On the spectrum of fractional Brownian motion. **IEEE Transactions on Information Theory**. 35(1): 197-199.

Flandrin, P. (1992). Wavelet analysis and synthesis of fractional Brownian motion. **IEEE Transactions on Information Theory**. 38(2): 910-917.

Folland, G.B. (1999). **Real analysis: modern techniques and their applications**. New York: John Wiley & Sons.

Gasquet, C. and Witomski, P. (1999). **Fourier analysis and applications; filtering, numerical computation, Wavelets**. New York: Springer-Verlag.

Grimmett, G.R. and Stirzaker, D.R. (1998). **Probability and random processes**. New York: Oxford University Press.

Heneghan, C., Lowen, S.B. and Teich, M. (1996). Two dimensional fractional Brownian motion wavelet analysis and synthesis, in Proc. **IEEE Southwest Symp. Image Analysis and Interpretation (San Antonio, TX)**. 213-217.

Hernández (1995). **Lectures on probability and second order random fields**. Singapore: World Scientific Publishing.

Jouini, A., Kratou, M. and Ajmi, N. (2008). General wavelet bases on the cube. **International Journal of Mathematical Analysis**. 2: 647-662.

Kato, T. and Masry, E. (1999). On the spectral density of the wavelet transform of fractional Brownian motion. **Journal Time Series Analysis** 20(5): 559-563.

Klebaner F.C. (1998). **Introduction to stochastic calculus with applications**. London: Imperial College Press.

Ludeman, L.C. (2003). **Random processes; filtering, estimation and detection**. New Jersey: John Wiley & Sons.

Martin, W. and Flandrin, P. (1985). Wigner-Ville spectral analysis of nonstationary processes. **IEEE Trans. Acoust. Speech Signal Process**. 33(6): 1461-1470.

Masry, E. (1993). The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion. **IEEE Transactions on Information Theory**. 39(1): 260-264.

Masry, E. (1995). Random fields and their wavelet transforms and representation: covariance and spectral properties. **IEEE Transactions on Information Theory**. Proceedings, 1995 IEEE International Symposium: 427.

Masry, E. (1998). Covariance and spectral properties of the wavelet transform and discrete wavelet coefficients of second-order random fields. **Signal Processing**. 64: 131-143.

Meyer, Y. (1987). **Wavelets with compact support**. Zygmund Lectures, U.Chicago.

Meyer, Y. (1991). Ondelettes sur l'intervalle. **Revista Mathematica Iberoamericana**. 7(2): 115-134.

Meyer, Y., Sellan, F. and Taqqu, M.S. (1999). Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion. **The Journal of Fourier Analysis and Applications**. 5(5): 465-494.

Pathak, R.S. (2001). **A course in distribution theory and application**. India: Alpha Science.

Paulo, G. and Rudolf, R. (1999). Wavelet analysis of fractional Brownian motion in multi fractal time. **Proceeding of The 17^{th} Colloquium GRETSI**.

Pinsky, M.A. (2009). **Introduction to Fourier analysis and wavelets**. American Mathematical Society Providence, Rhode Island.

Ramanathan, J. and Zeitouni, O. (1991). On the wavelet transform of fractional Brownian motion. **IEEE Transactions on Information Theory**. 37(4): 1156-1158.

Reed, I.S., Lee, P.C. and Truong, T.K. (1995). Spectral representation of fractional Brownian motion in n dimensions and its properties. **IEEE Transactions on Information Theory**. 41(5): 1439-1451.

Steele, J.M. (2000). **Stochastic calculus and financial applications**. New York: Springer-Verlag.

Strichartz, R.S. (2003). **A guide to distribution theory and fourier transforms**. Singapore: World Scientific Publishing.

Tewfik, A.H. and Kim, M. (1992). Correlation structure of the discrete wavelet coefficients of fractional Brownian motion. **IEEE Transactions on Information Theory**. 38(2): 904-909.

Viniotis, Y. (1998). **Probability and random processes for electrical engineering**. New York: The McGraw-Hill Companies, Inc.

Wade, W.R. (1999). **An introduction to analysis**. New Jersey: Prentice-Hall, Inc.

Walnut, D.F. (2001). **An introduction to wavelet analysis**. Boston: Birkhauser.

Yaglom, A.M. (1962), **An introduction to the theory of stationary random functions**, USA: Dover Publications, Inc.

Yaglom, A.M. (1987). **Correlation theory of stationary and related random functions I; basic results**. New York: Springer.

APPENDICES

APPENDIX A

FOURIER TRANSFORM OF SOME FUNCTIONS

Example A.1. If $f(x) = e^{-\alpha|x|^2}$, $x \in \mathbb{R}^d$ where $\alpha > 0$ then $\hat{f}(\xi) = \left(\frac{\pi}{\alpha}\right)^{\frac{d}{2}} e^{-\frac{|\xi|^2}{4\alpha}}$.

Proof. Since $f \in L^1(\mathbb{R}^d)$ we consider the L^1 Fourier transform.

First let $d = 1$. Then, $\frac{d}{dx}(e^{-\alpha x^2}) = -2\alpha x e^{-\alpha x^2}$. Theorem 2.3 (1) says that, $(\hat{f})'(\xi) = i\mathcal{F}(xf)(\xi)$ and $\mathcal{F}(f')(\xi) = -i\xi\hat{f}(\xi)$, hence we have

$$\begin{aligned}(\hat{f})'(\xi) &= i\mathcal{F}(xe^{-\alpha x^2})(\xi) = -\frac{i}{2\alpha}\mathcal{F}\left(-2\alpha x e^{-\alpha x^2}\right)(\xi)\\ &= -\frac{i}{2\alpha}\mathcal{F}\left(f'\right)(\xi) = -\frac{i}{2\alpha}(-i\xi)(\hat{f})(\xi) = -\frac{\xi}{2\alpha}(\hat{f})(\xi)\end{aligned}$$

so that we obtain the linear differential equation $(\hat{f})'(\xi) + \frac{\xi}{2\alpha}\hat{f}(\xi) = 0$. It follows that

$$\frac{d}{d\xi}\left(e^{\frac{\xi^2}{4\alpha}}\hat{f}(\xi)\right) = e^{\frac{\xi^2}{4\alpha}}(\hat{f})'(\xi) + \hat{f}(\xi)e^{\frac{\xi^2}{4\alpha}}(\frac{2\xi}{4\alpha}) = e^{\frac{\xi^2}{4\alpha}}\left[(\hat{f})'(\xi) + \frac{\xi}{2\alpha}\hat{f}(\xi)\right] = 0.$$

Hence $e^{\frac{\xi^2}{4\alpha}}\hat{f}(\xi)$ is a constant (does not depend on the variable ξ). To obtain the value of this constant, let $\xi = 0$. We get $\hat{f}(0) = \int f(x)\,dx = \int e^{-\alpha x^2}\,dx = \sqrt{\frac{\pi}{\alpha}}$. Thus $e^{\frac{\xi^2}{4\alpha}}\hat{f}(\xi) = \sqrt{\frac{\pi}{\alpha}}$ and then $\hat{f}(\xi) = \sqrt{\frac{\pi}{\alpha}}e^{-\frac{\xi^2}{4\alpha}}$.

The d-dimensional case follows by Fubini's Theorem, since $|x|^2 = \sum_{j=1}^{d} x_j^2$, then

$$\begin{aligned}\hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-\alpha|x|^2}e^{-i\xi\cdot x}\,dx &= \prod_{j=1}^{d}\int_{\mathbb{R}} e^{-\alpha x_j^2}e^{-i\xi_j\cdot x_j}\,dx_j\\ &= \left(\frac{\pi}{\alpha}\right)^{\frac{d}{2}}\prod_{j=1}^{d} e^{-\frac{\xi_i^2}{4\alpha}}\\ &= \left(\frac{\pi}{\alpha}\right)^{\frac{d}{2}} e^{-\frac{|\xi|^2}{4\alpha}}.\end{aligned}$$

□

We next apply the result of Example A.1.

Example A.2. The Gaussian distribution with zero mean and variance parameter $\sigma > 0$ is the Borel measure on $\mathbb{R}^d$ with density $e^{-\frac{|x|^2}{2\sigma^2}}$. That is, it is the probability measure m defined by $m(A) = \displaystyle\int_A e^{-\frac{|x|^2}{2\sigma^2}}\, dx$ for Borel subsets A of $\mathbb{R}^d$ (see also more details in Definition 3.5 in Section 3.1). The Fourier transform $\hat{m}$ of this measure is given in Definition 2.8 of Section 2.4, and by Example A.1 we obtain that

$$\hat{m}(\xi) = \int_{\mathbb{R}^d} e^{-i\xi\cdot x} e^{-\frac{|x|^2}{2\sigma^2}}\, dx = (2\pi\sigma^2)^{\frac{d}{2}} e^{-\frac{|\xi|^2\sigma^2}{2}}.$$

We will use this formula in Appendix D to obtain the characteristic function of a Gaussian random field.

APPENDIX B

THE PROOF OF BOCHNER'S THEOREM

We begin by establishing some general properties of positive definite functions.

Proposition B.1. *Let φ be a positive definite function. Then*

1) $\varphi(0) \geq 0$

2) $\varphi(-x) = \overline{\varphi(x)}$ *and* $|\varphi(x)| \leq \varphi(0)$ *for all* $x \in \mathbb{R}$

3) the sums, products and limits of positive definite functions are positive definite. In addition, $\exp \varphi$ *is a positive definite function.*

Proof. Suppose that φ is a positive definite function. Then for all finite sequences of complex number $\{c_i\}_{i=1}^N$ and finite sequences $\{\xi_i\}_{i=1}^N \subset \mathbb{R}^d$ we have $\sum_{i,j=1}^{N} c_i \overline{c_j} \varphi(\xi_i - \xi_j) \geq 0$.

Then 1) setting $N = 1$ and $c_1 = 1$ we get $\varphi(0) \geq 0$,

2) Setting $N = 2$, $\xi_1 = x$ and $\xi_2 = 0$ we get

$$|c_1|^2 \varphi(0) + |c_2|^2 \varphi(0) + c_1 \overline{c_2} \varphi(x) + c_2 \overline{c_1} \varphi(-x) \geq 0.$$

Setting $c_1 = i, c_2 = -1$ we have $2\varphi(0) \geq i\varphi(x) - i\varphi(-x)$, then $0 = \mathrm{Im}(i\varphi(x)) - \mathrm{Im}(i\varphi(-x)) = \mathrm{Re}(\varphi(x)) - \mathrm{Re}(\varphi(-x))$, that is

$$\mathrm{Re}(\varphi(-x)) = \mathrm{Re}(\varphi(x)).$$

Setting $c_1 = 1, c_2 = -1$ we have $2\varphi(0) \geq \varphi(x) + \varphi(-x)$. Hence $0 = \mathrm{Im}(\varphi(x)) + \mathrm{Im}(\varphi(-x))$ that is

$$\mathrm{Im}(\varphi(-x)) = -\mathrm{Im}(\varphi(x)).$$

Hence $\varphi(-x) = \mathrm{Re}(\varphi(-x)) + i\mathrm{Im}(\varphi(-x)) = \overline{\mathrm{Re}(\varphi(x)) + i\mathrm{Im}(\varphi(x))} = \overline{\varphi(x)}$.

If $\varphi(x) \neq 0$, setting $c_1 = -\dfrac{|\varphi(x)|}{\varphi(x)}$, $c_2 = 1$, then $|c_1| = 1$ and hence

$$2\varphi(0) \geq \frac{|\varphi(x)|}{\varphi(x)}\varphi(x) + \frac{|\varphi(x)|}{\overline{\varphi(x)}}\varphi(-x) = |\varphi(x)| + \frac{|\varphi(x)|}{\overline{\varphi(x)}}\overline{\varphi(x)} = 2|\varphi(x)|.$$

Thus $|\varphi(x)| \leq \varphi(0)$.

3) Note that φ positive definite is equivalent to the matrix $[\varphi(\xi_i - \xi_j)]_{i,j}$ being positive semidefinite for all choice of $\xi_1, ..., \xi_N \in \mathbb{R}^d$ and $N \in \mathbb{N}$. From here it is not difficult to see that the product of two positive definite functions is again a positive definite function. The sums and pointwise limits of positive definite functions are obviously positive definite from the definition. Since the exponential function of a given function is obtained by sums, products, and limits, the exponential of a positive definite function is again positive definite. □

Remark B.1. The following two properties will be used in the next proposition.

1) For fixed $x \in \mathbb{R}^d$, the function $z \longmapsto \varphi(z)e^{iz\cdot x}$ is a positive definite function, whenever φ is a positive definite function.

Indeed, let $\{c_k\}_{k=1}^N \subset \mathbb{C}$ and $\{\xi_k\}_{k=1}^N \subset \mathbb{R}^d$. We have

$$\sum_{k,j=1}^N c_k\overline{c_j}e^{i(\xi_k-\xi_j)\cdot x} = \sum_{k,j=1}^N c_k\overline{c_j}e^{i\xi_k\cdot x}e^{-i\xi_j\cdot x} = \sum_{k,j=1}^N c_k e^{i\xi_k\cdot x}\overline{c_j e^{i\xi_j\cdot x}} = \Big|\sum_{k=1}^N c_k e^{i\xi_k\cdot x}\Big|^2 \geq 0.$$

Thus $z \longmapsto e^{iz\cdot x}$ is a positive definite function and hence $z \longmapsto \varphi(z)e^{iz\cdot x}$ is a positive definite function, as it is the product of two positive definite functions.

2) If f is an integrable even function, then $f*f$ is a positive definite function.

Indeed, let $\{c_i\}_{i=1}^N \subset \mathbb{C}$ and $\{\xi_i\}_{i=1}^N \subset \mathbb{R}^d$.

We have

$$\sum_{i,j=1}^N c_i\overline{c_j}(f*f)(\xi_i - \xi_j) = \sum_{i,j=1}^N c_i\overline{c_j}\int_{\mathbb{R}^d} f(y)f(\xi_i - \xi_j - y)\,dy.$$

Changing the variable $y \mapsto \xi_i - y$ we have

$$\sum_{i,j=1}^N c_i\overline{c_j}(f*f)(\xi_i - \xi_j) = \sum_{i,j=1}^N c_i\overline{c_j}\int_{\mathbb{R}^d} f(\xi_i - y)f(y - \xi_j)\,dy.$$

Since f is an even function, $f(y-\xi_j) = f(\xi_j - y)$ so that

$$\begin{aligned}\sum_{i,j=1}^{N} c_i\overline{c_j}(f*f)(\xi_i-\xi_j) &= \sum_{i,j=1}^{N} c_i\overline{c_j}\int_{\mathbb{R}^d} f(\xi_i-y)f(\xi_j-y)\,dy \\ &= \int_{\mathbb{R}^d}\sum_{i,j=1}^{N} c_i\overline{c_j}f(\xi_i-y)f(\xi_j-y)\,dy \\ &= \int_{\mathbb{R}^d}\left|\sum_{i=1}^{N} c_i f(\xi_i-y)\right|^2 dy \geq 0.\end{aligned}$$

Proposition B.2. *The Fourier transform of a finite Borel measure m on $\mathbb{R}^d$ has the following properties.*

1) $\hat{m}$ is a continuous function with $\hat{m}(0) = m(\mathbb{R}^d)$.

2) $\hat{m}$ is a positive definite function.

3) If m_1, m_2 are two measures, then the Fourier transform of their convolution is the product of their Fourier transforms. Recall that the convolution is defined by

$$(m_1 * m_2)(B) = \int_{\{(x,y):x+y\in B\}} 1\,dm_1(x)dm_2(y),$$

and hence for any bounded continuous function g

$$\int_{\mathbb{R}^d} g(z)\,d(m_1*m_2)(z) = \int_{\mathbb{R}^{2d}} g(x+y)\,dm_1(x)dm_2(y).$$

Proof. 1) Suppose that $\xi_n \to \xi$ as $n \to \infty$. Now for each $n \in \mathbb{N}$, $|e^{i\xi_n\cdot x}| = 1$, $\int_{\mathbb{R}^d} dm(x) = m(\mathbb{R}^d) < \infty$ and $e^{i\xi_n\cdot x} \longrightarrow e^{i\xi\cdot x}$. Then by the Dominated Convergence Theorem,

$$\hat{m}(\xi_n) = \int_{\mathbb{R}^d} e^{-i\xi_n\cdot x}\,dm(x) \longrightarrow \int_{\mathbb{R}^d} e^{-i\xi\cdot x}\,dm(x) = \hat{m}(\xi).$$

This shows that $\hat{m}$ is a continuous function.

Obviously, $\hat{m}(0) = \int_{\mathbb{R}^d} 1\,dm(x) = m(\mathbb{R}^d)$.

2) Let $\{c_j\}_{j=1}^{n}$ be a finite set of complex numbers and $\{\xi_j\}_{j=1}^{n}$ any corresponding finite subset of $\mathbb{R}^d$.

Then

$$\sum_{j,k=1}^{n} c_j\bar{c}_k\hat{m}(\xi_j-\xi_k) = \sum_{j,k=1}^{n} c_j\bar{c}_k\int_{\mathbb{R}^d} e^{-i\xi_j\cdot x}e^{i\xi_k\cdot x}\,dm(x) = \int_{\mathbb{R}^d}\sum_{j,k=1}^{n} c_je^{-i\xi_j\cdot x}\overline{c_ke^{-i\xi_k\cdot x}}\,dm(x)$$

$$= \int_{\mathbb{R}^d}\left|\sum_{j=1}^{n} c_je^{-i\xi_j\cdot x}\right|^2 dm(x) \quad \geq \quad 0$$

Hence $\hat{m}$ is a positive definite function.

3) Consider

$$\hat{m}_1(\xi)\hat{m}_2(\xi) = \int_{\mathbb{R}^{2d}} e^{-i\xi\cdot(x+y)}\,dm_1(x)\,dm_2(y) = \int_{\mathbb{R}^d} e^{-i\xi\cdot(z)}\,d(m_1{*}m_2)(z) = \widehat{m_1{*}m_2}(\xi).$$

□

Lemma B.3. *If φ is a measurable positive definite function on $\mathbb{R}^d$, then for every nonnegative Lebesgue integrable function f, one has*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\varphi(x-y)f(x)f(y)\,dx\,dy \geq 0. \tag{B.1}$$

If the function f is even, then

$$\int_{\mathbb{R}^d}\varphi(x)\,(f*f)(x)\,dx \geq 0. \tag{B.2}$$

In particular, for all $\alpha > 0$ we have

$$\int_{\mathbb{R}^d}\varphi(x)e^{-\alpha|x|^2}\,dx \geq 0. \tag{B.3}$$

Proof. Let $k \geq 2$ be arbitrary.

As φ is a positive definite function, $k\varphi(0) + \sum_{i\neq j}\varphi(y_i - y_j) \geq 0$ for any vector $y = (y_1, ..., y_k) \in (\mathbb{R}^d)^k$. Clearly, the function $\tilde{\varphi}(y) = k\varphi(0) + \sum_{i\neq j}\varphi(y_i - y_j)$ is Lebesge measurable on $(\mathbb{R}^d)^k$.

By using the boundedness and measurability of φ we can integrate this inequality with respect to the finite measure $f(y_1)f(y_2)...f(y_k)\,dy_1\,dy_2\,...\,dy_k$, and as $I(f) =$

$\int_{\mathbb{R}^d} f(x)\,dx \geq 0$ we have

$$\int_{\mathbb{R}^{kd}} k\varphi(0) \prod_{l=1}^{k} f(y_l)\,dy + \sum_{i\neq j} \int_{\mathbb{R}^{kd}} \varphi(y_i - y_j) \prod_{l=1}^{k} f(y_l)\,dy \geq 0$$

$$k\varphi(0) \prod_{l=1}^{k} \int_{\mathbb{R}^d} f(y_l)\,dy_l + \sum_{i\neq j} \left[\prod_{\substack{l=1 \\ l\neq i,j}}^{k} \int_{\mathbb{R}^d} f(y_l)\,dy_l \right] \int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(y_i - y_j) f(y_i) f(y_j)\,dy_i\,dy_j \geq 0$$

$$k\varphi(0) I(f)^k + I(f)^{k-2} \sum_{i\neq j} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(y_i - y_j) f(y_i) f(y_j)\,dy_i\,dy_j \geq 0$$

$$k\varphi(0) I(f)^k + I(f)^{k-2} (k)(k-1) \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x-y) f(x) f(y)\,dx\,dy \geq 0.$$

If $I(f) = 0$, the assertion is trivial. Thus, we may assume that $I(f) \neq 0$. Dividing by $k(k-1)I(f)^k$ we get

$$\frac{\varphi(0)}{k-1} + I(f)^{-2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x-y) f(x) f(y)\,dx\,dy \geq 0.$$

Letting $k \to \infty$,

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x-y) f(x) f(y)\,dx\,dy \geq 0. \tag{B.4}$$

Next, assume that f is even, that is $f(-y) = f(y)$ for all $y \in \mathbb{R}^d$.

Then

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x-y) f(x) f(y)\,dx\,dy = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x-y) f(x)\,dx\, f(y)\,dy.$$

Changing a variable, $x \mapsto x + y$, and applying Fubini's Theorem we have

$$\begin{aligned}\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x-y) f(x) f(y)\,dx\,dy &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x) f(x+y)\,dx\, f(y)\,dy \\ &= \int_{\mathbb{R}^d} \varphi(x) \int_{\mathbb{R}^d} f(x+y) f(y)\,dy\,dx.\end{aligned}$$

Again changing a variable, $y \mapsto -y$, and as f is even we have

$$\begin{aligned}\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(x-y) f(x) f(y)\,dx\,dy &= \int_{\mathbb{R}^d} \varphi(x) \int_{\mathbb{R}^d} f(x-y) f(-y)\,dy\,dx \\ &= \int_{\mathbb{R}^d} \varphi(x) \int_{\mathbb{R}^d} f(x-y) f(y)\,dy\,dx \\ &= \int_{\mathbb{R}^d} \varphi(x) (f * f)(x)\,dx.\end{aligned} \tag{B.5}$$

Thus, by equation (B.4) and (B.5),

$$\int_{\mathbb{R}^d} \varphi(x)(f*f)(x)\,dx \geq 0. \tag{B.6}$$

To prove the last assertion, we express $e^{-\alpha|x|^2}$ as a convolution. Set $f(x) = ce^{-2\alpha|x|^2}$ for $c, \alpha > 0$. This function is Lebesgue integrable, nonnegative and even on $\mathbb{R}^d$, and $\hat{f}(\xi) = c\left(\frac{\pi}{2\alpha}\right)^{\frac{d}{2}} e^{-\frac{|\xi|^2}{8\alpha}}$. Thus $\mathcal{F}(f*f)(\xi) = \left(\hat{f}(\xi)\right)^2 = c^2\left(\frac{\pi}{2\alpha}\right)^d e^{-\frac{|\xi|^2}{4\alpha}} = c^2\left(\frac{\pi}{2\alpha}\right)^{\frac{d}{2}}\left[\left(\frac{\pi}{2\alpha}\right)^{\frac{d}{2}} e^{-\frac{|\xi|^2}{4\alpha}}\right]$. Then, by Example A.1 in Appendix A we have $(f*f)(x) = c^2\left(\frac{\pi}{2\alpha}\right)^{\frac{d}{2}} e^{-\alpha|x|^2}$. If we let $c = \left(\frac{\pi}{2\alpha}\right)^{-\frac{d}{4}}$, then $f(x) = \left(\frac{\pi}{2\alpha}\right)^{-\frac{d}{4}} e^{-2\alpha|x|^2}$ with $c, \alpha > 0$ and hence $(f*f)(x) = e^{-\alpha|x|^2}$. Substituting into equation (B.6) we obtain

$$\int_{\mathbb{R}^d} \varphi(x)e^{-\alpha|x|^2}\,dx \quad\geq\quad 0.$$

□

Proof of Bochner's Theorem

Proof. First, suppose that φ is an integrable positive definite function on $\mathbb{R}^d$. Let $f = \hat{\varphi}$. Then f is bounded and continuous. We claim that $f \geq 0$. Let us consider the function

$$P_t(x) = (2\pi t)^{-\frac{d}{2}} e^{-\frac{|x|^2}{2t}} \quad \text{for} \quad t > 0.$$

As f is bounded and P_t integrable, $P_t * f$ exists. Then

$$\begin{aligned}
(P_t * f)(x) &= \int_{\mathbb{R}^d} f(y)P_t(x-y)\,dy = (2\pi t)^{-\frac{d}{2}} \int_{\mathbb{R}^d} f(y)e^{-\frac{|x-y|^2}{2t}}\,dy \\
&= (2\pi t)^{-\frac{d}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \varphi(z)e^{-iz\cdot y}\,dz\, e^{-\frac{|x-y|^2}{2t}}\,dy \\
&= (2\pi t)^{-\frac{d}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \varphi(z)e^{-iz\cdot y}\, e^{-\frac{|x-y|^2}{2t}}\,dy\,dz,
\end{aligned}$$

where Fubini's theorem applies by integrability of φ. Changing the variable $y \mapsto$

$x-y$ we have

$$(P_t * f)(x) = (2\pi t)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(z) e^{-iz\cdot(x-y)}\, e^{-\frac{|y|^2}{2t}}\, dy\, dz$$
$$= (2\pi t)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \varphi(z) e^{-iz\cdot x} \left(\int_{\mathbb{R}^d} e^{iz\cdot y}\, e^{-\frac{|y|^2}{2t}}\, dy \right) dz.$$

By Example A.1 we have

$$(P_t * f)(x) = (2\pi t)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \varphi(z) e^{-iz\cdot x}\, (2t\pi)^{\frac{d}{2}}\, e^{-\frac{t|z|^2}{2}}\, dz.$$

By Lemma B.3 and Remark B.1 it turns out that

$$(P_t * f)(x) = \int_{\mathbb{R}^d} \varphi(z) e^{-iz\cdot x} e^{-\frac{t|z|^2}{2}}\, dz \geq 0.$$

As $|\varphi(z) e^{iz\cdot x} e^{-\frac{t|z|^2}{2}}| \leq |\varphi(z)|$, and φ is integrable, we have by the Dominated Convergence Theorem

$$\lim_{k\to\infty} P_{\frac{1}{k}} * f(x) = \int \varphi(z) e^{-iz\cdot x} \lim_{k\to\infty} \left(e^{-\frac{|z|^2}{2k}} \right) dz = \int \varphi(z) e^{-iz\cdot x}\, dz = f(x).$$

Hence $f \geq 0$ as $P_{\frac{1}{k}} * f(x) \geq 0$ for all $k > 0$, and the claim is proved. Next we show that f is integrable. For each $k > 0$, we have by Fubini's Theorem

$$\int_{\mathbb{R}^d} f(x) e^{-\frac{|x|^2}{2k}}\, dx = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(z) e^{-ix\cdot z}\, dz\, e^{-\frac{|x|^2}{2k}}\, dx \quad \text{as } f = \hat{\varphi}$$
$$= \int_{\mathbb{R}^d} \varphi(z) \left[\int_{\mathbb{R}^d} e^{-ix\cdot z}\, e^{-\frac{|x|^2}{2k}}\, dx \right] dz$$
$$= \int_{\mathbb{R}^d} \varphi(z) \left[(2k\pi)^{\frac{d}{2}}\, e^{-\frac{k|z|^2}{2}} \right] dz.$$

Since φ is positive definite and applying Proposition B.1 (1) we have

$$\int_{\mathbb{R}^d} f(x) e^{-\frac{|x|^2}{2k}}\, dx \leq \varphi(0)\, (2k\pi)^{\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{k|z|^2}{2}}\, dz.$$

Since $\int_{\mathbb{R}^d} e^{-\frac{k|z|^2}{2}}\, dz = \left(\frac{2\pi}{k} \right)^{\frac{d}{2}}$ it follows that

$$\int_{\mathbb{R}^d} f(x) e^{-\frac{|x|^2}{2k}}\, dx \leq \varphi(0)\, (2\pi)^d.$$

Since $e^{-\frac{|x|^2}{2k}} \leq 1$ and $\lim\limits_{k\to\infty} e^{-\frac{|x|^2}{2k}} = 1$ we have by Fatou's Lemma

$$\int_{\mathbb{R}^d} f(x)\,dx \leq \overline{\lim_{k\to\infty}} \int_{\mathbb{R}^d} f(x)e^{-\frac{|x|^2}{2k}}\,dx \leq \lim_{k\to\infty} \varphi(0)(2\pi)^d < \infty.$$

That is f is integrable. Since $f = \hat{\varphi}$ and f, φ are integrable functions, we have $\check{f}(x) = \varphi(x)$ a.e. x. That is

$$\varphi(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\xi\cdot x} f(\xi)\,d\xi \quad \text{a.e.}$$

which shows that the assertion holds, with $d\mu(\xi) = \frac{1}{(2\pi)^d} f(\xi)d\xi$. Note that the measure μ is finite as f is an integrable function.

In the general case, suppose that φ is a Lebesgue measurable positive definite function on $\mathbb{R}^d$. Let $\epsilon > 0$, and consider the function $x \longmapsto \varphi(x)e^{-\epsilon|x|^2}$. We have

$$\int_{\mathbb{R}^d} |\varphi(x)e^{-\epsilon|x|^2}|\,dx \leq \varphi(0) \int_{\mathbb{R}^d} e^{-\epsilon|x|^2}\,dx < \infty$$

that is, $x \longmapsto \varphi(x)e^{-\epsilon|x|^2}$ is an integrable function. As shown in the proof of Lemma B.3, $e^{-\epsilon|x|^2}$ is a positive definite function. Since the product of positive definite functions is again positive definite, then $\varphi(x)e^{-\epsilon|x|^2}$ is an integrable positive definite function, hence as shown above, coincides almost every where with a continuous function. Hence the function φ has a continuous modification ψ. We show that ψ is also a positive definite function. Indeed, since φ is bounded and P_t integrable, $\varphi * P_t$ exists. By continuity (see Theorem 8.4 in Folland (1999)) one has $\psi(x) = \lim\limits_{t\to 0} \psi * P_t(x)$ for each x. As $\psi * P_t(x) = \varphi * P_t(x)$ for all x and $t > 0$, in view of proposition B.1 (3), it remains to note that $\varphi * P_t$ is a positive definite function. Indeed, applying the Dominated Convergence Theorem we have $\varphi * P_t(x) = \lim\limits_{\epsilon\to 0} \varphi_\epsilon * P_t(x)$, where $\varphi_\epsilon(x) = \varphi(x)e^{-\epsilon|x|^2}$. We already know by the first part that the integrable positive definite function φ_ϵ coincides almost everywhere with the Fourier transform of some nonnegative integrable function g_ϵ. Hence

$\varphi_\epsilon * P_t$ is the Fourier transform of the nonnegative function $g_\epsilon \hat{P}_t$, i.e., is positive definite. Therefore, ψ is a continuous positive definite function, almost everywhere equal to φ. The remaining part of this proof requires tools of functional analysis and can be found in Pinsky (2009). □

APPENDIX C

THE PROOF OF SUBSECTION 2.6.3

The proof of Lemma 2.8

Proof. Since $\text{supp}(\varphi) \subset [0, 2N-1]$, the support of $\varphi(x-k)$ is contained in $[k, k+2N-1]$. Now, for $k \leq -2N+1$, we have $[k, k+2N-1] \subset (-\infty, 0]$, and we get

$$c_k = \int_{-\infty}^{\infty} f(x)\overline{\varphi(x-k)}\, dx = \int_{k}^{k+2N-1} f(x)\overline{\varphi(x-k)}\, dx = 0$$

as $f(x) = 0$ for $x \in [k, k+2N-1] \subset (-\infty, 0]$.

Let p be the smallest integer k such that $c_k \neq 0$ (p exists as $c_k = 0$ for $k \leq -2N+1$). Suppose to contrary that $p < 0$. Then $p+1 \leq 0$, and it follows that $f(x) = 0$ for all $x \in [p, p+1] \subset (-\infty, 0]$. Observe however that $f(x) = c_p\varphi(x-p)$ for $x \in [p, p+1]$. Then $\varphi(x-p) = 0$ for all $x \in [p, p+1]$, that is, $\varphi(x) = 0$ for all $x \in [0, 1]$, which contradicts the fact that $\text{supp}(\varphi) \cap [0, 1] \neq \emptyset$. Hence the smallest integer p such that $c_p \neq 0$ is greater than or equal to 0. Therefore, $c_k = 0$ for $k \leq -1$. □

The proof of Theorem 2.9

Proof. Keep the notation $S_l(j), l = 1, 2, 3$ from the paragraph below Lemma 2.8. For each j, let us set

$$X_j^{(1)} = \text{span}\{\varphi_{j,k} : k \in S_1(j)\}$$

$$X_j^{(2)} = \text{span}\{\varphi_{j,k} : k \in S_2(j)\}$$

$$X_j^{(3)} = \text{span}\{\varphi_{j,k} : k \in S_3(j)\}$$

Furthermore, let us set

$$Y_j^{(1)} = \text{span}\{\varphi_{j,k}\big|_{[0,1]} : k \in S_1(j)\}$$
$$Y_j^{(2)} = \text{span}\{\varphi_{j,k}\big|_{[0,1]} : k \in S_2(j)\}$$
$$Y_j^{(3)} = \text{span}\{\varphi_{j,k}\big|_{[0,1]} : k \in S_3(j)\}$$

These are all finite dimensional spaces. Our first goal is to show that $P : X_j^{(1)} \longrightarrow Y_j^{(1)}$, $P : X_j^{(2)} \longrightarrow Y_j^{(2)}$, $P : X_j^{(3)} \longrightarrow Y_j^{(3)}$ are one-to-one. P being linear and surjective, it will follow that these two maps are linear isomorphisms. Clearly $P : X_j^{(2)} \longrightarrow Y_j^{(2)}$ is one-to-one, as $\text{supp}(\varphi_{j,k}) \subset [0,1]$ for all $k \in S_2(j)$.

Next, let $f_1(x) = \sum_{k\in S_1(j)} c_k\varphi_{j,k}(x) \in X_j^{(1)}$ and suppose that $Pf_1 = 0$. That is,

$$f_1(x) = \sum_{k=-2N+2}^{-1} c_k\varphi_{j,k}(x) = 0 \text{ for all } x \in [0,1]. \tag{C.1}$$

As $\text{supp}(\varphi_{j,k}) \subset [2^{-j}k, 2^{-j}(2N+k-1)]$ and $j \geq j_0$, it follows that (C.1) holds for all $x \in [0,\infty)$. Equivalently,

$$f_1(2^{-j}x) = \sum_{k=-2N+2}^{-1} 2^{\frac{j}{2}} c_k\varphi(x-k) = 0 \quad \text{for all } x \geq 0.$$

Let us first show that $c_{-1} = 0$.

If $-2N+2 \leq k \leq -2$, then $\text{supp}(\varphi(x-k)) \subset (-\infty, 2N-3]$ while for $k_0 = -1$, $\text{supp}(\varphi(x-k_0)) \subset [-1, 2N-2]$. Hence for $x \in (2N-3, 2N-2)$ we have $0 = f_1(2^{-j}x) = 2^{\frac{j}{2}}c_{-1}\varphi(x-(-1))$. As $(2N-3, 2N-2) \cap \text{supp}(\varphi(x-(-1))) \neq \emptyset$, then $c_{-1} = 0$. Thus,

$$f_1(2^{-j}x) = \sum_{k=-2N+2}^{-2} 2^{\frac{j}{2}} c_k\varphi(x-k) = 0 \quad \text{for all } x \geq 0.$$

We repeat the above argument, with $k_0 = -2$. If $-2N+2 \leq k \leq -3$, then $\text{supp}(\varphi(x-k)) \subset (-\infty, 2N-4]$ while for $k_0 = -2$, $\text{supp}(\varphi(x-k_0)) \subset [-2, 2N-3]$. Hence for $x \in (2N-4, 2N-3)$ we have $0 = f_2(2^{-j}x) = 2^{\frac{j}{2}}c_{-2}\varphi(x-(-2))$. As

$[2N-4, 2N-3] \cap \operatorname{supp}(\varphi(x-(-2))) \neq \emptyset$, then $c_{-2} = 0$.

Continuing this way, we obtain that $c_k = 0$ for all $-2N+2 \leq k \leq -1$, that is, $f_1 = 0$. It follows that

$$P : X_j^{(1)} \longrightarrow Y_j^{(1)}$$

is one to one.

Next let $f_3(x) = \sum_{k \in S_3(j)} c_k \varphi_{j,k}(x) \in X_j^{(3)}$ and suppose that $Pf_3 = 0$. That is,

$$f_3(x) = \sum_{k=2^j-2N+2}^{2^j-1} c_k \varphi_{j,k}(x) = 0 \text{ for all } x \in [0,1]. \tag{C.2}$$

As $\operatorname{supp}(\varphi_{j,k}) \subset [2^{-j}k, 2^{-j}(2N+k-1)]$ and $j \geq j_0$, it follows that (C.2) holds for all $x \in (-\infty, 1]$.

Set

$$\begin{aligned}\tilde{f}_3(x) = f_3(x+1) &= \sum_{k=2^j-2N+2}^{2^j-1} c_k 2^{\frac{j}{2}} \varphi(2^j x + 2^j - k) \\ &= 2^{\frac{j}{2}} \sum_{k=-2N+2}^{-1} c_{k+2^j} \varphi(2^j x - k).\end{aligned}$$

Then $f_3 \in V_j(\mathbb{R})$ and $\tilde{f}_3(2^{-j}x) = 2^{\frac{j}{2}} \sum_{k=-2N+2}^{-1} c_{k+2^j} \varphi(x-k) = 0$ for all $x \in (-\infty, 0]$. By lemma 2.8 we have $c_{k+2^j} = 0$ for $k \leq -1$, that is, $c_k = 0$ for $2^j - 2N + 2 \leq k \leq 2^j - 1$. This shows that

$$P : X_j^{(3)} \longrightarrow Y_j^{(3)}$$

is one to one.

Next we claim that there exist $C_2 > C_1 > 0$ so that for any $j \geq j_0$ and for any sequence $\{a_{j,k}\}_{k \in S(j)}$ of coefficients

$$C_1 \left(\sum_{k \in S(j)} |a_{j,k}|^2 \right)^{\frac{1}{2}} \leq \left\| \sum_{k \in S(j)} a_{j,k} \varphi_{j,k}(x) \right\|_{L^2[0,1]} \leq C_2 \left(\sum_{k \in S(j)} |a_{j,k}|^2 \right)^{\frac{1}{2}}. \tag{C.3}$$

Let any $j \geq j_0$ be fixed. Let $X_j = X_j^{(1)} + X_j^{(2)} + X_j^{(3)}$ and $Y_j = Y_j^{(1)} + Y_j^{(2)} + Y_j^{(3)}$. Then as $\{\varphi_{j,k}\}_{k \in S(j)}$ are orthonormal, (C.3) is equivalent to

$$C_1 \|f\|_{L^2(\mathbb{R})} \leq \|Pf\|_{L^2([0,1])} \leq C_2 \|f\|_{L^2(\mathbb{R})} \tag{C.4}$$

for $f = \sum_{k \in S(j)} a_{j,k} \varphi_{j,k} \in X_j$.

We first show that (C.4) holds for all $f \in X_j^{(1)}$. Observe that

1) $X_j^{(1)}$ and $Y_j^{(1)}$ are $2N-2$ dimensional vector spaces.

2) If $f(x) = \sum_{k=-2N+2}^{-1} c_k \varphi_{j,k}(x) \in X_j^{(1)}$, then as $\text{supp}(\varphi_{j,k}) \subset [2^{-j}k, 2^{-j}(2N+k-1)]$ we have that $\varphi_{j,k}(x) = 0$ for all $x \geq 2^{-j}(2N-2)$, and as $j \geq j_0$ then, $f(x) = 0$ for all $x \geq 1$. Observe that the following diagram is commutative:

$$\begin{array}{lcl} X_j^{(1)}, \|\cdot\|_{L^2(\mathbb{R})} & \xrightarrow{D_{2^j}} & X_0^{(1)}, \|\cdot\|_{L^2(\mathbb{R})} \\ P \downarrow & & \downarrow Q \\ Y_j^{(1)}, \|\cdot\|_{L^2[0,1]} & \xrightarrow{D_{2^j}} & Y_0^{(1)}, \|\cdot\|_{L^2[0,\infty)} \end{array}$$

where $X_0^{(1)} = \text{span}\{\varphi(x+1), ..., \varphi(x+2N-2)\}$

$Y_0^{(1)} = \text{span}\{\varphi(x+1)\big|_{[0,\infty)}, ..., \varphi(x+2N-2)\big|_{[0,\infty)}\}$

$Q : f \longmapsto f\big|_{[0,\infty)}$ the restriction map

$(D_{2^j} f)(x) = 2^{-\frac{j}{2}} f(2^{-\frac{j}{2}} x)$.

Since $f_1 = \sum_{k=-2N+2}^{-1} c_k \varphi_{j,k} \in X_j^{(1)}$ then $(D_{2^j} f_1)(x) = \sum_{k=-2N+2}^{-1} c_k \varphi(x-k) \in X_0^{(1)}$. Thus, clearly, D_{2^j} is an isometry of $X_j^{(1)}$ onto $X_0^{(1)}$ in the norm $\|\cdot\|_{L^2(\mathbb{R})}$. Similarly, D_{2^j} is an isometry of $Y_j^{(1)}$ onto $Y_0^{(1)}$, when $Y_j^{(1)}$ is given $\|\cdot\|_{L^2[0,1]}$ and $Y_0^{(1)}$ is given $\|\cdot\|_{L^2[0,\infty)}$. In fact

1) Since $\varphi_{j,k}(x) = 0$ for $x \geq \frac{1}{2}$ we may consider an element of $Y_j^{(1)}$ as defined on $[0,\infty)$.

2) If $\tilde{f}(x) = \sum_{k=-2N+2}^{-1} c_k \varphi_{j,k}(x)\big|_{[0,infty)} \in Y_j^{(1)}$ for $x \geq 0$, then $(D_{2^j} \tilde{f})(x) = \sum_{k=-2N+2}^{-1} c_k \varphi(x-k)\big|_{[0,\infty)} \in Y_0^{(1)}$ for $x \geq 0$.

3) $\|D_{2^j}\tilde{f}\|_{L^2[0,\infty)} = \|\tilde{f}\|_{L^2[0,\infty)} = \|\tilde{f}\|_{L^2[0,1]}$ as $\mathrm{supp}(\tilde{f}) \subset [0,1]$.

Now Q maps basis vector to basis vector and is linear, hence defines an isomorphism between the two finite dimensional vector spaces $X_0^{(1)}$ and $Y_0^{(1)}$. Hence there exist $C_1^{(1)}, C_2^{(1)} > 0$ such that

$$C_1^{(1)}\|f_1\|_{L^2(\mathbb{R})} \leq \|Qf_1\|_{L^2[0,\infty)} \leq C_2^{(1)}\|f_1\|_{L^2(\mathbb{R})}$$

for all $f_1 \in X_0^{(1)}$. As $D_{2^j}, D_{2^{-j}}$ are isometries, we have

$$C_1^{(1)}\|f_1\|_{L^2(\mathbb{R})} \leq \|Pf_1\|_{L^2[0,1]} \leq C_2^{(1)}\|f_1\|_{L^2(\mathbb{R})}$$

for all $f_1 \in X_j^{(1)}$ for $j \geq j_0$.

A similar argument shows that there exist $C_1^{(3)}, C_2^{(3)} > 0$ such that

$$C_1^{(3)}\|f_3\|_{L^2(\mathbb{R})} \leq \|Pf_3\|_{L^2[0,\infty)} \leq C_2^{(3)}\|f_3\|_{L^2(\mathbb{R})}$$

for all $f_3 \in X_j^{(3)}$ for $j \geq j_0$.

Observe that for $f_2 \in X_j^{(2)}$, as $\mathrm{supp}(\varphi_{j,k}) \subset [0,1]$ for all $k \in S_2(j)$, then $\mathrm{supp}(f) \subset [0,1]$ and we have

$$\|f_2\|_{L^2(\mathbb{R})} = \|f_2\|_{L^2[0,1]}.$$

Next let $f \in X_j$ be arbitrary, say $f(x) = f_1(x) + f_2(x) + f_3(x)$ where $f_i \in X_j^{(i)}$, that is, $f_i(x) = \sum_{k \in S_i(j)} a_{j,k}\varphi_{j,k}(x); i = 1, 2, 3.$

Now

$$\begin{aligned}
\|f_2\|_{L^2[0,1]} = \|f_2\|_{L^2(\mathbb{R})} &= \sum_{k \in S_2(j)} |a_{j,k}|^2 \\
&= \sum_{k \in S_2(j)} \left|< f, \varphi_{j,k} >_{L^2(\mathbb{R}^d)}\right|^2 \quad \text{as } \{\varphi_{j,k}\} \text{ is an orthonormal basis of } X_j \\
&= \sum_{k \in S_2(j)} \left|< f, \varphi_{j,k} >_{L^2[0,1]}\right|^2 \quad \text{as } \mathrm{supp}(\varphi_{j,k}) \subset [0,1] \\
&\leq \|f\|^2_{L^2[0,1]} \quad \text{by Bessel's inequality and since } \{\varphi_{j,k}\}_{k \in S_2(j)} \text{ is} \\
&\qquad\qquad \text{orthonormal in } L^2[0,1].
\end{aligned}$$

It follow that

$$\begin{aligned}\|f_1+f_3\|_{L^2[0,1]} &= \|f-f_2\|_{L^2[0,1]}\\ &\le \|f\|_{L^2[0,1]}+\|f_2\|_{L^2[0,1]}\\ &\le \|f\|_{L^2[0,1]}+\|f\|_{L^2[0,1]} = 2\|f\|_{L^2[0,1]}.\end{aligned}$$

Now as f_1 and f_3 have disjoint supports then

$$\|f_1\|^2_{L^2[0,1]}+\|f_3\|^2_{L^2[0,1]} = \|f_1+f_3\|^2_{L^2[0,1]}$$

then

$$\begin{aligned}\|f_1\|_{L^2[0,1]} &\le \|f_1+f_3\|_{L^2[0,1]} \le 2\|f\|_{L^2[0,1]}\\ \|f_3\|_{L^2[0,1]} &\le \|f_1+f_3\|_{L^2[0,1]} \le 2\|f\|_{L^2[0,1]}\end{aligned}$$

Thus,

$$\begin{aligned}\|f\|_{L^2(\mathbb{R})} &= \|f_1+f_2+f_3\|_{L^2(\mathbb{R})}\\ &\le \|f_1\|_{L^2(\mathbb{R})}+\|f_2\|_{L^2(\mathbb{R})}+\|f_3\|_{L^2(\mathbb{R})}\\ &\le \frac{1}{C_1^{(1)}}\|f_1\|_{L^2[0,1]}+\|f_2\|_{L^2[0,1]}+\frac{1}{C_1^{(3)}}\|f_3\|_{L^2[0,1]}\\ &\le \frac{2}{C_1^{(1)}}\|f\|_{L^2[0,1]}+\|f\|_{L^2[0,1]}+\frac{2}{C_1^{(3)}}\|f\|_{L^2[0,1]}\\ &= \left(\frac{2}{C_1^{(1)}}+1+\frac{2}{C_1^{(3)}}\right)\|f\|_{L^2[0,1]}\end{aligned}$$

where we have used the fact that f_1, f_2 and f_3 are mutually orthonormal in $L^2(\mathbb{R}^d)$. Thus, the left inequality holds with $\frac{1}{C_1} = \frac{2}{C_1^{(1)}}+1+\frac{2}{C_1^{(3)}}$. For the right hand side,

$$\begin{aligned}\|f\|_{L^2[0,1]} &\le \|f_1\|_{L^2[0,1]}+\|f_2\|_{L^2[0,1]}+\|f_3\|_{L^2[0,1]}\\ &\le C_2^{(1)}\|f_1\|_{L^2(\mathbb{R})}+\|f_2\|_{L^2(\mathbb{R})}+C_2^{(3)}\|f_3\|_{L^2(\mathbb{R})}\\ &\le C_2^{(1)}\|f\|_{L^2(\mathbb{R})}+\|f\|_{L^2(\mathbb{R})}+C_2^{(3)}\|f\|_{L^2(\mathbb{R})}\\ &= \left(C_2^{(1)}+1+C_2^{(3)}\right)\|f\|_{L^2(\mathbb{R})}\end{aligned}$$

and the right inequality holds with $C_2 = C_2^{(1)} + 1 + C_2^{(3)}$.

Hence inequality (C.3) holds. $\square$

The proof of Lemma 2.11

Proof. Below, we will replace x by $x + 2N - 1 - l$ where $l = 1, 2, ..., N - 1$ in equation (2.19) . Note that, for each l, we have for $-l > 1 - N$, if $x \geq 0$,

$$2(x + 2N - 1 - l) \geq 2(0 + 2N - 1 + 1 - N) = 2N,$$

so that by supp$\varphi \subset [0, 2N - 1]$,

$$\varphi\left(2(x + 2N - 1 - l)\right) = 0 \quad \text{for all} \ \ l = 1, 2, ..., N - 1. \tag{C.5}$$

First, we replace x by $x + 2N - 2$ in equation (2.19) and get

$$\begin{aligned}&\sqrt{2}\varphi(2(x + 2N - 2))\\ &\quad = \overline{h}_0\varphi(x + 2N - 2) + \overline{h}_2\varphi(x + 2N - 1) + ... + \overline{h}_{2N-2}\varphi(x + 3N - 3)\\ &\quad + \overline{g}_0\psi(x + 2N - 2) + \overline{g}_2\psi(x + 2N - 1) + ... + \overline{g}_{2N-2}\psi(x + 3N - 3).\end{aligned}$$

Now, if $x \geq 0$ and $k \leq 1$, then $x + 2N - k \geq 2N - 1$, and as suppφ, suppψ lie in $[0, 2N - 1]$ we have $\varphi(x + 2N - k) = 0 = \psi(x + 2N - k)$ for all $k \leq 1$. Applying (C.5), the above equation yields

$$\psi(x - (-2N + 2)) = \frac{-\overline{h}_0}{\overline{g}_0}\varphi(x - (-2N + 2)). \tag{C.6}$$

Next, we replace x by $x + 2N - 3$ in equation (2.19),

$$\begin{aligned}&\sqrt{2}\varphi(2(x + 2N - 3))\\ &\quad = \overline{h}_0\varphi(x + 2N - 3) + \overline{h}_2\varphi(x + 2N - 2) + ... + \overline{h}_{2N-2}\varphi(x + 3N - 4)\\ &\quad + \overline{g}_0\psi(x + 2N - 3) + \overline{g}_2\varphi(x + 2N - 2) + ... + \overline{g}_{2N-2}\psi(x + 3N - 4).\end{aligned}$$

Again, suppφ, suppψ are contained in $[0, 2N - 1]$ and we have $\varphi(x + 2N - k) = 0 = \psi(x + 2N - k)$ for $x \geq 0$ and $k \leq -1$, and hence by (C.5), $0 = \overline{h}_0\varphi(x + 2N -$

$3) + \overline{h}_2\varphi(x+2N-2) + \overline{g}_0\psi(x+2N-3) + \overline{g}_2\psi(x+2N-2)$. Then by equation (C.6) we have

$$\psi(x-(-2N+3)) = \frac{-\overline{h}_0}{\overline{g}_0}\varphi(x-(-2N+3)) + \left(\frac{\overline{h}_0\overline{g}_2}{g_0^2} - \frac{\overline{h}_2}{\overline{g}_0}\right)\varphi(x-(-2N+3)+1). \tag{C.7}$$

Continuing by induction, we finally replace x by $x+N$ in equation (2.19),

$$\begin{aligned}\sqrt{2}\varphi\left(2(x+N)\right) = {} & \overline{h}_0\varphi(x+N) + \overline{h}_2\varphi(x+N+1) + ... + \overline{h}_{2N-2}\varphi(x+2N-1) \\ & + \overline{g}_0\psi(x+N) + \overline{g}_2\psi(x+N+1) + ... + \overline{g}_{2N-2}\psi(x+2N-1).\end{aligned}$$

Then, as $\varphi(x+2N-1) = 0 = \psi(x+2N-1)$ for $x \geq 0$ and applying (C.5) we have

$$\begin{aligned}0 = {} & \overline{h}_0\varphi(x+N) + \overline{h}_2\varphi(x+N+1) + ... + \overline{h}_{2N-4}\varphi(x+2N-2) \\ & + \overline{g}_0\psi(x+N) + \overline{g}_2\psi(x+N+1) + ... + \overline{g}_{2N-4}\psi(x+2N-2).\end{aligned}$$

That is,

$$\begin{aligned}\psi\left(x+N\right) = {} & -\frac{\overline{h}_0}{\overline{g}_0}\varphi\left(x+N\right) - \frac{\overline{h}_2}{\overline{g}_0}\varphi\left(x+N+1\right) - ... \\ & - \frac{\overline{h}_{2N-4}}{\overline{g}_0}\varphi\left(x+2N-2\right) - \frac{\overline{g}_2}{\overline{g}_0}\psi\left(x+N+1\right) - ... \\ & - \frac{\overline{g}_{2N-4}}{\overline{g}_0}\psi\left(x+2N-2\right),\end{aligned}$$

Applying the results of the previous induction steps, we see that each $\psi(x+N+k), 0 \leq k \leq N-2$ can be expressed on $[0,\infty)$ as a linear combinations of function $\varphi(x+N+r); k < r \leq N-2$,

$$\psi(x+N+k) = \sum_{r=k+1}^{N-2} a_r^{(k)}\varphi(x+N+r)$$

for some coefficients $a_r^{(k)}$ determined by the wavelet and scaling filters. That is, each $\psi(x-k)\big|_{[0,\infty)}, -2N+2 \leq k \leq -N$, is a linear combination

$$\sum_{r=2-2N}^{k-1} c_r^{(k)}\varphi(x-r).$$

We conclude that the functions $\psi(x-k)\big|_{[0,\infty)}, \quad -2N+2 \le k \le -N$ belong to $V_0[0,\infty)$. This proves the lemma. □

The proof of Lemma 2.12

Proof. Part I: By Lemma 2.11, the functions $\psi(x-k), -2N+2 \le k \le -N$, when restricted to $[0,\infty)$, belong to $V_0[0,\infty)$. That is, each $\psi(x-k)$ is a linear combination of functions $\varphi_l(x) = \varphi(x-l)$ when x is restricted to $[0,\infty)$.

Replacing now x by $2^j x$, we have the function $\psi(2^j x - k), -2N+2 \le k \le -N$, when restricted to $[0,\infty)$, belong to $V_j[0,\infty)$ and in fact, are linear combinations of functions $\varphi_{j,l}(x)\big|_{[0,\infty)}$. Now, if we restrict x further to $[0,1]$ we obtain $\psi(2^j x - k)\big|_{[0,1]}$ and clearly, $\psi(2^j x - k)\big|_{[0,1]}$ is a linear combination of $\varphi_{j,l}\big|_{[0,1]}$. That is, $\psi(2^j x - k), -2N+2 \le k \le -N$, when restricted to $[0,1]$, belong to $V_j[0,1]$.

Part II: By part I, the functions $\psi(x-k)\big|_{[0,1]}, -2N+2 \le k \le -N$, belong to $V_0[0,1]$.

Next, the same argument of Part I, we have $\psi(2^j x - k)\big|_{[0,1]}$ for $2^j - N + 1 \le k \le 2^j - 1$. □

The proof of theorem 2.13

Proof. Note that by Theorem 2.9, the functions $\varphi_{j,k}\big|_{[0,1]}, -2N+2 \le k \le 2^j - 1$, form a Riesz basis of the space $V_j([0,1])$. That is, the dimension of $V_{j+1}[0,1]$ is $2^{j+1} + 2N - 2$, and now card$\{\varphi_{j,k}, -2N+2 \le k \le 2^j - 1\} = 2^j + 2N - 2$ and card$\{\psi_{j,k}, -N+1 \le k \le 2^j - N\} = 2^j$.

It remains to show that, for an arbitrary function f of V_{j+1}, the restriction of f to $[0,1]$ can we written as $g + h$ where $g \in V_j[0,1]$ and

$$h(x) = \sum_{k=-N+1}^{2^j - N} \alpha_{j,k} \psi_{j,k}(x).$$

In fact, let $f \in V_{j+1}[0,1]$. Then $f = g + h$ where $g \in V_j[0,1]$ and $h \in V_j[0,1]^{\perp}$. As such, $h = \sum_k \beta_{j,k}\psi_{j,k}$. Consider the following.

If $k \leq -2N+1$, then $2^j x - k \geq -k \geq 2N-1$. Thus $\psi_{j,k}\big|_{[0,1]} = 0$.

If $-2N+2 \leq k \leq -N$, then by Lemma 2.12, $\psi_{j,k}\big|_{[0,1]} \in V_j[0,1]$.

If $k \geq 2^j$, then $2^j x - k \leq 2^j - k \leq 2^j < 0$. Thus $\psi_{j,k}\big|_{[0,1]} = 0$.

If $2^j - N + 1 \leq k \leq 2^j - 1$, then $\psi_{j,k}\big|_{[0,1]} \in V_j[0,1]$.

These show that $\psi(x-k)\big|_{[0,1]} \neq 0$ and does not belong to $V_j[0,1]$ if $-N+1 \leq k \leq 2^j - N$, and hence

$$h(x) = \sum_{k=-N+1}^{2^j-N} \alpha_{j,k}\psi_{j,k}(x).$$

□

APPENDIX D

GAUSSIAN RANDOM FIELD

Definition D.1. (Multivariate Gaussian Distribution)

Let $V = [V_1, V_2, ..., V_m]^T$ be an m-dimensional random variable vector. We define the mean vector of V to be the vector $E[V] = [E[V_1], ..., E[V_m]]^T$ and the covariance matrix V to be $\Sigma = [\sigma_{ij}]_{m\times m}$ where $\sigma_{ij} = \mathrm{Cov}(V_i, V_j)$ for $i, j = 1, 2, ..., m$. V is called multivariate Gaussian with mean $E[V]$ and covariance matrix σ if the density function of V is given by

$$f(x) = (2\pi)^{-\frac{m}{2}} (\det \sigma)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-E[V])^T \sigma^{-1}(x-E[V])} \quad \text{for all } x = (x_1, ..., x_m) \in \mathbb{R}^m$$

Remark D.1. If $V = (V_1, ..., V_n)$ is a multivariate Gaussian with mean $E[V]$ and covariance matrix σ, then one can show (see in Grimmett and Stirzaker (1998)) $Y = a_1V_1 + a_2V_2 + ... + a_nV_n$, for constants $a_1, ..., a_n$, has the Gaussian distribution with mean $\sum_{i=1}^{n} a_i E[V_i]$ and variance $\sum_{i=1}^{n} a_i^2 \mathrm{Var}[V_i] + 2\sum_{i<j} a_i a_j \mathrm{Cov}[V_i, V_j]$.

Remark D.2. If $\sigma = \mathrm{diag}(\sigma_{ii})$ we get

$$\begin{aligned} f(x) &= (2\pi)^{-\frac{m}{2}} \prod_{i=1}^{m} (\sigma_{ii})^{-\frac{1}{2}} \, e^{-\frac{1}{2}(x-E[V])^T \sigma^{-1}(x-E[V])} \\ &= \prod_{i=1}^{m} (2\pi\sigma_{ii})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{m}(x_i - E[V_i])^2 \sigma_{ii}^{-1}\right) \\ &= \prod_{i=1}^{m} (2\pi\sigma_{ii})^{-\frac{1}{2}} e^{-\frac{(x_i - E[V_i])^2}{2\sigma_{ii}}} \\ &= \prod_{i=1}^{m} f_{V_i}(x_i) \end{aligned}$$

where $f_X(x) = (2\pi\rho)^{-\frac{1}{2}} e^{-\frac{(x-E[X])^2}{2\rho}}$ is a normal density function with mean $E[X]$ and variance ρ. This shows that the components in a multivariate Gaussian

$(V_1, V_2, ..., V_m)$ are independent, if it has a diagonal covariance matrix.

Definition D.2. (Characteristic Functions)

Let V be any m-dimensional vector of random variables of distribution μ, and $t = (t_1, t_2, ..., t_m)$ a vector of real numbers. The characteristic function of V is defined by the function $\phi : \mathbb{R}^m \to \mathbb{C}$ given by $\phi(t) = E(e^{it^T V})$. Note that if μ is the distribution of V, then $\phi(t)$ is the inverse Fourier transform $\check{\mu}$ of μ, up to a scaling factor. In particular ϕ uniquely determines the distribution μ.

Remark D.3.

By the identity $\int_{\mathbb{R}} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}}$ we have by the covariance matrix σ is a Hermitian matrix $\int_{\mathbb{R}^d} e^{-x^T \sigma x}\, dx = \sqrt{\frac{\pi^d}{|\det \sigma|}}$ and for a m-dimensional multivariate Gaussian distribution V with mean $E[V]$ and covariance matrix σ we have characteristic function

$$E[e^{it^T V}] = \frac{1}{(2\pi)^{\frac{m}{2}} |\det \sigma|^{\frac{1}{2}}} \int_{\mathbb{R}^m} e^{it^T x} e^{-\frac{1}{2}(x - E[V])^T \sigma^{-1} (x - E[V])}\, dx.$$

By a change of variables, $x \mapsto \sigma x + E[V]$ we have

$$\begin{aligned}
E[e^{it^T V}] &= \frac{1}{(2\pi)^{\frac{m}{2}} |\det \sigma|^{\frac{1}{2}}} \int_{\mathbb{R}^m} e^{it^T(\sigma x + E[V])} e^{-\frac{1}{2}(\sigma x)^T x} |\det \sigma|\, dx \\
&= \frac{|\det \sigma|^{\frac{1}{2}}}{(2\pi)^{\frac{m}{2}}} e^{it^T E[V]} \int_{\mathbb{R}^m} e^{it^T \sigma x} e^{-\frac{1}{2} x^T \sigma x}\, dx \\
&= \frac{|\det \sigma|^{\frac{1}{2}}}{(2\pi)^{\frac{m}{2}}} e^{it^T E[V] - \frac{1}{2} t^T \sigma t} \int_{\mathbb{R}^m} e^{-\frac{1}{2}(x - it)^T \sigma (x - it)}\, dx \\
&= \frac{|\det \sigma|^{\frac{1}{2}}}{(2\pi)^{\frac{m}{2}}} e^{it^T E[V] - \frac{1}{2} t^T \sigma t} \left(\frac{\pi^{\frac{m}{2}}}{\left(\frac{1}{2}\right)^{\frac{m}{2}} |\det(\sigma)|^{\frac{1}{2}}} \right) \\
&= e^{it^T E[V] - \frac{1}{2} t^T \sigma t}.
\end{aligned}$$

Definition D.3. (Gaussian Random Field)

A random field $\{X_t\}_{t \in \mathbb{R}^d}$ is called a Gaussian random field if for any $m \in \mathbb{N}$, and any choice of $t_1, ..., t_m,\ \ t_i \in \mathbb{R}^d$, the random vector $\{X_{t_1}, X_{t_2}, ..., X_{t_m}\}$, has the multivariate Gaussian distribution.

Lemma D.1. *Let $X_1, X_2, ..., X_n$ be independent Normal Random variables with distributions $N(E[X_1], \sigma^2_{X_1}), ..., N(E[X_n], \sigma^2_{X_n})$, respectively. Then the distribution of $(X_1, X_1 + X_2, X_1 + X_2 + X_3, ..., X_1 + X_2 + ... + X_n)$ is multivariate Gaussian with mean $(E[X_1], E[X_1] + E[X_2], ..., E[X_1] + ... + E[X_n])$ and covariance matrix*

$$\sigma = \begin{bmatrix} \sigma^2_{X_1} & \sigma^2_{X_1} & \sigma^2_{X_1} & \cdots & \sigma^2_{X_1} \\ \sigma^2_{X_1} & \sigma^2_{X_1} + \sigma^2_{X_2} & \sigma^2_{X_1} + \sigma^2_{X_2} & \cdots & \sigma^2_{X_1} + \sigma^2_{X_2} \\ \sigma^2_{X_1} & \sigma^2_{X_1} + \sigma^2_{X_2} & \sigma^2_{X_1} + \sigma^2_{X_2} + \sigma^2_{X_3} & \cdots & \sigma^2_{X_1} + \sigma^2_{X_2} + \sigma^2_{X_3} \\ \vdots & \vdots & \vdots & & \vdots \\ \sigma^2_{X_1} & \sigma^2_{X_1} + \sigma^2_{X_2} & \sigma^2_{X_1} + \sigma^2_{X_2} + \sigma^2_{X_3} & \cdots & \sigma^2_{X_1} + \sigma^2_{X_2} + ... + \sigma^2_{X_n} \end{bmatrix}.$$

Proof. Note that $E[X_1 + X_2] = E[X_1] + E[X_2]$. Since $\text{Var}[X_1] = \sigma^2_{X_1}$, we have by independence, $\text{Cov}[X_1, X_2] = E[X_1X_2] - E[X_1][X_2] = 0$, and hence $\text{Cov}[X_1, X_1 + X_2] = \text{Cov}[X_1, X_1] + \text{Cov}[X_2, X_1] = \sigma^2_{X_1}$, and $\text{Var}[X_1 + X_2] = \text{Var}[X_1] + \text{Var}[X_2] + 2\text{Cov}[X_1, X_2] = \sigma^2_{X_1} + \sigma^2_{X_2} + 2(0) = \sigma^2_{X_1} + \sigma^2_{X_2}$. Next, set $V = (X_1, X_1 + X_2)^T$, $E[V] = (E[X_1], E[X_1] + E[X_2])$, and $\sigma = \begin{bmatrix} \sigma^2_{X_1} & \sigma^2_{X_1} \\ \sigma^2_{X_1} & \sigma^2_{X_1} + \sigma^2_{X_2} \end{bmatrix}$, consider

$$E[e^{it^TV}] = E[e^{i(t_1X_1 + t_2(X_1+X_2))}] = E[e^{i(t_1+t_2)X_1}e^{it_2X_2}].$$

Since X_1 and X_2 are independent we have

$$E[e^{it^TV}] = E[e^{i(t_1+t_2)X_1}]E[e^{it_2X_2}] = e^{i(t_1+t_2)E[X_1] - \frac{(t_1+t_2)^2}{2}\sigma^2_{X_1}} e^{i(t_2)E[X_2] - \frac{t_2^2}{2}\sigma^2_{X_2}}$$
$$= e^{i(t_1E[X_1] + t_2(E[X_1]+E[X_2]))} e^{-\frac{1}{2}[(t_1^2 + 2t_1t_2 + t_2^2)\sigma^2_{X_1} + t_2^2\sigma^2_{X_2}]} = e^{it^TE[V]}e^{-\frac{1}{2}t^T\sigma t}.$$

Hence $\phi(t) = e^{it^TE[V] - \frac{1}{2}t^T\sigma t}$ is the characteristic function of $V = (X_1, X_1 + X_2)$, which shows that V is bivariate Gaussian with mean $(E[X_1], E[X_1] + E[X_2])$ and covariance matrix

$$\sigma = \begin{bmatrix} \sigma^2_{X_1} & \sigma^2_{X_1} \\ \sigma^2_{X_1} & \sigma^2_{X_1} + \sigma^2_{X_2} \end{bmatrix}.$$

Now $E[X_1+...+X_k]=E[X_1]+...+E[X_k]$ for all $2\le k\le n$. For each $1\le k\le n$,

$$\mathrm{Cov}[X_1,\sum_{i=1}^{k}X_i]=\mathrm{Cov}[\sum_{i=1}^{k}X_i,X_1]=\mathrm{Var}(X_1)+\sum_{i=2}^{k}\mathrm{Cov}[X_i,X_1]=\sigma^2_{X_1}$$

as $\sum_{i=2}^{k}\mathrm{Cov}[X_i,X_1]=0$ by independence.
Furthermore, for each $2\le k\le n$,

$$\begin{aligned}\mathrm{Cov}[X_1+X_2,\sum_{i=1}^{k}X_i]&=\mathrm{Cov}[X_1,\sum_{i=1}^{k}X_i]+\mathrm{Cov}[X_2,\sum_{i=1}^{k}X_i]\\&=\mathrm{Cov}[X_1,X_1]+\sum_{i=2}^{k}\mathrm{Cov}[X_1,X_i]+\mathrm{Cov}[X_2,X_2]\\&\quad+\sum_{i=1,i\neq 2}^{k}\mathrm{Cov}[X_2,X_i]\\&=\mathrm{Var}(X_1)+\mathrm{Var}(X_2)\\&=\sigma^2_{X_1}+\sigma^2_{X_2}\end{aligned}$$

as $\sum_{i=1,i\neq j}^{k}\mathrm{Cov}[X_i,X_j]=0$ by independence.
Continuing by induction, we get

$$\sigma=\begin{bmatrix}\sigma^2_{X_1} & \sigma^2_{X_1} & \sigma^2_{X_1} & \dots & \sigma^2_{X_1}\\ \sigma^2_{X_1} & \sigma^2_{X_1}+\sigma^2_{X_2} & \sigma^2_{X_1}+\sigma^2_{X_2} & \dots & \sigma^2_{X_1}+\sigma^2_{X_2}\\ \sigma^2_{X_1} & \sigma^2_{X_1}+\sigma^2_{X_2} & \sigma^2_{X_1}+\sigma^2_{X_2}+\sigma^2_{X_3} & \dots & \sigma^2_{X_1}+\sigma^2_{X_2}+\sigma^2_{X_3}\\ \vdots & \vdots & \vdots & & \vdots\\ \sigma^2_{X_1} & \sigma^2_{X_1}+\sigma^2_{X_2} & \sigma^2_{X_1}+\sigma^2_{X_2}+\sigma^2_{X_3} & \dots & \sigma^2_{X_1}+\sigma^2_{X_2}+...+\sigma^2_{X_n}\end{bmatrix}$$

Set $V=(X_1,X_1+X_2,...,X_1+X_2+...+X_n)^T$, then

$$E[V]=(E[X_1],E[X_1]+E[X_2],...,E[X_1]+E[X_2]+...+E[X_n])^T.$$

Consider for all $t=(t_1,t_2,...,t_n)^T$

$$\begin{aligned}E[e^{it^TV}]&=E[\exp(i(t_1X_1+t_2(X_1+X_2)+...+t_n(X_1+X_2+...+X_n))]\\&=E[\exp(i([(t_1+t_2+...+t_n)X_1+(t_2+t_3+...+t_n)X_2+...+t_nX_n]).\end{aligned}$$

By independence of the random variables $X_1, X_2, ..., X_n$ and continuity of the exponential function, we may use Theorem 3.1 to obtain that

$$\begin{aligned} E[e^{it^TV}] &= E[\exp(i[t_1 + ... + t_n]X_1)]E[\exp(i[t_2 + ... + t_n]X_2)]...E[\exp(it_nX_n)] \\ &= \exp\left(i(t_1 + t_2 + ... + t_n)E[X_1] - \frac{1}{2}(t_1 + t_2 + ... + t_n)^2\sigma^2_{X_1}\right) \\ &\quad \exp\left(i(t_2 + t_3 + ... + t_n)E[X_2] - \frac{1}{2}(t_2 + t_3 + ... + t_n)^2\sigma^2_{X_2}\right)... \\ &\quad \exp\left(it_nE[X_n] - \frac{1}{2}t_n^2\sigma^2_{X_n}\right). \end{aligned}$$

We can rewrite this in term of $E[V]$ and σ as

$$\begin{aligned} E[e^{it^TV}] &= \exp\left(i(t_1E[X_1] + t_2(E[X_1] + E[X_2]) + ... + t_n(E[X_1] + ... + E[X_n]))\right) \\ &\quad \exp\left(-\frac{1}{2}((t_1 + t_2 + ... + t_n)^2\sigma^2_{X_1} + (t_2 + t_3 + ... + t_n)^2\sigma^2_{X_2} + ... + t_n^2\sigma_{X_n})\right) \\ &= \exp\left(it^TE[V] - \frac{1}{2}t^T\sigma t\right). \end{aligned}$$

This shows that $(X_1, X_1 + X_2, ..., X_1 + X_2 + ... + X_n)$ is multivariate Gaussian, by remark D.3. □

APPENDIX E

BROWNIAN MOTION

Theorem E.1. *Standard Brownian motion is a Gaussian process with mean function zero and covariance function* $Cov[X_t, X_s] = \min(t, s)$.

Proof. Let $\{B_t\}_{t\geq 0}$ be a Brownian motion. Let $n \in \mathbb{N}$ and for $t_1 < t_2 < ... < t_n$, consider the vector $\{B_{t_1}, B_{t_2}, ..., B_{t_n}\}$.

By (B1), (B2) , the increments $B_t - B_s$ are independent and Gaussian distributed with mean 0 and variance $t-s$ for all $t > s$, that is $B_{t_1}, B_{t_2}-B_{t_1}, B_{t_3}-B_{t_2}, ..., B_{t_n}-B_{t_{n-1}}$ are independent and $N(0, t_{i+1} - t_i)$ for all $i = 1, 2, 3, ..., n$.

We now apply Lemma D.1 in Appendix D, by choosing

$$
\begin{aligned}
X_1 &= B_{t_1} \\
X_2 &= B_{t_2} - B_{t_1} \\
&\vdots \\
X_n &= B_{t_n} - B_{t_{n-1}}.
\end{aligned}
$$

We thus obtain that

$$
\begin{aligned}
X_1 + X_2 &= B_{t_2} \\
X_1 + X_2 + X_3 &= B_{t_3} \\
&\vdots \\
X_1 + X_2 + ... + X_n &= B_{t_n}
\end{aligned}
$$

and hence

$$
\begin{aligned}
E[X_1] = E[B_{t_1}] &= 0 & \sigma^2_{X_1} &= \sigma^2_{B_{t_1}} = t_1 \\
E[X_1 + X_2] = E[B_{t_2}] &= 0 & \sigma^2_{X_2} &= \sigma^2_{B_{t_2} - B_{t_1}} = t_2 - t_1 \\
&\vdots & &\vdots \\
E[X_1 + X_2 + ... + X_n] = E[B_{t_n}] &= 0 & \sigma^2_{X_n} &= \sigma^2_{B_{t_n} - B_{t_{n-1}}} = t_n - t_{n-1}.
\end{aligned}
$$

It follows from Lemma D.1, that $(B_{t_1}, ..., B_{t_n})$ is multivariate Gaussian, with $E(B_{t_i}) = 0$ for all i and covariance matrix

$$
\begin{bmatrix}
t_1 & t_1 & ... & t_1 \\
t_1 & t_2 & ... & t_2 \\
t_1 & t_2 & ... & t_3 \\
\vdots & \vdots & & \vdots \\
t_1 & t_2 & ... & t_n
\end{bmatrix}
=
\begin{bmatrix}
\min(t_1, t_1) & \min(t_1, t_2) & ... & \min(t_1, t_n) \\
\min(t_2, t_1) & \min(t_2, t_2) & ... & \min(t_2, t_n) \\
\min(t_3, t_1) & \min(t_3, t_2) & ... & \min(t_3, t_n) \\
\vdots & \vdots & & \vdots \\
\min(t_n, t_1) & \min(t_n, t_2) & ... & \min(t_n, t_n)
\end{bmatrix}.
$$

Let $s, t \in [0, \infty)$. Consider

$$
\begin{aligned}
\text{Cov}(B_t, B_s) &= E[(B_t - E[B_t])(B_s - E[B_s])] \\
&= E[B_t B_s] - 2E[B_t]E[B_s] + E\left[E[B_t]E[B_s]\right] \\
&= E[B_t B_s] \quad \text{as} \quad E[B_t] = 0 \quad \text{for all } t.
\end{aligned}
$$

Now, if $t < s$ then $B_s = B_s + B_t - B_t$ and hence

$$\text{Cov}(B_t, B_s) = E[B_t(B_s + B_t - B_t)] = E[B_t^2] + E[B_t(B_s - B_t)].$$

Since $\{B_t\}$ has independent increments of distribution $N(0, s-t)$, we have

$$\text{Cov}(B_t, B_s) = E[B_t^2] + E[B_t]E[B_s - B_t] = E[B_t^2] = t = \min(t, s).$$

This shows that $\{B_t\}_{t \geq 0}$ is a Gaussian process with mean 0 and variance $\min(t, s)$.

$\square$

Lemma E.2. *A Gaussian process* $\{X_t\}_{t\geq 0}$ *with the property* $E[X_t] = 0$ *for all* $t \geq 0$ *and*

$$Cov(X_s, X_t) = \min(s,t) \quad \text{for all } s,t \geq 0,$$

has independent increments. If in addition, the process has continuous paths , then it is a standard Brownian motion on $[0,\infty)$.

Proof. First we can see that for all $s,t \geq 0$;

$$E[X_sX_t] = \text{Cov}(X_s, X_t) + E[X_s]E[X_t] = \text{Cov}(X_s, X_t) = \min(s,t), \tag{E.1}$$

since $E[X_t] = 0$ for all $t \geq 0$. Then for all $0 \leq s \leq t$, $E[X_t - X_s] = E[X_t] - E[X_s] = 0$ and hence

$$\begin{aligned} \text{Var}(X_t - X_s) &= E\left[(X_t - X_s)^2\right] - E[X_t - X_s]^2 = E[(X_t - X_s)^2] \\ &= E[X_t^2] + E[X_s^2] - 2E[X_tX_s] = t + s - 2s = t - s. \end{aligned} \tag{E.2}$$

Next, consider the process of increments $\{X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, ..., X_{t_n} - X_{t_{n-1}}\}$ for each $n > 0$. Note that $\{X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, ..., X_{t_n} - X_{t_{n-1}}\}$ is a multivariate Gaussian by Remark D.1 in Appendix D. Let $i < j \leq n$ we have

$$\begin{aligned} \text{Cov}(X_{t_i} - X_{t_{i-1}}, X_{t_j} - X_{t_{j-1}}) &= E[(X_{t_i} - X_{t_{i-1}})(X_{t_j} - X_{t_{j-1}})] \\ &= E[X_{t_i}X_{t_j}] - E[X_{t_i}X_{t_{j-1}}] - E[X_{t_{i-1}}X_{t_j}] + E[X_{t_{i-1}}X_{t_{j-1}}] \\ &= t_i - t_i - t_{i-1} + t_{i-1} \quad \text{by Equation } (E.1) \\ &= 0 \end{aligned} \tag{E.3}$$

This shows that $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, ..., X_{t_n} - X_{t_{n-1}}$ has a diagonal covariance matrix, then we have by remark D.2, $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, ..., X_{t_n} - X_{t_{n-1}}$ are independent. By equations (E.1), (E.2), (E.3), if it has continuous sample paths, then $\{X_t\}_{t\in\mathbb{R}}$ is a standard Brownian motion on $[0,\infty)$. □

APPENDIX F

FRACTIONAL BROWNIAN FIELD

Definition F.1. (Self similarity)

A random field $\{X_t\}_{t\in\mathbb{R}^d}$ is H-selfsimilar if for any $a > 0$, $X_{at} \overset{d}{=} a^H X_t$; we call H the exponent of self-similarity.

Lemma F.1. *If $\{B_t\}_{t\geq 0}$ is a standard Brownian motion then we have*

1) $\{B_t\}_{t\in\mathbb{R}}$ is a Gaussian process with zero mean and

$$Cov(B_s, B_t) = E[B_s B_t] = min(s,t) \qquad \textit{for all } \ s,t$$

2) $\{B_t\}_{t\in\mathbb{R}}$ is $\frac{1}{2}$-selfsimilar, and

3) $\{B_t\}_{t\in\mathbb{R}}$ has stationary increments.

Proof. 1) See Theorem E.1 in Appendix E.

2) Let $a > 0$. First we show that $\left\{a^{-\frac{1}{2}}B_{at}\right\}$ is a standard Brownian motion.
(B1) sample paths $t \mapsto a^{-\frac{1}{2}}B_{at}$ are continuous a.e. ω as the dilation function is continuous.
(B2) Since the increments $\{B_t - B_s\}$ are independent then $a^{-\frac{1}{2}}B_{at} - a^{-\frac{1}{2}}B_{as}$ are also independent as dilation is a Borel (continuous) function.
(B3) Let $0 \leq s < t$, then $E[a^{-\frac{1}{2}}B_{at} - a^{-\frac{1}{2}}B_{as}] = a^{-\frac{1}{2}}E[B_{at} - B_{as}] = a^{-\frac{1}{2}}(0) = 0$ and $Var(a^{-\frac{1}{2}}B_{at} - a^{-\frac{1}{2}}B_{as}) = \left(a^{-\frac{1}{2}}\right)^2 Var(B_{at} - B_{as}) = a^{-1}(at - as) = t - s$. Thus $a^{-\frac{1}{2}}B_{at} - a^{-\frac{1}{2}}B_{as}$ is $N(0, t-s)$.
(B4) Let $t \geq 0$, then $a^{-\frac{1}{2}}B_0 = a^{-\frac{1}{2}}(0) = 0$
This shows that $\{X_t\} = \left\{a^{-\frac{1}{2}}B_{at}\right\}_{t\in\mathbb{R}}$ is a standard Brownian motion. In particular, the process $\{X_t\}$ is $N(0,t)$ which shows that $X_t \overset{d}{=} B_t$, i.e. $a^{-\frac{1}{2}}B_{at} \overset{d}{=} B_t$, that

is $B_{at} \overset{d}{=} a^{\frac{1}{2}} B_t$.

3) Let $h \in \mathbb{R}$. Then $B_{t+h} - B_h$ is $N(0,t)$ as $\{B_t\}_{t\in\mathbb{R}}$ is a standard Brownian motion. Since B_t is also $N(0,t)$, then $B_{t+h} - B_h$ and B_t has normal distribution with the same mean and variance, we get that they have the same density function. Therefore $B_{t+h} - B_h \overset{d}{=} B_t$. □

Theorem F.2. *A fractional Brownian motion* $\left\{B_t^{\frac{1}{2}}\right\}_{t\geq 0}$ *which has continuous paths and satisfies* $B_0^{\frac{1}{2}} = 0$ *is a standard Brownian motion up to a multiplicative constant.*

Proof. Now for all $s, t \geq 0$ we have $E[B_t^{\frac{1}{2}}] = 0$ and

$$E[B_t^{\frac{1}{2}} B_s^{\frac{1}{2}}] = \tfrac{V_H}{2}\left(|t| + |s| - |t-s|\right) = V_H \min(s,t)$$

then $\dfrac{1}{V_H} E[B_t^{\frac{1}{2}} B_s^{\frac{1}{2}}] = \min(s,t)$. By Lemma E.2 in Appendix E, $\left\{\frac{1}{V_H} B_t^{\frac{1}{2}}\right\}_{t\geq 0}$ is a standard Brownian motion. □

Theorem F.3. *Let* $\left\{B_t^H\right\}_{t\in\mathbb{R}^d}$ *be a fractional Brownian field. Then*

1) $\left\{B_t^H\right\}_{t\in\mathbb{R}^d}$ *has stationary increments.*

2) If $\left\{B_t^H\right\}_{t\in\mathbb{R}^d}$ *has independent increments, then* $H = \frac{1}{2}$.

Proof. 1) Let $h \in \mathbb{R}$, then

$$\begin{aligned}
E\left[B_{t+h}^H - B_h^H)(B_{s+h}^H - B_h^H)\right] &= E\left[(B_{t+h}^H B_{s+h}^H\right] - E\left[B_{t+h}^H B_h^H\right] - E\left[B_h^H B_{s+h}^H\right] \\
&\quad + E\left[B_h^H B_h^H\right] \\
&= \frac{V_H}{2}\left(\|t+h\|^{2H} + \|s+h\|^{2H} - \|t+h-s-h\|^{2H}\right) \\
&\quad - \frac{V_H}{2}\left(\|t+h\|^{2H} + \|h\|^{2H} - \|t+h-h\|^{2H}\right) \\
&\quad - \frac{V_H}{2}\left(\|h\|^{2H} + \|s+h\|^{2H} - \|h-s-h\|^{2H}\right) \\
&\quad + \frac{V_H}{2}\left(\|h\|^{2H} + \|h\|^{2H} - \|h-h\|^{2H}\right) \\
&= \frac{V_H}{2}\left(\|t\|^{2H} + \|s\|^{2H} - \|t-s\|^{2H}\right) \\
&= E[B_t^H B_s^H].
\end{aligned}$$

Now $E[B_{t+h}^H - B_h^H] = E[B_{t+h}^H] - E[B_h^H] = 0-0 = 0 = E[B_t^H]$, and Var $\left(B_{t+h}^H - B_t^H\right) = E[(B_{t+h}^H - B_h^H)^2] - \left(E[B_{t+h}^H - B_h^H]\right)^2 = E[(B_t^H)^2] = Var[B_t^H]$. Note that $\{B_{t+h}^H - B_h^H\}$ is a multivariate Gaussian by Remark D.1 in Appendix D. Then $B_{t+h}^H - B_h^H$ and B_t^H has normal distribution with the same mean and variance, we get that they have the same density function. Therefore $B_{t+h}^H - B_h^H \stackrel{d}{=} B_t^H$.

2) Suppose that $\{B_t^H\}$ has independent increments . Then, for $s, t \in \mathbb{R}^d$,

$$\begin{aligned} 0 = E[B_s^H]E[B_t^H - B_s^H] &= E[B_s^H(B_t^H - B_s^H)] \\ &= E[B_s^H B_t^H] - E[(B_s^h)^2] \\ &= \frac{V_H}{2}\left(\|s\|^{2H} + \|t\|^{2H} - \|t-s\|^{2H} - 2\|s\|^{2H}\right) \\ &= \frac{V_H}{2}\left(\|t\|^{2H} - \|s\|^{2H} - \|t-s\|^{2H}\right). \end{aligned}$$

Then $\|t\|^{2H} - \|s\|^{2H} - \|t-s\|^{2H} = 0$ implies $\|t\|^{2H} - \|s\|^{2H} = \|t-s\|^{2H}$ and hence $H = \frac{1}{2}$. Indeed, if $H \neq \frac{1}{2}$, set $t = 2s$ we have $(2\|s\|)^{2H} - \|s\|^{2H} = \|2s - s\|^{2H}$, then $(2^{2H} - 1)\|s\|^{2H} = \|s\|^{2H}$ so that $2^{2H} = 2$ that is $H = \frac{1}{2}$ which is a contradiction. □

APPENDIX G

BOREL CANTELLI LEMMA

Lemma G.1. *(Borel-Centelli Lemma)*
If $\{A_i\}_{i=1}^{\infty}$ *is any sequence of events and* $\sum_{i=1}^{\infty} P(A_i) < \infty$ *then*

$$P(\bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k) = 0.$$

Proof. Since $\sum_{i=1}^{\infty} P(A_i) < \infty$ we have

$$\begin{aligned} P(\bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k) &= \lim_{m\to\infty} P(\bigcup_{k=m}^{\infty} A_k) \leq \overline{\lim_{m\to\infty}} \sum_{k=m}^{\infty} P(A_k) \\ &= \overline{\lim_{m\to\infty}} \left(\sum_{k=1}^{\infty} P(A_k) - \sum_{k=1}^{m-1} P(A_k) \right) = 0. \end{aligned}$$

□

Note that

$$\begin{aligned} \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k &= \{\omega : \omega \in \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k\} = \{\omega : \forall m \in \mathbb{N}\ \exists k \geq m \text{ such that } \omega \in A_k\} \\ &= \{\omega : \omega \text{ belongs to infinitely many } A_k\text{'s}\}. \end{aligned}$$

Lemma G.2. *If* $\{Z_n\}$ *is a sequence of Gaussian random variables with mean 0 and variance 1, then there is a random variable* C *such that* $|Z_n| \leq C\sqrt{\ln n}$ *a.s. for* $n \geq 2$ *and* $P(C < \infty) = 1$.

Proof. Let $x \geq 1$.
Then $P(|Z_n| \geq x) = \dfrac{2}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{u^2}{2}}\, du \leq \sqrt{\dfrac{2}{\pi}} \int_x^{\infty} u e^{-\frac{u^2}{2}}\, du = e^{-\frac{x^2}{2}} \sqrt{\dfrac{2}{\pi}}.$
Hence for all $\alpha > 1$ and $n \geq 2$,

$$P\left(|Z_n| \geq \sqrt{2\alpha \ln n}\right) \quad \leq \quad e^{-\frac{(2\alpha \ln n)}{2}} \sqrt{\frac{2}{\pi}} \quad = \quad n^{-\alpha} \sqrt{\frac{2}{\pi}}.$$

Now $\sum_{n=2}^{\infty} n^{-\alpha}\sqrt{\frac{2}{\pi}} < \infty$ as $\alpha > 1$. Then by the Borel-Cantelli Lemma (Lemma G.1) we have $P(|Z_n| \geq \sqrt{2\alpha \ln n}$ for infinitely many $n) = 0$.
Setting $C = \sup_{n\geq 2} \frac{|Z_n|}{\sqrt{\ln n}}$ then consider, for some fixed $\alpha > 1$

$$P(C=\infty)=P\left(\sup_{n\geq 2}\frac{|Z_n|}{\sqrt{\ln n}}=\infty\right) \leq P\left(|Z_n| > \sqrt{2\alpha\ln n} \text{ for infinitely many } n\right) = 0.$$

Thus $P(C < \infty) = 1$. □

A modification of this proof yields:

Lemma G.3. *Let $J = \{j = (j_1, ..., j_d) : j_i \in \mathbb{N} \cup \{0\}\}$ and $\mathcal{K}^j = \{k = (k_1, ..., k_d) : k_i = -2^{j_i} + 1, ..., 2^{j_i} - 1\}$. If $\{Z_{j,k}, j \in J$ and $k \in \mathcal{K}^j\}$ is a sequence of Gaussian $N(0, 1)$ random variables, then there exists a random variable C such that $|Z_{_{j,k}}| \leq C\left(\sum_{i=1}^{d}\ln(2^{j_i} + |k_i|)\right)^{\frac{1}{2}}$ a.s. for all $j \in J$ and $k \in \mathcal{K}^J$, and $P(C < \infty) = 1$.*

Proof. Let $x \geq 1$. Then as in the proof of Lemma G.2

$$P\left(|Z_{_{j,k}}| \geq x\right) \leq e^{-\frac{x^2}{2}}\sqrt{\frac{2}{\pi}} \text{ for each } j \in J, k \in \mathcal{K}^j.$$

Hence for each $\alpha > 1$,

$$\begin{aligned} P\left(|Z_{_{j,k}}| \geq \left[2\alpha\sum_{i=1}^{d}\ln(2^{j_i} + |k_i|)\right]^{\frac{1}{2}}\right) &\leq \sqrt{\frac{2}{\pi}}\exp\left(-\alpha\sum_{i=1}^{d}\ln(2^{j_i} + |k_i|)\right) \\ &= \sqrt{\frac{2}{\pi}}\prod_{i=1}^{d}\exp(\ln(2^{j_i} + |k_i|)^{-\alpha}) \\ &= \sqrt{\frac{2}{\pi}}\prod_{i=1}^{d}(2^{j_i} + |k_i|)^{-\alpha}. \end{aligned}$$

Now for each $1 \leq i \leq d$, $\sum_{j_i=0}^{\infty}\sum_{k_i=-2^{j_i}+1}^{2^{j_i}-1}(2^{j_i} + |k_i|)^{-\alpha} < \infty$ as $\alpha > 1$, then

$\sum_{j\in J}\sum_{k\in\mathcal{K}^J}\prod_{i=i}^{d}(2^{j_i} + |k_i|)^{-\alpha} < \infty$, and hence by the Borel-Cantelli Lemma we have

$$P\left(|Z_{_{j,k}}| \geq \left[2\alpha\sum_{i=1}^{d}\ln(2^{j_i} + |k_i|)\right]^{\frac{1}{2}} \text{ for infinitely many } j \text{ and } k\right) = 0.$$

Set $C = \sup\limits_{\substack{j \in J \\ k \in \mathcal{K}^j}} \dfrac{|Z_{j,k}|}{\left[\sum\limits_{i=1}^{d} \ln(2^{j_i} + |k_i|)\right]^{\frac{1}{2}}}$, then

$$P(C = \infty) \leq P\left(\sup_{\substack{j \in J \\ k \in \mathcal{K}^j}} |Z_{j,k}| > [2\alpha \sum_{i=1}^{d} \ln(2^{j_i} + |k_i|)]^{\frac{1}{2}} \text{ for infinitely many } j, k\right) = 0.$$

□

CURRICULUM VITAE

NAME: Kitipol Nualtong. **GENDER:** Male. **NATIONALITY:** Thai.

EDUCATION BACKGROUND:

- Bachelor of Science in Mathematics, Prince of Songkla University, Thailand, 1997 - 2000.

- Master of Science in Mathematics, Chulalongkorn University, Thailand, 2003 - 2004.

WORK EXPERIENCES:

- Lecturer, Department of Mathematics, Prince of Songkla University, Thailand, June 2002 - May 2003, June 2005 - May 2007.

PRESENTATIONS:

- *Continuous wavelet transform of some classes of random fields*, The 16th Annual Meeting in Mathematics, March 10-11, 2011, Faculty of Science, Khon Kaen University, Khon Kaen, Thailand.

SCHOLARSHIP:

- The Ministry of University Affairs of Thailand (MUA), 2007-2010.