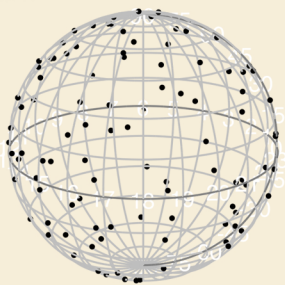


Mathematical Developments on Isotropic Positive Definite Functions on Spheres

Ahmed Arafat Hassan Mohammed

Ph.D. Student



$$\psi(\theta) = \sum_{n=0}^{\infty} b_{n,d} \frac{C_n^{(d-1)/2}(\cos \theta)}{C_n^{(d-1)/2}(1)}$$

$$b_{n,d} = \frac{2n + d - 1}{2^{3-d}\pi} \frac{(\Gamma(\frac{d-1}{2}))^2}{\Gamma(d-1)} \int_0^\pi \psi(\theta) \frac{C_n^{(d-1)/2}(\cos \theta)}{(\sin \theta)^{1-d}} d\theta$$

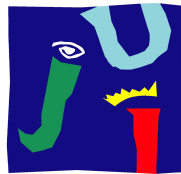


Department of Mathematics
School of Technology and Experimental
Sciences
2017

*Mathematical Developments on Isotropic
Positive Definite Functions on Spheres*

AHMED ARAFAT HASSAN MOHAMMED
PH.D. STUDENT

A DISSERTATION SUBMITTED IN PARTIAL
FULFILLMENT OF
THE REQUIREMENTS FOR THE DOCTOR OF PHILOSOPHY
DEGREE IN MATHEMATICS
IN THE GRADUATE COLLEGE OF
THE UNIVERSITY OF JAUME I



**UNIVERSITAT
JAUME·I**

UNIVERSITY OF JAUME I
SCHOOL OF TECHNOLOGY AND EXPERIMENTAL
SCIENCES
DEPARTMENT OF MATHEMATICS
CASTELLÓN, SPAIN.
MAY 29, 2017

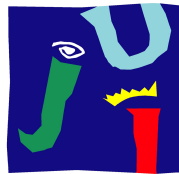
*Mathematical Developments on Isotropic
Positive Definite Functions on Spheres*

AHMED ARAFAT HASSAN MOHAMMED
PH.D. STUDENT

A DISSERTATION SUBMITTED IN PARTIAL
FULFILLMENT OF
THE REQUIREMENTS FOR THE DOCTOR OF PHILOSOPHY
DEGREE IN MATHEMATICS
IN THE GRADUATE COLLEGE OF
THE UNIVERSITY OF JAUME I

ADVISOR
JORGE MATEU
FULL PROFESSOR OF MATHEMATICS

COADVISOR
EMILIO PORCU
FULL PROFESSOR OF MATHEMATICS



**UNIVERSITAT
JAUME·I**

UNIVERSITY OF JAUME I
SCHOOL OF TECHNOLOGY AND EXPERIMENTAL
SCIENCES
DEPARTMENT OF MATHEMATICS
CASTELLÓN, SPAIN.
MAY 29, 2017

Title

Mathematical Developments on Isotropic Positive Definite Functions on Spheres

Abstract: In this thesis, we investigate some problems related to positive definite functions on the sphere. As known, positive definite functions play an important role in representation theory, probability theory, stochastic processes, harmonic analysis and machine learning. This role is emphasized a long the thesis.

First, equivalence of Gaussian measures represents a fundamental tool to establish the properties of maximum likelihood estimators as well as kriging predictions under fixed domain asymptotics. Classical results for the equivalence and orthogonality of measures associated to Gaussian fields on bounded sets of \mathbb{R}^d are available in the literature. The present work considers Gaussian fields defined over spheres of \mathbb{R}^{d+1} , with covariance functions depending on the great circle distance. We provide necessary and sufficient conditions for the equivalence of two Gaussian measures with two different covariance models with associated d -Schoenberg sequences. As an example we study equivalence of Gaussian measures for some parametric families of covariance functions valid on spheres. A simulation study explores the consistency of the maximum likelihood estimator associated to the covariance parameters of some covariance models on the sphere.

Second, we consider the Schoenberg class Ψ_d of continuous functions $\psi : [0, \pi] \rightarrow \mathbb{R}$, with $\psi(0) = 1$ such that the mapping $C(\xi, \eta) = \psi(\theta(\xi, \eta))$, $\xi, \eta \in \mathbb{S}^d$, with θ being the geodesic distance, is positive definite on the product of two d -dimensional spheres \mathbb{S}^d embedded in \mathbb{R}^{d+1} . We face Problems 1 and 3 proposed in the essay by [Gneiting \(2013b\)](#) and related to the d -Schoenberg coefficients in the series expansion of members of Ψ_d for a given d . Such problems have precise implications for the simulation of Gaussian fields on spheres as well as for applications to geostatistical data, with special emphasis to atmospheric sciences. We also show how to deduce the 2-Schoenberg coefficients of given parametric families of members of the class Ψ_d , called respectively exponential and Askey families, for which the 1-coefficients were available but the 2-dimensional case was still elusive.

Third, we propose and define a family of marked point processes in a noncompact semisimple Lie groups. We first generate Lévy processes via marked point processes by using jump-diffusion processes. We then build a family of Markov processes in a maximal compact subgroup of a given semisimple Lie group.

Keywords: Covariance Operator, Gaussian Measures, Lie Groups, Positive Definite Functions, Random Fields, Schoenberg Coefficients, Spherical Harmonics.

Acceptation Note

Thesis Work

Approved

“Laureatus mention”

Jury

Salvador Hernández Muñoz

Jury

Morten Nielsen

Jury

Stefano De Marchi

Advisor

Jorge Mateu

Coadvisor

Emilio Porcu

Castellón, Spain., May 29, 2017

Dedication

I dedicate my thesis work to my family and many friends. A special feeling of gratitude to my loving parents, **Arafat** and **Fayza** whose words of encouragement and push for tenacity ring in my ears. My brothers **Yasser** and **Mohamed** have never left my side and are very special.

I also dedicate this work and give special thanks to my wife **Aml** and my wonderful son **Hamza** for being there for me throughout the entire doctoral program. I miss both of you so much.

I dedicate this work and give special thanks to my best friends **Francisco**, **Jonatan**, **Mehdi**, **Carlos**, **Iulian** and all members of Mathematics department at Jaume I University.

With all my heart, thank to you all.

Acknowledgments

I would like to express my deep gratitude and sincere thanks to my advisor Jorge Mateu, he was very kind with me throughout this work from the start to the end of this work. I feel markedly indebted to him for his continuous stimulation, unlimited help, valuable instructions and very close supervision. He devoted much of his precious time for meticulous guidance that made this review possible.

I would like to express my deep appreciation and gratitude to Emilio Porcu who helped me to get a better understanding of the subject. During this work his continuous support, endless experience and help have done very much to accomplish this work.

Alos, I would like to thank Pablo Gregori for being very supportive and patient, during my PhD studies.

Thanks to you all.

Contents

Contents	III
Introduction	V
1. Background and Preliminaries	1
1.1 Positive Definiteness of Functions	1
1.1.1 Positive Definite Kernels	2
1.1.2 Positive Definite Functions	7
1.1.3 Reproducing Kernel Hilbert Spaces	9
1.2 Random Field Theory	12
1.2.1 Gaussian Random Fields	12
1.2.2 Spectral Representation for a Random Field	15
1.3 Spherical Harmonics and $L^2(\mathbb{S}^d)$ -Space	16
1.3.1 Fourier Representation From \mathbb{R}^d to \mathbb{S}^d	17
1.3.2 Zonal Spherical Harmonics Functions	20
1.3.3 Schoenberg's Coefficients on Spheres	22
2. Equivalence and Orthogonality of Gaussian Measures on Spheres	27
2.1 Some Basic Structures	28
2.2 Necessary and Sufficient Condition on Spheres	32
2.3 Selected Examples	37
2.3.1 Multiquadric Covariance Function	38
2.3.2 The Sine Power Covariance Functions	40
2.4 A Simulation Study	42
2.5 Conclusions	45
3. Gneiting's Problems and the Class Ψ_d of Positive Definite Functions over Hyperspheres	47
3.1 The class Ψ_d and d -Schoenberg coefficients	48
3.2 Gneiting's problems with their solutions	50
3.2.1 Statements of the Problems	50
3.2.2 Main results	52
3.3 On the 2-Schoenberg coefficients	56

3.3.1	Exponential Family	57
3.3.2	Askey Family	59
4.	A family of Markov Processes in Maximal Compact Subgroups of a Semisimple Lie Groups	63
4.1	Introduction	63
4.2	Marked Point Processes and the Multivariate Case	65
4.3	Lévy Processes in Semisimple Lie Groups	72
4.4	Conclusions	76
5.	Conclusions and Future Work	77
5.1	General Conclusions	77
5.2	Ongoing and Future Work	78
5.2.1	Equivalence of Gaussian Measures of Multivariate Random Fields on Sphere	78
5.2.2	When a Function from the Class Ψ_d Belongs to Ψ_{d+1} ?	79
5.2.3	Turán Constant of \mathbb{S}^d	79
	Bibliography	81

Introduction

Isotropic positive definite functions on spheres play a significant role in Mathematics, and they are of particular interest in geostatistics. For example, when modeling large-scale spatial datasets, it might be of interest to work with Gaussian processes at a planetary scale, being the Earth a perfect sphere, and using the geodesic distance instead of the Euclidean one. In this context, isotropic positive definite functions on spheres are important. In this thesis, our primary interest is to study isotropic positive definite functions on spheres and their representation.

There has been a fervent research activity around positive definite functions on spheres in the last five years. The seminal paper by [Gneiting \(2013a\)](#) offers an impressive overview of the problem as well as some connections between mathematical, complex and harmonic analysis, as well as approximation theory, with the theory of stochastic processes, Gaussian random fields, and geostatistics.

Schoenberg's theorem ([Schoenberg, 1942](#), Theorem 2) in concert with the orthonormality properties of spherical harmonics imply that a very natural assumption on positive definite functions over d -dimensional spheres of \mathbb{R}^{d+1} is that they depend on the geodesic (great circle) distance between any two points located over the d -dimensional spherical shell. Such an assumption is known as *geodesic isotropy* and it is the building block for more sophisticated constructions, such

as in [Berg and Porcu \(2016\)](#), [Estrade et al. \(2016\)](#) as well as [Porcu et al. \(2016\)](#). More technical approaches based on complex spheres and locally compact groups have been proposed in [Berg et al. \(2016\)](#).

As an application of the spectral representation of the isotropic positive definite functions on spheres provided by [Schoenberg \(1942\)](#), we studied the equivalence of two Gaussian measures on the sphere. Equivalence and orthogonality of probability measures are useful tools when assessing asymptotic properties of both prediction and estimation for Gaussian fields. For Gaussian fields defined over bounded sets of \mathbb{R}^d and with a given stationary covariance function, the problem has been studied by [Gikhman and Skorokhod \(1966\)](#), [Skorokhod and Yadrenko \(1973\)](#) and [Da Prato and Zabczyk \(1992\)](#). The results obtained in this direction motivated the tours de force in ([Stein, 1988, 1990, 1993, 1999a, 2004](#)), where conditions under which predictions under a misspecified covariance function are asymptotically efficient are provided.

A simple condition for the equivalence of two Gaussian measures, given in [Stein \(2004\)](#), and based on the celebrated [Skorokhod and Yadrenko \(1973\)](#) results, is the following: two stationary and isotropic (radially symmetric) covariance functions C_i , with associated spectral densities f_i , $i = 1, 2$, are compatible if $f_2(z)z^a$, $a > 0$ is bounded away from 0 and ∞ as $z \rightarrow \infty$ and

$$\int_r^\infty z^{d-1} \left(\frac{f_1(z) - f_2(z)}{f_2(z)} \right)^2 dz < \infty, \quad (1)$$

for any $r > 0$. Here, isotropy is meant in the sense that the covariance function depends only on the Euclidean distance. The argument above is the key for evaluating equivalence of Gaussian measures with all the implications in terms of estimation and prediction under fixed domain asymptotics. [Zhang \(2004\)](#), using (1), showed conditions for the compatibility of two Matérn covariance functions with different variance and scale parameters, and with the same (and fixed) level of smoothness. Such a result has huge implications in terms of consistency of the maximum likelihood estimators associated to the covariance parameters. Specif-

ically, in the Matérn case not all the parameters can be estimated consistently but the so-called microergodic parameter (Zhang, 2004) can be estimated consistently. Similar results can be found in Bevilacqua et al. (2016) when using the Generalized Wendland covariance function.

Stationary Gaussian fields on spheres (Marinucci and Peccati, 2011) have covariance functions that depend naturally on the great circle distance. Their different mathematical structure has been described in (Gneiting, 2013a), and their extension to space-time has been achieved in (Berg and Porcu, 2016) as well as in (Porcu et al., 2016).

Fixed domain asymptotics is a natural asymptotics framework for Gaussian fields defined on spheres when considering both estimation and prediction. The lack of results for establishing the equivalence of Gaussian measures defined on the sphere motivates the present thesis: we give necessary and sufficient conditions based on the d -Schoenberg sequences associated to covariance functions depending on the great circle distance. We find that two Gaussian measures with these covariance models are never compatible except for the trivial case. A simulation study explores the consistency of the maximum likelihood estimator associated to the covariance parameters of Multiquadric and Sine power covariance models. In the simulation study, we also consider the exponential model using the geodesic distance.

The plan of the thesis is the following. In the first chapter, we present the main concepts needed for results which we show in subsequent chapters, such as positive definite kernels and spherical harmonics. We focus on the Schoenberg's representation of the isotropic positive definite functions on spheres.

The second chapter is devoted to the main results related to the necessary and sufficient condition for the equivalence relation between two Gaussian measures on the sphere. We then apply the results presented in that chapter to study the equivalence or orthogonality of some parametric families of covariance func-

tions on the sphere. Specifically, we focus on the Multiquadric and Sine power covariance models (Gneiting, 2013a).

In the third chapter we address two important open problems. The former being related to the representation of the d -Schoenberg's coefficients in terms of 1-Schoenberg coefficients. Such a problem is parenthetical to the celebrated Matheron's turning bands operator (Matheron, 1963) proposed in Euclidean space only. The latter problem finds instead motivation in atmospheric data assimilation, where locally supported isotropic correlation functions are used for the distance-dependent reduction of global scale covariance estimates in ensemble Kalman filter settings (Buehner and Charron, 2007; Hamill et al., 2001).

In the fourth chapter, we derive some results related to stochastic processes on the Lie groups. Finally, in the last chapter we give a general conclusion and the ongoing work.

CHAPTER 1

Background and Preliminaries

In this chapter, we present the main concepts that are needed to provide our results which we show in subsequent chapters. All these results are standard, and we provide a motivation and references when necessary.

1.1 Positive Definiteness of Functions

The study of positive definiteness (kernels or functions) in Mathematical analysis began with the work of [Mathias \(1923\)](#). The fundamental point of the characterization of the positive definiteness regarding Fourier transforms appeared a few decades later by [Bochner \(1948\)](#) and [Schoenberg \(1938\)](#). So we will begin with a brief introduction for positive definite kernels and then for positive definite functions.

1.1.1 Positive Definite Kernels

A positive definite kernel is a generalization of a positive definite function or a positive definite matrix. We will start with the definition of the positive definite kernel.

Definition 1.1.1. *Suppose a nonempty set \mathcal{X} and the mapping $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$. The mapping K is said to be positive definite kernel if for every $n \in \mathbb{N}$ and for every n -tuple (x_1, \dots, x_n) of elements from \mathcal{X} the matrix $[K(x_i, x_j)]_{i,j=1}^n$ is positive definite in the sense that*

$$\sum_{i,j=1}^n K(x_i, x_j) \alpha_i \bar{\alpha}_j \geq 0,$$

for all complex numbers $\alpha_1, \dots, \alpha_n$.

We consider $K(x_i, x_j)$ for $i, j = 1, \dots, n$ as entries of $n \times n$ matrix. Note that if the square matrix $A = [a_{ij}]$ with $a_{ij} \in \mathbb{C}$ is positive definite, then A is Hermitian (i.e., $a_{ij} = \bar{a}_{ji}$) with non-negative eigenvalues, see (Johnson, 1970). For example, Assume that $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ is positive definite kernel. Then for any $x_1, x_2, x_3 \in \mathcal{X}$ the matrix

$$\begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) & K(x_1, x_3) \\ K(x_2, x_1) & K(x_2, x_2) & K(x_2, x_3) \\ K(x_3, x_1) & K(x_3, x_2) & K(x_3, x_3) \end{bmatrix}$$

is positive definite (i.e., Hermitian with non-negative eigenvalues). Thus

$$K(x_i, x_j) = \overline{K(x_j, x_i)}, \quad \text{for all } i, j = 1, 2, 3, \quad (1.1)$$

and the equation

$$\det \left([K(x_i, x_j)]_{i,j=1}^3 - \lambda I \right) = 0 \quad (1.2)$$

has non-negative solutions in λ , where I is the 3×3 identity matrix.

The following proposition deals with operations that preserve positive definiteness.

Proposition 1.1.2. *Let $K_i : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$, $i = 1, 2, \dots$ be positive definite kernels. Then the following properties are hold (Berg et al., 1984).*

1. *The positive combination $\sum_i a_i K_j$ is also positive definite for all $a_i \geq 0$, $i = 1, 2, \dots$*
2. *The product of two positive definite kernels is also positive definite.*
3. *If the limit*

$$K(x, y) = \lim_{i \rightarrow \infty} K_i(x, y)$$

exists, then $K(x, y)$ is positive definite kernel, too.

From Proposition 1.1.2, we see that the set of all positive definite kernels is closed under point-wise convergence and convex cone stable under multiplication (Schoenberg, 1942). Furthermore, if $K(x, x) > 0$ for all $x \in \mathcal{X}$, then the normalization of $K(x, y)$

$$\frac{K(x, y)}{\sqrt{K(x, x) K(y, y)}}$$

is a positive definite kernel, see (Rasmussen, 2006).

Example 1.1.3. *Consider $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ a positive definite kernel. Then by Proposition 1.1.2 we have*

$$\exp(K(x, y)) = 1 + \frac{1}{1!} K(x, y) + \frac{1}{2!} (K(x, y))^2 + \frac{1}{3!} (K(x, y))^3 + \dots$$

is a positive definite kernel.

We can define a positive definite kernel by using any arbitrary function on \mathcal{X} , and this is noted in the following proposition.

Proposition 1.1.4. *Suppose $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ is a positive definite kernel and $f : \mathcal{X} \rightarrow \mathbb{C}$ a function. Then,*

$$\widetilde{K}(x, y) = f(x)K(x, y)\overline{f(y)} \tag{1.3}$$

is a positive definite kernel.

Indeed, let $x_1, \dots, x_n \in \mathcal{X}$ and let $\alpha_1, \dots, \alpha_n \in \mathbb{C}$. We have

$$\begin{aligned} \sum_{i,j=1}^n \widetilde{K}(x_i, x_j) \alpha_i \overline{\alpha_j} &= \sum_{i,j=1}^n f(x_i) K(x_i, x_j) \overline{f(x_j)} \alpha_i \overline{\alpha_j} \\ &= \sum_{i,j=1}^n K(x_i, x_j) \alpha_i \overline{f(x_j)} \overline{f(x_i)} \alpha_j \geq 0. \end{aligned}$$

In particular, $K(x, y) = f(x) \overline{f(y)}$ is a positive definite kernel. Hence,

$$\begin{aligned} \sum_{i,j=1}^n K(x_i, x_j) \alpha_i \overline{\alpha_j} &= \sum_{i,j=1}^n f(x_i) \overline{f(x_j)} \alpha_i \overline{\alpha_j} \\ &= \left| \sum_{i=1}^n f(x_i) \alpha_i \right|^2 \geq 0. \end{aligned}$$

Proposition 1.1.5. *Let $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ be a positive definite kernel such that $|K(x, y)| < \epsilon$ for all $x, y \in \mathcal{X}$, $\epsilon > 0$. If the radius of convergence of the power series $f(\xi) = \sum_{i=1}^{\infty} a_i \xi^i$ with non-negative coefficients is $\epsilon > 0$, then the composition $f \circ K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ is a positive definite kernel.*

Example 1.1.3 is a particular case of Proposition 1.1.5, since the analytic function in this case is the exponential $f(x) = \exp(x)$. Definition 1.1.1 can be formulated for a real-valued positive definite kernel. Then we have the following definition.

Definition 1.1.6. (*Berg et al., 1984*) *Suppose \mathcal{X} is a nonempty set and let K be a symmetric mapping $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ (i.e., $K(x, y) = K(y, x)$ for all $x, y \in \mathcal{X}$). We say that K is a positive definite kernel if and only if it is symmetric, and the following inequality*

$$\sum_{i,j=1}^n K(x_i, x_j) \alpha_i \alpha_j \geq 0 \tag{1.4}$$

is hold for any $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ and for any $x_1, \dots, x_n \in \mathcal{X}$.

Remark. *Let \mathcal{X} be a nonempty set and \mathcal{Y} is a subset of \mathcal{X} . If $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ is positive definite kernel, then the restriction of K to $\mathcal{Y} \times \mathcal{Y}$ is a positive definite kernel.*

Let X be a set and $\mathcal{B}(X)$ a σ -algebra over X . A function μ from $\mathcal{B}(X)$ to \mathbb{R}_+ is called a measure if for all $E \in \mathcal{B}(X)$, $\mu(E) \geq 0$, $\mu(\emptyset) = 0$ and for all countable

collections of pairwise disjoint sets $\{E_i\}_{i=1}^{\infty}$ in $\mathcal{B}(X)$:

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i).$$

Then the triple $(X, \mathcal{B}(X), \mu)$ is said to be measurable space.

Example 1.1.7. *Suppose \mathcal{X} is a nonempty set and $(X, \mathcal{B}(X), \mu)$ is a measurable space. Then*

1. *Let $\Lambda(X) := \{K_{\omega} : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R} : \omega \in X\}$ i.e., $K_{\omega}(x, y) \in L^1(X)$. Then*

$$K(x, y) = \int_X K_{\omega}(x, y) \, d\mu(\omega),$$

where $x, y \in \mathcal{X}$, is positive definite kernel.

2. *Let $k : \mathcal{X} \times X \rightarrow \mathbb{R}$ be a function such that for any $x \in \mathcal{X}$ the function $k(x, \cdot) : X \rightarrow \mathbb{R}$ is a member of the space $L^2(d\mu)$. Then the function*

$$K(x, y) = \int_X k(x, s) \overline{k(y, s)} \, d\mu(s)$$

is a positive definite kernel.

Let \mathcal{H} be a Hilbert space of real-valued function on some nonempty sets \mathcal{X} (i.e., $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$). We denote the inner product on \mathcal{H} by $\langle f, g \rangle_{\mathcal{H}}$ and the associated norm by $\|f\|^2 = \langle f, f \rangle_{\mathcal{H}}$ for all $f, g \in \mathcal{H}$. The sets that admit a kernel can be embedded in a Hilbert space, this is the so-called Moore-Aronszajn theorem, see (Lloyd, 2003).

Theorem 1.1.8 (Moore-Aronszajn Theorem). *Let \mathcal{X} be a nonempty set, the mapping $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ is a positive definite kernel if and only if there exists a Hilbert space \mathcal{H} and a function $\varphi : \mathcal{X} \rightarrow \mathcal{H}$ (so-called feature map) such that $K(x, y) = \langle \varphi(x), \varphi(y) \rangle_{\mathcal{H}}$ for all $x, y \in \mathcal{X}$.*

Proof. Let $K(x, y)$ be a positive definite kernel. Define $\varphi : \mathcal{X} \rightarrow \mathbb{R}^{\mathcal{X}}$ by $\varphi(x) = K(x, \cdot)$, for all $x \in \mathcal{X}$. Let $\mathcal{K} \subset \mathbb{R}^{\mathcal{X}}$ be the space of all linear combinations of functions of the form $\varphi(x)$ for some x . Following Proposition 1.1.5 for every

$f = \sum_{i=1}^n \alpha_i K(x_i, \cdot)$ and $g = \sum_{j=1}^m \beta_j K(y_j, \cdot)$, we can define the inner product on \mathcal{K} by

$$\langle f, g \rangle_{\mathcal{K}} = \sum_{i=1}^n \sum_{j=1}^m \alpha_i \beta_j K(x_i, y_j).$$

Since K is a positive definite kernel, then $\langle f, g \rangle_{\mathcal{K}} = \langle g, f \rangle_{\mathcal{K}}$ and

$$\|f\|_{\mathcal{K}} = \langle f, f \rangle_{\mathcal{K}} = \sum_{i,j=1}^n \alpha_i \alpha_j K(x_i, x_j) \geq 0.$$

Moreover, $K(x, y) = \langle K(x, \cdot), K(y, \cdot) \rangle_{\mathcal{K}} = \langle \varphi(x), \varphi(y) \rangle_{\mathcal{H}}$, for all $x, y \in \mathcal{X}$. And the Hilbert space \mathcal{H} is the completion of \mathcal{K} .

Conversely, let $x, y \in \mathcal{X}$. Since $K(x, y) = \langle \varphi(x), \varphi(y) \rangle_{\mathcal{H}}$, by the properties of the inner product function, then $K(x, y) = K(y, x)$. Thus K is symmetric. Let $n \in \mathbb{N}$, $x_1, \dots, x_n \in \mathcal{X}$ and $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. Then

$$\sum_{i,j=1}^n K(x_i, x_j) \alpha_i \overline{\alpha_j} = \sum_{i,j=1}^n \langle \varphi(x_i), \varphi(x_j) \rangle_{\mathcal{H}} \alpha_i \overline{\alpha_j} = \left\| \sum_{i=1}^n \alpha_i \varphi(x_i) \right\|_{\mathcal{H}}^2 \geq 0.$$

Thus K is a positive definite kernel. □

Let $\{e_i, i \in \mathcal{I}\}$ be an orthonormal basis of the Hilbert space \mathcal{H} , where \mathcal{I} is an index set. Then for every $x, y \in \mathcal{X}$

$$K(x, y) = \sum_{i \in \mathcal{I}} \langle \varphi(x), e_i \rangle_{\mathcal{H}} \overline{\langle \varphi(y), e_i \rangle_{\mathcal{H}}}.$$

Therefore, if we define $\varphi_i : \mathcal{X} \rightarrow \mathbb{R}$ by $\varphi_i(x) = \langle \varphi(x), e_i \rangle_{\mathcal{H}}$, then we can rewrite the positive definite kernel as $K(x, y) = \sum_{i \in \mathcal{I}} \varphi_i(x) \overline{\varphi_i(y)}$, for all $x, y \in \mathcal{X}$, see (Berg et al., 1984; Lloyd, 2003; Lang and Schwab, 2015).

Example 1.1.9. Consider $\mathcal{X} = \mathbb{R}$ and $K(x, y) = \frac{xy}{2}$. Then $K(x, y)$ is a positive definite kernel, since according to Theorem 1.1.8, the Hilbert space could be $\mathcal{H} = \mathbb{R}$ with $\varphi(x) = \frac{x}{\sqrt{2}}$. Also, we can redefine $K(x, y) = \left(\frac{x}{2}, \frac{x}{2}\right) \left(\frac{y}{2}, \frac{y}{2}\right)^T$, in this case the Hilbert space will be $\mathcal{H} = \mathbb{R}^2$ with $\varphi(x) = \left(\frac{x}{2}, \frac{x}{2}\right)$, where $(\cdot)^T$ refers to matrix transpose.

1.1.2 Positive Definite Functions

Let (\mathcal{S}, \star) be an abelian semigroup equipped with a binary operation \star , an identity element (denoted by e), and involution $*$: $\mathcal{S} \rightarrow \mathcal{S}$ such that $(s \star t)^* = s^* \star t^*$ and $(s^*)^* = s$ for all $s, t \in \mathcal{S}$, see (Ressel and Ricker, 1998). We will denote the abelian semigroup with identity e together with the involution $*$ by $*$ -semigroup.

Note that, if (\mathcal{G}, \star) is an abelian group, then we can consider two different involutions $*$: $g \mapsto g^{-1}$ and the identical involution $*$: $g \mapsto g$ for all $g \in \mathcal{G}$, see (Berg et al., 1984). Now, we give the definition of the positive definite function on an abelian $*$ -semigroup (\mathcal{S}, \star) with e .

Definition 1.1.10. *Let \mathcal{S} be a $*$ -semigroup. A function $f : \mathcal{S} \rightarrow \mathbb{C}$ is said to be a positive definite function if $K : (s, t) \mapsto f(s \star t^*)$ is a positive definite kernel on $\mathcal{S} \times \mathcal{S}$, i.e., if*

$$\sum_{i,j=1}^n f(s_i \star s_j^*) \alpha_i \bar{\alpha}_j \geq 0,$$

for every n -tuple (s_1, \dots, s_n) of members from \mathcal{S} and for every $\alpha_1, \dots, \alpha_n \in \mathbb{C}$.

Thus, a positive definite function comes in conjunction with the existence of positive definite kernel $K(s, t)$ such that

$$K(s, t) = f(s \star t^*), \quad \text{for all } s, t \in \mathcal{S}. \quad (1.5)$$

Furthermore, if f is a positive definite function on a $*$ -semigroup \mathcal{S} , then Propositions 1.1.2, 1.1.5 and Theorem 1.1.8 are hold for f .

Example 1.1.11. *Let $(\mathcal{G}, +)$ be an abelian group and G be a subset of \mathcal{G} . Then the involution map is $*$: $g \mapsto -g$, moreover $G - G$ is a symmetric subset. Hence, the function $f : G \rightarrow \mathbb{C}$ is positive definite on G if $K(g, h) = f(g - h)$ is a positive definite kernel for all $g, h \in G$.*

The next example is a particular case of Example 1.1.11.

Example 1.1.12. *Assume that $(\mathcal{G}, +) = (\mathbb{R}^d, +)$ and $f : \mathbb{R}^d - \mathbb{R}^d \rightarrow \mathbb{C}$ is defined by $f(\mathbf{x}) = \exp(i\langle \mathbf{x}, \omega \rangle)$. Then f is positive definite for any fixed $\omega \in \mathbb{R}^d$. This is*

due to the inequality in Definition 1.1.10 as

$$\begin{aligned}
\sum_{i,j=1}^n f(\mathbf{x}_i - \mathbf{x}_j) \alpha_i \bar{\alpha}_j &= \sum_{i,j=1}^n \exp(i\langle \mathbf{x}_i - \mathbf{x}_j, \omega \rangle) \alpha_i \bar{\alpha}_j \\
&= \sum_{i,j=1}^n \exp(i\langle \mathbf{x}_i, \omega \rangle) \exp(i\langle -\mathbf{x}_j, \omega \rangle) \alpha_i \bar{\alpha}_j \\
&= \sum_{i,j=1}^n \exp(i\langle \mathbf{x}_i, \omega \rangle) \overline{\exp(i\langle \mathbf{x}_j, \omega \rangle)} \alpha_i \bar{\alpha}_j \\
&= \left| \sum_{i=1}^n \exp(i\langle \mathbf{x}_i, \omega \rangle) \alpha_i \right|^2 \geq 0,
\end{aligned}$$

for every n -tuple $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ of members from \mathbb{R}^d and for every $\alpha_1, \dots, \alpha_n \in \mathbb{C}$.

The following theorem is one of the most important results of positive definite functions on locally compact abelian groups in general and on \mathbb{R}^d in particular, that is, the so-called Bochner's theorem, see (Adler and Taylor, 2007; Da Prato and Zabczyk, 1992; Rudin, 2011; Heyer, 2012). Let $\mathcal{B}(\mathbb{R}^d)$ denote the Borel σ -algebra on \mathbb{R}^d .

Theorem 1.1.13. *Let $f : \mathbb{R}^d \rightarrow \mathbb{C}$ be a bounded continuous function. Then, f is positive definite if and only if there exists a positive and finite Borel measure μ on $\mathcal{B}(\mathbb{R}^d)$ such that*

$$f(\omega) = \int_{\mathbb{R}^d} \exp(i\langle \omega, \mathbf{x} \rangle) d\mu(\mathbf{x}), \quad (1.6)$$

for all $\omega \in \mathbb{R}^d$.

Thus, Bochner theorem state that any continuous positive definite function on \mathbb{R}^d is a Fourier transform of some finite Borel measure $\mu \in \mathcal{B}(\mathbb{R}^d)$. To show that the function $f(\omega)$ in (1.6) is positive definite function, let $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$ and $\alpha_1, \dots, \alpha_n \in \mathbb{C}$. Then

$$\begin{aligned}
\sum_{i,j=1}^n f(\mathbf{x}_i - \mathbf{x}_j) \alpha_i \bar{\alpha}_j &= \sum_{i,j=1}^n \alpha_i \bar{\alpha}_j \int_{\mathbb{R}^d} \exp(i\langle \mathbf{x}_i - \mathbf{x}_j, \mathbf{x} \rangle) d\mu(\mathbf{x}) \\
&= \int_{\mathbb{R}^d} \sum_{i,j=1}^n \exp(i\langle \mathbf{x}_i, \mathbf{x} \rangle) \exp(i\langle -\mathbf{x}_j, \mathbf{x} \rangle) \alpha_i \bar{\alpha}_j d\mu(\mathbf{x}) \\
&= \int_{\mathbb{R}^d} \left| \sum_{i=1}^n \exp(i\langle \mathbf{x}_i, \mathbf{x} \rangle) \alpha_i \right|^2 d\mu(\mathbf{x}) \geq 0.
\end{aligned}$$

In fact, the importance of Bochner's theorem (Theorem 1.1.13) is that it is providing the characterization of the positive definite function on the Euclidean space. In Section 1.3 below, we will give the equivalent theorem which characterizes positive definite functions on the d -dimensional sphere, the so-called Schoenberg's theorem.

1.1.3 Reproducing Kernel Hilbert Spaces

A reproducing kernel Hilbert space is a Hilbert space of functions \mathcal{H} in addition to all evaluation operators defined on \mathcal{H} being bounded and linear, see (Berlinet and Thomas-Agnan, 2011). The first one who presented the concept of reproducing kernels (for short, RK) is Barnard and Moore (1935), whereas they referred to RK as positive Hermitian matrices. Aronszajn (1950) defined the reproducing kernel Hilbert space (for short, RKHS). In this section, we shall describe in general what is the meaning of a reproducing kernel Hilbert space and its positive definite kernel.

Let \mathcal{X} be a nonempty set and \mathcal{H} be a Hilbert space of real-valued functions $f : \mathcal{X} \rightarrow \mathbb{R}$ with associated inner product and norm

$$\langle f, g \rangle_{\mathcal{H}} = \int_{\mathcal{X}} f(x) g(x) d\mu(x), \quad (1.7)$$

$$\|f\|_{\mathcal{H}} = \left(\int_{\mathcal{X}} |f(x)|^2 d\mu(x) \right)^{1/2}, \quad (1.8)$$

for all $f, g \in \mathcal{H}$, where μ is a positive and finite measure on the Borel σ -algebra $\mathcal{B}(\mathcal{X})$. We say that the Hilbert space \mathcal{H} is $L^2(\mathcal{X}, \mu)$ (so-called squared-integrable space), if $\|f\|^2 < \infty$, see (Dai and Xu, 2013).

The evaluation operator on a Hilbert space \mathcal{H} is a mapping δ_x that evaluates each function in \mathcal{H} at the point $x \in \mathcal{X}$, i.e., $\delta_x : f \mapsto f(x)$. It is clear that the evaluation operator is linear in the sense that for any $\alpha, \beta \in \mathbb{R}$, $\delta_x[\alpha f + \beta g] = \alpha \delta_x[f] + \beta \delta_x[g]$.

We say that δ_x is bounded evaluation operator, if there exists $M > 0$ such that

$$|\delta_x[f]| = |f(x)| \leq M \|f\|_{\mathcal{H}},$$

for all $x \in \mathcal{X}$ and for all $f \in \mathcal{H}$, see (Steinwart and Christmann, 2008). Now, we can define the reproducing kernel Hilbert space as follows.

Definition 1.1.14. *A Hilbert space \mathcal{H} defined above with associated inner product (1.7) and the norm (1.8) is said to be a reproducing kernel Hilbert space, if every evaluation operator δ_x is bounded on \mathcal{H} for all $x \in \mathcal{X}$.*

Since the bounded operator is continuous, then the Hilbert space \mathcal{H} is an RKHS if the evaluation operator is continuous for all $x \in \mathcal{X}$. Furthermore, the RKHS has an interesting property that is, if $f, g \in \mathcal{H}$ are closed to each other, then their evaluation $f(x)$ and $g(x)$ are closed for all $x \in \mathcal{X}$.

Let \mathcal{H} be a Hilbert space, in general the convergence with respect to $\|\cdot\|_{\mathcal{H}}$ does not imply the pointwise convergence. But if \mathcal{H} is an RKHS, then the convergence with respect to $\|\cdot\|_{\mathcal{H}}$ does not necessarily implies pointwise convergence, i.e., let $\{f_n\}_{n=0}^{\infty}$ be a sequence of functions in \mathcal{H} , then $\lim_{n \rightarrow \infty} \|f_n - f\| = 0$ imply that $f_n \rightarrow f$ as $n \rightarrow \infty$, this is because of the continuity of the evaluation operator. To show the meaning of the kernel in Definition 1.1.14 we need to define the reproducing kernel (RK), see (Berlinet and Thomas-Agnan, 2011).

Definition 1.1.15. *Let \mathcal{H} be a Hilbert space defined as above with associated inner product (1.7) and the norm (1.8). We say that the kernel $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ is a reproducing kernel if for all $x \in \mathcal{X}$, $f \in \mathcal{H}$, then $K(x, \cdot) \in \mathcal{H}$ and $\langle f, K(x, \cdot) \rangle_{\mathcal{H}} = f(x)$.*

Thus, the reproducing kernel acts as evaluation operator for all $f \in \mathcal{H}$. This coincides with what we have said above in Theorem 1.1.8, where for all $x, y \in \mathcal{X}$, we can define $K(x, y) = \langle K(x, \cdot), K(y, \cdot) \rangle_{\mathcal{H}}$. Using Riesz representation theorem, we have the following theorem.

Theorem 1.1.16. (*Sejdinovic and Gretton, 2012*) *The Hilbert space \mathcal{H} of real-valued functions is RKHS if and only if \mathcal{H} has a unique reproducing kernel.*

Note that, the Hilbert space in Theorem 1.1.8 is not necessarily a reproducing kernel Hilbert space. Also, it is clear that the reproducing Kernel is the positive definite kernel. If \mathcal{H} is a RKHS, then the subset $\text{span}\{\delta_x : x \in \mathcal{X}\}$ is dense in \mathcal{H} .

Example 1.1.17. (*Pesenson, 1999*) *The set of band limited functions $\Xi(\mathbb{R})$ is the set of all square-integrable functions on \mathbb{R} such that the Fourier transform defined by*

$$\mathcal{F}[f](\xi) = \int_{-\infty}^{\infty} \exp(-ix\xi) f(x) dx,$$

for any $\xi \in \mathbb{R}$ has a compact support. Then we can define the space of band-limited square-integrable functions on \mathbb{R} as follows

$$\Xi(\mathbb{R}) := \left\{ f \in L^2(\mathbb{R}) : \text{Supp}\mathcal{F}[f] \subset [0, 1] \right\}. \quad (1.9)$$

It is clear that the space $\Xi(\mathbb{R})$ is a Hilbert space. To prove that $\Xi(\mathbb{R})$ is a RKHS, we consider the positive definite kernel

$$K(x, y) = \frac{\sin(x - y)}{\pi(x - y)},$$

from $\mathbb{R} \times \mathbb{R}$ to \mathbb{R} . So we have to show that $K(x, y)$ is a reproducing kernel, i.e., $\langle f, K(x, \cdot) \rangle_{\mathcal{H}} = f(x)$ as follows,

$$\begin{aligned} \langle f, K(x, \cdot) \rangle_{\mathbb{R}} &= \int_{-\infty}^{\infty} f(y) \frac{\sin(y - x)}{\pi(y - x)} dy \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) \int_{-1}^1 \exp(-i(y - x)\xi) d\xi dy \\ &= \frac{1}{2\pi} \int_{-1}^1 \int_{-\infty}^{\infty} f(y) \exp(-iy\xi) \exp(ix\xi) dy d\xi \\ &= \frac{1}{2\pi} \int_{-1}^1 \exp(ix\xi) \mathcal{F}[f](\xi) d\xi = f(x). \end{aligned}$$

Thus, the positive definite kernel $K(x, y)$ is a reproducing kernel. Moreover, the space $\Xi(\mathbb{R})$ is RKHS.

Consider that \mathcal{H} is a RKHS with a function $f \in \mathcal{H}$ such that $f(x) \neq f(y)$ for all $x, y \in \mathcal{X}$. Then we can define a distance function $d : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}_+$ by

$$\begin{aligned} d(x, y) &= \inf \{ |f(x) - f(y)| : \|f\|_{\mathcal{H}} \leq 1 \}, \\ &= \inf \{ \langle f, K(x, \cdot) - K(y, \cdot) \rangle_{\mathcal{H}} : \|f\|_{\mathcal{H}} \leq 1 \}. \end{aligned}$$

Thus, the distance satisfies $d(x, y) \leq \|K(x, \cdot) - K(y, \cdot)\|_{\mathcal{H}}$. Let \mathcal{X} be a nonempty set. Then the main conclusion of this section is, by the positive definite kernel $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ (function) we can construct a reproducing kernel Hilbert space \mathcal{H} , hence we can define a distance function $d(x, y)$.

1.2 Random Field Theory

The intention of this section is to set up a framework that contains the important concepts which we need to present the results. Some of the main references in this section are [Adler and Taylor \(2007\)](#); [Preston \(1976\)](#); [Da Prato and Zabczyk \(1992\)](#) and [Spodarev \(2013\)](#). Let Ω be a nonempty set, $\mathcal{B}(\Omega)$ a σ -algebra on Ω and \mathbb{P} is a probability measure on $\mathcal{B}(\Omega)$, i.e., $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$, for all $B \in \mathcal{B}(\Omega)$ then $0 \leq \mathbb{P}(B) \leq 1$ and for the collection $\{B_i\}_{i=1}^{\infty}$ of pairwise disjoint sets in $\mathcal{B}(\Omega)$

$$\mathbb{P} \left(\bigcup_{i=1}^{\infty} B_i \right) = \sum_{i=1}^{\infty} \mathbb{P}(B_i).$$

Then the triple $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ is said to be a probability space.

1.2.1 Gaussian Random Fields

Let $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ be a probability space and $(X, \mathcal{B}(X), \mu)$ be a measurable space.

Let us start with fundamental definitions for random fields, in general.

Definition 1.2.1. (*Da Prato and Zabczyk, 1992*) A random variable $Z : \Omega \rightarrow X$ is a measurable mapping from $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ to $(X, \mathcal{B}(X), \mu)$, i.e.,

$$Z^{-1}(E) = \{\omega \in \Omega : Z(\omega) \in E\} \in \mathcal{B}(\Omega) \quad \text{for all } E \in \mathcal{B}(X). \quad (1.10)$$

The random field (random function or stochastic process) on a probability space $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ is the family $Z = \{Z(t) : t \in \mathcal{T}\}$, where \mathcal{T} is an index space (for example $\mathcal{T} = \mathbb{R}_+, \mathbb{R}$). In short, the random field $Z(t)$ is a function whose values are random variables for any $\omega \in \Omega$. The family $\{Z(\omega, t) : t \in \mathcal{T}\}$ for fixed $\omega \in \Omega$ is called a trajectory (realization) of Z , see ([Preston, 1976](#)).

Note that if $X = \mathbb{R}$, then we call Z is a univariate random field, and if $X = \mathbb{R}^d$, then we call it multivariate random field.

Definition 1.2.2. We define \mathbb{P}_Z as the distribution of a random variable $Z(\omega)$ as a probability measure on the measurable space $(X, \mathcal{B}(X), \mu)$, i.e., $\mathbb{P}_Z(E) = \mathbb{P}_Z(Z^{-1}(E))$, for $E \in \mathcal{B}(X)$.

The next definition to define the mean value and covariance functions for a given random field Z on a probability space $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$.

Definition 1.2.3. Let $Z = \{Z(t) : t \in \mathcal{T}\}$ be a random field, the mean function of Z is given by

$$m(t) := \mathbb{E}[Z(t)] \quad (1.11)$$

for all $t \in \mathcal{T}$, where $\mathbb{E}[\cdot]$ is the expectation. And the covariance function of the random field Z is defined by

$$C(t, s) = \text{cov}(Z(t), Z(s)) = \mathbb{E}[Z(t) Z(s)] - \mathbb{E}[Z(t)]\mathbb{E}[Z(s)], \quad (1.12)$$

for all $t, s \in \mathcal{T}$.

We say that the covariance function is centered if the last term in (1.12) is equal to zero, otherwise it is called a non-centered covariance function. The variance of the random field Z is given by $C(t, t)$ for all $t \in \mathcal{T}$. It is clear that the covariance

function is a symmetric positive definite i.e., it holds that

$$\sum_{i,j=1}^n C(t_i, t_j) \alpha_i \bar{\alpha}_j \geq 0,$$

for all $n \in \mathbb{N}$, n -tuple (t_1, t_2, \dots, t_n) of members from \mathcal{T} and $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{C}$.

Also, the correlation function of the random field is defined by

$$R(t, s) = \frac{C(t, s)}{\sqrt{C(t, t)C(s, s)}}.$$

It is clear that by the Cauchy-Schwarz inequality we conclude that the correlation function is bounded by 1, i.e., $|R(t, s)| \leq 1$, see (Spodarev, 2013). The core of this section will be the Gaussian random field.

Definition 1.2.4. *We say that the random variable $Z(t)$ is a Gaussian random variable it has a density function*

$$\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \quad x \in \mathbb{R},$$

for the mean value m of $Z(t)$ and the variance $\sigma^2 = C(t, t)$. We denote it by $Z(t) \sim N(m, \sigma^2)$.

Furthermore, $Z = \{Z(t) : t \in \mathcal{T}\}$ is said to be a Gaussian random field if $Z(t)$ is a Gaussian random variable for all $t \in \mathcal{T}$. In the case that $\mathcal{T} = \mathbb{R}^d$ and each $Z(t)$ is \mathbb{R}^d -valued variable, then $Z = \{Z_i(t) : t \in \mathcal{T}, i = 1, \dots, n\}$ is multivariate Gaussian random field if it has a density function

$$\varphi(\mathbf{x}) = \frac{1}{(\sqrt{2\pi})^n \sqrt{\det(\mathbf{C})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m})\mathbf{C}^{-1}(\mathbf{x} - \mathbf{m})^T\right), \quad \mathbf{x} \in \mathbb{R}^d,$$

for the mean vector $\mathbf{m} = (\mathbb{E}(Z_1(t)), \dots, \mathbb{E}(Z_n(t)))$ and covariance matrix $\mathbf{C} = [C_{ij}]$, where $C_{ij} = \mathbb{E}[Z_i Z_j] - \mathbb{E}[Z_i]\mathbb{E}[Z_j]$, for $i, j = 1, \dots, n$ and we denote it by $N_n(\mathbf{m}, \mathbf{C})$, see (Spodarev, 2013; Adler and Taylor, 2007).

1.2.2 Spectral Representation for a Random Field

We shall introduce the notions of stationary (homogeneous) and isotropic random fields in the following definition.

Definition 1.2.5. Let $Z = \{Z(t) : t \in \mathcal{T}\}$ be a random field.

1. Z is said to be **strongly stationary** if for any $n \in \mathbb{N}$, $\tau, t_1, t_2, \dots, t_n \in \mathcal{T}$, the distribution of $(Z(t_1), \dots, Z(t_n))$ is the same as the distribution of $(Z(t_1 + \tau), \dots, Z(t_n + \tau))$.
2. Let $\mathbb{E}|Z(t)|^2 < \infty, t \in \mathcal{T}$, we say that the random field Z is **weakly stationary** if it has mean function $m(t)$, $t \in \mathcal{T}$ and covariance function $C(t, s) = C(t + \tau, s + \tau) = C(s - t)$, for all $\tau, t, s \in \mathcal{T}$, i.e., the covariance function depends only on the difference $s - t$.
3. A stationary random field is said to be **isotropic** if the covariance function of Z depends only on the distance, i.e., $C(t, s) = C(\|t - s\|)$, where $\|t - s\|$ denotes the distance between t and s , see Table 1.1.

For example, the next table contains some examples of isotropic covariance functions on \mathbb{R}^d , see [Gneiting \(2013a\)](#).

Model	$C(\ t - s\)$	Parameter
Matérn	$\frac{2^{\nu-1}\sigma^2}{\Gamma(\nu)} \left(\frac{\ t-s\ }{c}\right)^\nu K_\nu\left(\frac{\ t-s\ }{c}\right)$	$c > 0; \nu > 0$
Powered exponential	$\sigma^2 \exp\left(-\left(\frac{\ t-s\ }{c}\right)^\alpha\right)$	$c > 0; 0 < \alpha \leq 2$
Askey	$\sigma^2 \left(1 - \frac{\ t-s\ }{c}\right)_+^\alpha$	$c > 0; \alpha \geq 2$

TABLE 1.1. Parametric models of isotropic covariance functions with the range parameter, where K_ν denotes the modified Bessel function of the second kind of order ν ([Gradshteyn and Ryzhik, 2007](#)).

Note that we can also say that the random field is strongly isotropic if the distribution of Z is invariant under rotation, i.e., let $\mathcal{T} = \mathbb{R}^d$ and $Q \in \text{SO}(d)$, the the distribution of $(Z(t_1), \dots, Z(t_n))$ is the same as the distribution of $(Z(Qt_1), \dots, Z(Qt_n))$, see ([Spodarev, 2013](#)). Let $Z = \{Z(t) : t \in \mathbb{R}^d\}$ be a centered (i.e., with a zero mean function) complex-valued random field on

the probability space $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ with $\mathbb{E}|Z(t)|^2 < \infty$ and covariance function $C(t, s) = \mathbb{E}[Z(t)Z(s)]$, $t, s \in \mathcal{T}$. Following [Spodarev \(2013\)](#) the spectral representation of $Z(t)$ is given by

$$Z(t) = \int_{\mathbb{R}^d} \exp(i\langle t, x \rangle) d\mu(x), \quad t \in \mathbb{R}^d, \quad (1.13)$$

where μ is an independently centered random measure on $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ with control measure ν . This construction allows us to state the so-called spectral representation theorem.

Theorem 1.2.6. ([Spodarev, 2013](#)) *The random field Z has spectral representation if and only if its covariance function $C(t, s)$ can be represented in the form*

$$C(t, s) = \int_{\mathbb{R}^d} \exp(i\langle (t - s), x \rangle) d\nu(x), \quad t, s \in \mathbb{R}^d. \quad (1.14)$$

Similarly, in Section [2.2](#) we state the representation of Gaussian random fields on spheres.

1.3 Spherical Harmonics and $L^2(\mathbb{S}^d)$ –Space

Harmonic analysis on the circle means Fourier analysis. In general, [Fourier \(1822\)](#) stated that, every square-integrable \mathbf{T} –periodic function f on the interval $[x_0, x_0 + \mathbf{T}]$ can be expressed uniquely as sums of the fundamental exponential functions, and has the form

$$f(x) = \sum_{n \in \mathbb{Z}} c_n \exp\left(-\frac{2\pi i n x}{\mathbf{T}}\right), \quad (1.15)$$

where c_n denotes the Fourier coefficients, and are given by

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot \exp\left(-\frac{2\pi i n x}{\mathbf{T}}\right) dx, \quad \text{for all } n \in \mathbb{Z}. \quad (1.16)$$

1.3.1 Fourier Representation From \mathbb{R}^d to \mathbb{S}^d

In fact, the basic idea of Fourier series is the so-called orthogonal decomposition theorem, which states that the basis of Fourier series is orthogonal, see (Cooperstein, 2010).

Theorem 1.3.1 (The Orthogonal Decomposition Theorem). *Let W be a subspace of \mathbb{R}^d . Then each $\mathbf{y} \in \mathbb{R}^d$ can be uniquely represented in the form $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}} \in W$ and $\mathbf{z} \in W$. In fact, if $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ is any orthogonal basis for W , then*

$$\hat{\mathbf{y}} = \left(\frac{\langle \mathbf{y} \cdot \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1 \cdot \mathbf{u}_1 \rangle} \right) \mathbf{u}_1 + \dots + \left(\frac{\langle \mathbf{y} \cdot \mathbf{u}_n \rangle}{\langle \mathbf{u}_n \cdot \mathbf{u}_n \rangle} \right) \mathbf{u}_n,$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$, where $\langle \cdot, \cdot \rangle$ is the standard inner product. The vector $\hat{\mathbf{y}}$ is called the orthogonal projection of \mathbf{y} onto W .

By Theorem 1.3.1, we can drive the generalized Fourier series for an arbitrary function based on the basis of orthogonal functions. Let $f(x)$ and $g(x)$ be functions defined on an interval I , then we define the inner product with respect to the weight function $w(x) > 0$ (Rudin, 2011; Grafakos, 2008) as follows

$$\langle f, g \rangle = \int_I f(x)g(x)w(x)dx. \quad (1.17)$$

The space of square-integrable functions on I , denoted as $L^2(I)$, is formed by function on I such that $\langle f, f \rangle < \infty$. On the circle, \mathbb{S}^1 , the inner product (1.17) for given functions f and g on \mathbb{S}^1 with respect the weight function $w(x) = 1$ is given by

$$\langle f, g \rangle = \int_{\mathbb{S}^1} f(x)g(x)dx. \quad (1.18)$$

It is clear that the space $L^2(\mathbb{S}^1)$ associated to the inner preduct (1.18) is Hilbert space. Let the subspace $\mathcal{H}_n(\mathbb{S}^1) = \text{span}\{\cos nx, \sin nx\}$ for all $n \in \mathbb{N}$ with $\mathcal{H}_0(\mathbb{S}^1) = \mathbb{R}$, then by Hilbert sum decomposition, we have

$$L^2(\mathbb{S}^1) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n(\mathbb{S}^1),$$

the subspaces $\{\mathcal{H}_n(\mathbb{S}^1)\}_{n=0}^\infty$ are pairwise orthogonal. The next definition is the generalized Fourier series (Fekete, 1935; Grafakos, 2008; Marks, 2009; Rudin, 2011).

Definition 1.3.2 (Generalized Fourier Series). *Let $\{f_n(x)\}_{n=0}^\infty$ be a sequence of orthogonal basis functions, and $f(x)$ is represented by*

$$f(x) = \sum_{n=0}^{\infty} c_n f_n(x). \quad (1.19)$$

Then, the general Fourier coefficients c_n are determined as

$$c_n = \frac{\langle f, f_n \rangle_w}{\langle f_n, f_n \rangle_w}, \quad (1.20)$$

where $\langle \cdot, \cdot \rangle_w$ means the inner product with respect to the weight function $w(x) > 0$.

After this quick glimpse of the Fourier series for a function defined on \mathbb{R} , now, we shall introduce the so-called spherical harmonics (i.e., the Fourier series on the sphere). Let $d \geq 1$ be an integer, the unit sphere denoted by \mathbb{S}^d defined as follows

$$\mathbb{S}^d := \{(x_1, \dots, x_{d+1}) \in \mathbb{R}^{d+1} \mid x_1^2 + \dots + x_{d+1}^2 = 1\}.$$

We start with $\mathcal{P}_n(d+1)$, the space of homogeneous polynomials of degree n in \mathbb{R}^{d+1} with real coefficients, and let $\mathcal{H}_n(d+1)$, the space of real harmonic polynomials in $\mathcal{P}_n(d+1)$, i.e.,

$$\mathcal{H}_n(d+1) = \{P \in \mathcal{P}_n(d+1) \mid \Delta P = 0\},$$

where, Δ denotes the Laplace operator. Let $\mathcal{H}(\mathbb{S}^d)$ be the restrictions of $\mathcal{H}_n(d+1)$ to \mathbb{S}^d , the so-called spherical harmonics. The next lemma gives the relations between the spaces $\mathcal{H}_n(d+1)$ and $\mathcal{H}(\mathbb{S}^d)$, see (Atkinson and Han, 2012).

Lemma 1.3.3. *The space $\mathcal{H}_n(d+1)$ is isomorphic to their restrictions to \mathbb{S}^d .*

We conclude from Lemma 1.3.3 that $\dim(\mathcal{H}_n(d+1)) = \dim(\mathcal{H}(\mathbb{S}^d))$. Since the elements in $\mathcal{H}(\mathbb{S}^d)$ are homogeneous polynomials P on spheres with $\Delta_{\mathbb{S}^d} P \equiv 0$,

we conclude that the dimension of this space $\mathcal{H}(\mathbb{S}^d)$ is

$$h(n, d+1) = \dim(\mathcal{H}(\mathbb{S}^d)) = \frac{(2n+d-1)(n+d-2)!}{n!(d-1)!},$$

with $h(0, d+1) = 1$ for all $d \geq 1$, and its asymptotic behavior is $h(n, d+1) = \mathcal{O}(n^{d-1})$ (Morimoto, 1998, Theorem 2.4) or (Jean, 1980, Chapter 7). Let $L^2(\mathbb{S}^d)$ be the space of squared-integrable real-valued functions on the sphere \mathbb{S}^d , which is a Hilbert space with inner product defined by

$$\langle f, g \rangle_{L^2(\mathbb{S}^d)} := \frac{1}{\mathfrak{S}(\mathbb{S}^d)} \int_{\mathbb{S}^d} f(\mathbf{x})g(\mathbf{x})d(\mathbb{S}^d), \quad (1.21)$$

for every $f, g \in L^2(\mathbb{S}^d)$ and $\mathbf{x} \in \mathbb{S}^d$, where $d(\mathbb{S}^d)$ is the surface measure and $\mathfrak{S}(\mathbb{S}^d)$ is the surface area of \mathbb{S}^d (Wang et al., 2016),

$$\mathfrak{S}(\mathbb{S}^d) := \int_{\mathbb{S}^d} d(\mathbb{S}^d) = \frac{2\pi^{\frac{d+1}{2}}}{\Gamma(\frac{d+1}{2})}. \quad (1.22)$$

Similarly, we define the square-integrable functions on spheres, $L^2(\mathbb{S}^d)$, by the Hilbert sum decomposition

$$L^2(\mathbb{S}^d) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n(\mathbb{S}^d),$$

the subspaces $\{\mathcal{H}_n(\mathbb{S}^d)\}_{n=0}^{\infty}$ are pairwise orthogonal, and the set of all finite linear combination of elements in $\bigcap_{n=0}^{\infty} \mathcal{H}_n(\mathbb{S}^d)$ is dense in $L^2(\mathbb{S}^d)$. Following Lang (1993, Corollary 1.9), we conclude that for every $f \in L^2(\mathbb{S}^d)$ can be uniquely represented as sum of a converging series

$$f = \sum_{n=0}^{\infty} f_n, \quad (1.23)$$

where $f_n \in \mathcal{H}_n(\mathbb{S}^d)$. Let $\{Y_n^\ell \mid 1 \leq \ell \leq h(n, d+1)\}$ be an orthonormal basis of $\mathcal{H}_n(\mathbb{S}^d)$, i.e.,

$$\langle Y_n^{\ell_1}, Y_n^{\ell_2} \rangle_{L^2(\mathbb{S}^d)} = \delta_{\ell_1 \ell_2}, \quad 1 \leq \ell_1, \ell_2 \leq h(n, d+1),$$

where $\delta_{\ell_1 \ell_2}$ is the Kronecker delta. Then we have,

$$f_n = \sum_{\ell=1}^{h(n,d+1)} c_{n,\ell} Y_n^\ell, \quad \text{where } c_{n,\ell} = \langle f, Y_n^\ell \rangle. \quad (1.24)$$

The coefficients $c_{n,\ell}$ is the generalized Fourier coefficients with respect to the Hilbert space $L^2(\mathbb{S}^d)$.

1.3.2 Zonal Spherical Harmonics Functions

In this section, we study an important concept in harmonic analysis, that is, zonal function and its relation with positive definiteness. Also, we shall study the Schoenberg's coefficients of a given function on spheres \mathbb{S}^d . Let $\{Y_n^\ell \mid 1 \leq \ell \leq h(n, d+1)\}$ be an orthonormal basis of $\mathcal{H}_n(\mathbb{S}^d)$ and define the zonal harmonic polynomial of degree n by

$$Z_n(\mathbf{x}, \mathbf{y}) = \sum_{\ell=1}^{h(n,d+1)} Y_n^\ell(\mathbf{x}) \overline{Y_n^\ell(\mathbf{y})}, \quad \mathbf{x}, \mathbf{y} \in \mathbb{S}^d. \quad (1.25)$$

Following (Morimoto, 1998, Lemma 2.23), we have the next proposition, that gives an interesting property of zonal harmonics polynomials.

Proposition 1.3.4. *For all $\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}' \in \mathbb{S}^d$, with $d \geq 1$, if $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}', \mathbf{y}' \rangle$, then $Z_n(\mathbf{x}, \mathbf{y}) = Z_n(\mathbf{x}', \mathbf{y}')$. Furthermore, there exists some function, $\rho : \mathbb{R} \rightarrow \mathbb{R}$, such that $Z_n(\mathbf{x}, \mathbf{y}) = \rho(\langle \mathbf{x}, \mathbf{y} \rangle)$.*

We conclude from Proposition 1.3.4 that the function $Z_n(\mathbf{x}, \mathbf{y})$ depends only on the inner product $\langle \mathbf{x}, \mathbf{y} \rangle$. Let $C_n^d(\cdot)$ denote Gegenbauer polynomials of degree n and dimension $d+1$ with $C_n^d(1) = 1$. Theorem 2.24 in Morimoto (1998) studied the relation between Gegenbauer polynomials $C_n^d(\cdot)$ and zonal harmonics $Z_n(\mathbf{x}, \mathbf{y})$.

Theorem 1.3.5. *For any fixed point $\mathbf{x} \in \mathbb{S}^d$ and for any constant $c \in \mathbb{R}$, there is a unique homogeneous harmonic polynomial $Z_n \in \mathcal{H}_n(d+1)$, such that $\mathbf{y} \in \mathbb{S}^d$*

is invariant under rotation, and $Z_n(\mathbf{x}, \mathbf{x}) = c$. Moreover, we have

$$Z_n(\mathbf{x}, \mathbf{y}) = c \|\mathbf{y}\|^n C_n^d \left(\left\langle \frac{\mathbf{y}}{\|\mathbf{y}\|}, \mathbf{x} \right\rangle \right).$$

Since $\{Y_n^\ell \mid 1 \leq \ell \leq h(n, d+1)\}$ are the orthonormal basis of $\mathcal{H}_n(\mathbb{S}^d)$, then we can calculate the constant c in Theorem 1.3.5 as follows

$$c = \frac{h(n, d+1)}{\mathfrak{S}(\mathbb{S}^d)}.$$

Thus we can define the zonal harmonics in terms of Gegenbauer polynomials C_n^d , as follows

$$Z_n(\mathbf{x}, \mathbf{y}) = \frac{h(n, d+1)}{\mathfrak{S}(\mathbb{S}^d)} C_n^d(\langle \mathbf{x}, \mathbf{y} \rangle). \quad (1.26)$$

Then by the definition of zonal harmonics (1.25), we have

$$C_n^d(\langle \mathbf{x}, \mathbf{y} \rangle) = \frac{\mathfrak{S}(\mathbb{S}^d)}{h(n, d+1)} \sum_{\ell=1}^{h(n, d+1)} Y_n^\ell(\mathbf{x}) \overline{Y_n^\ell(\mathbf{y})}. \quad (1.27)$$

Going back to Equation (1.23), the following proposition will give a closed form of the sequence f_n , for all $n \geq 0$.

Proposition 1.3.6. *Consider a continuous function $f \in L^2(\mathbb{S}^d)$, if f can be uniquely represented as $f = \sum_n f_n$ with $f_n \in \mathcal{H}_n(\mathbb{S}^d)$, then f_n is given by*

$$f_n(\mathbf{x}) = \frac{h(n, d+1)}{\mathfrak{S}(\mathbb{S}^d)} \langle f(\mathbf{y}), C_n^d(\langle \mathbf{x}, \mathbf{y} \rangle) \rangle_{\mathbb{S}^d}, \quad (1.28)$$

for all $\mathbf{x} \in \mathbb{S}^d$.

Following (Morimoto, 1998, Theorem 2.53), and denoting $C_n^\lambda(t)$, $t \in [-1, 1]$ for the ultraspherical polynomials of degree n associated with $\lambda = (d-1)/2 > 0$, then we have

$$C_n^d(t) = \frac{2n+d-1}{(d-1)h(n, d+1)} C_n^\lambda(t), \quad (1.29)$$

and $C_n^\lambda(1) = \Gamma(n + 2\lambda)/(\Gamma(n + 1)\Gamma(2\lambda))$. Thus, we can rewrite Equation (1.28) as follows

$$f_n(\mathbf{x}) = \frac{2n + d - 1}{(d - 1) \mathfrak{S}(\mathbb{S}^d)} \langle f(\mathbf{y}), C_n^\lambda(\langle \mathbf{x}, \mathbf{y} \rangle) \rangle_{\mathbb{S}^d}. \quad (1.30)$$

1.3.3 Schoenberg's Coefficients on Spheres

Now, by Definition 1.3.2 we can summarize the generalized Fourier series as follows.

Proposition 1.3.7. *For every continuous and square-integrable function f on the interval $[-1, 1]$, $f(t)$ can uniquely represented by*

$$f(t) = \sum_{n=0}^{\infty} \tilde{b}_{n,d} C_n^\lambda(t). \quad (1.31)$$

Then, the general Fourier coefficients $\tilde{b}_{n,d}$ determined as

$$\tilde{b}_{n,d} = \frac{\langle f(t), C_n^\lambda(t) \rangle_w}{\langle C_n^\lambda(t), C_n^\lambda(t) \rangle_w}, \quad \lambda = \frac{d - 1}{2}, \quad (1.32)$$

where $\langle \cdot, \cdot \rangle_w$ means the inner product with respect to the weight function $w(t) = (1 - t^2)^{(d-2)/2}$.

The ultraspherical polynomials satisfies the following property

$$\int_{-1}^1 C_n^\lambda(t) C_m^\lambda(t) (1 - t^2)^{\frac{d-2}{2}} dt = \delta_{mn} \frac{(d - 1)^2 h(n, d + 1) \mathfrak{S}(\mathbb{S}^d)}{(2n + d - 1)^2 \mathfrak{S}(\mathbb{S}^{d-1})}, \quad (1.33)$$

where δ_{mn} is the Kronecker delta.

Now, if we considered two points \mathbf{x} and \mathbf{y} on the surface of \mathbb{S}^d , then the spherical or geodesic distance between \mathbf{x} and \mathbf{y} is given by

$$\theta(\mathbf{x}, \mathbf{y}) = \arccos(\langle \mathbf{x}, \mathbf{y} \rangle).$$

Thus, we conclude that the geodesic distance between any two points on the surface of \mathbb{S}^d depends on their inner product. Next proposition gives the Fourier series for every continuous square-integrable function of geodesic distance.

Proposition 1.3.8. *Let f be a continuous and square-integrable function on $[0, \pi]$, $f(\cos \theta)$ can uniquely represented by*

$$f(\cos \theta) = \sum_{n=0}^{\infty} \tilde{b}_{n,d} C_n^\lambda(\cos \theta). \quad (1.34)$$

Then, the coefficients $\tilde{b}_{n,d}$ are given by

$$\tilde{b}_{n,d} = \frac{(2n + d - 1) \Gamma(\lambda)}{2 \sqrt{\pi} \Gamma\left(\frac{d}{2}\right) C_n^\lambda(1)} \int_0^\pi f(\cos \theta) C_n^\lambda(\cos \theta) \sin^{(d-1)} \theta \, d\theta, \quad (1.35)$$

where $\lambda = \frac{d-1}{2} > 0$.

Proof. To obtain Equation (1.34), since $|\cos \theta| \leq 1$ then we can substitute $t = \cos \theta$ in Proposition 1.3.7. By applying Propodition 1.3.7 with weight function $w(\cos \theta) = \sin^{(d-2)} \theta$, then we have

$$\tilde{b}_{n,d} = \frac{\langle f(\cos \theta), C_n^\lambda(\cos \theta) \rangle_w}{\langle C_n^\lambda(\cos \theta), C_n^\lambda(\cos \theta) \rangle_w}. \quad (1.36)$$

Using Equation (1.33) and changing the variables ($t = \cos \theta$), then we have

$$\langle C_n^\lambda(\cos \theta), C_n^\lambda(\cos \theta) \rangle_w = \frac{(d-1)^2 h(n, d+1) \mathfrak{S}(\mathbb{S}^d)}{(2n+d-1)^2 \mathfrak{S}(\mathbb{S}^{d-1})}, \quad (1.37)$$

and since

$$\langle f(\cos \theta), C_n^\lambda(\cos \theta) \rangle_w = \int_0^\pi f(\cos \theta) C_n^\lambda(\cos \theta) \sin^{(d-1)} \theta \, d\theta, \quad (1.38)$$

then substituting (1.38) and (1.37) in (1.36), we have

$$\tilde{b}_{n,d} = \frac{(2n+d-1)^2 \mathfrak{S}(\mathbb{S}^{d-1})}{(d-1)^2 h(n, d+1) \mathfrak{S}(\mathbb{S}^d)} \int_0^\pi f(\cos \theta) C_n^\lambda(\cos \theta) \sin^{(d-1)} \theta \, d\theta. \quad (1.39)$$

By definition of $h(n, d + 1)$ and surface measure on \mathbb{S}^d , we have

$$\begin{aligned}
\tilde{b}_{n,d} &= \frac{2(2n + d - 1)^2 \pi^{\frac{d}{2}} \Gamma\left(\frac{d+1}{2}\right) n! (d - 1)!}{2(d - 1)^2 (2n + d - 1) (n + d - 2)! \pi^{\frac{d+1}{2}} \Gamma\left(\frac{d}{2}\right)} \\
&\quad \times \int_0^\pi f(\cos \theta) C_n^\lambda(\cos \theta) \sin^{(d-1)} \theta \, d\theta \\
&= \frac{(2n + d - 1) \Gamma\left(\frac{d+1}{2}\right) n! (d - 2)!}{(d - 1) \sqrt{\pi} (n + d - 2)! \Gamma\left(\frac{d}{2}\right)} \\
&\quad \times \int_0^\pi f(\cos \theta) C_n^\lambda(\cos \theta) \sin^{(d-1)} \theta \, d\theta \\
&= \frac{(2n + d - 1) (d - 1) \Gamma\left(\frac{d-1}{2}\right) \Gamma(n + 1) \Gamma(d - 1)}{2(d - 1) \sqrt{\pi} \Gamma(n + d - 1) \Gamma\left(\frac{d}{2}\right)} \\
&\quad \times \int_0^\pi f(\cos \theta) C_n^\lambda(\cos \theta) \sin^{(d-1)} \theta \, d\theta \\
&= \frac{(2n + d - 1) \Gamma\left(\frac{d-1}{2}\right) \Gamma(n + 1) \Gamma(d - 1)}{2 \sqrt{\pi} \Gamma(n + d - 1) \Gamma\left(\frac{d}{2}\right)} \\
&\quad \times \int_0^\pi f(\cos \theta) C_n^\lambda(\cos \theta) \sin^{(d-1)} \theta \, d\theta \\
&= \frac{(2n + d - 1) \Gamma(\lambda)}{2 \sqrt{\pi} \Gamma\left(\frac{d}{2}\right) C_n^\lambda(1)} \int_0^\pi f(\cos \theta) C_n^\lambda(\cos \theta) \sin^{(d-1)} \theta \, d\theta.
\end{aligned}$$

This completes the proof. \square

The representation of the square-integrable function defined on the sphere in Proposition 1.3.8 is called Schoenberg's representation. The coefficients $\tilde{b}_{n,d}$ in (1.35) have interesting properties as they can be used to determine whether a given function defined on \mathbb{S}^d is a positive definite function (see Schoenberg (1942); Gneiting (2013a)).

Now, we can define the concept of positive definite function on the sphere. We say that the function $C : \mathbb{S}^d \times \mathbb{S}^d \rightarrow \mathbb{R}$ is positive definite if

$$\sum_{i,j=1}^n \alpha_i \alpha_j C(\mathbf{x}_i, \mathbf{x}_j) \geq 0, \tag{1.40}$$

for any $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ and for every $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{S}^d$. Schoenberg (1942) characterized positive definite functions defined on the spheres of any dimension.

Theorem 1.3.9. *Schoenberg (1942) A necessary and sufficient condition for a continuous mapping $\psi : [0, \pi] \rightarrow \mathbb{R}$, with $\psi(0) = 1$ to belong to the class Ψ_d is that the ultraspherical expansion*

$$\sum_{n=0}^{\infty} \left\{ \frac{(n+\lambda)\Gamma(\lambda)}{\Gamma(\lambda+\frac{1}{2})\Gamma(\frac{1}{2})} \times \int_0^\pi C_n^\lambda(\cos \theta') \psi(\cos \theta') \sin^{d-1} \theta' d\theta' \right\} C_n^\lambda(\cos \theta) \quad (1.41)$$

has non-negative coefficients and converges absolutely and uniformly throughout $0 \leq \theta \leq \pi$ to $\psi(\theta)$.

When $\lambda \rightarrow 0$ the ultraspherical polynomials tends to Chebyshev polynomials of the first kind, that is,

$$\lim_{\lambda \rightarrow 0} \frac{C_n^\lambda(\cos \theta)}{C_n^\lambda(1)} = T_n(\cos \theta) = \cos(n\theta).$$

Thus, by Proposition 1.3.7, any continuous function on the circle \mathbb{S}^1 has Fourier's coefficients as follows,

$$\begin{aligned} \tilde{b}_{0,1} &= \frac{1}{\pi} \int_0^\pi f(\cos \theta) \cos(n\theta) d\theta, \\ \tilde{b}_{n,1} &= \frac{2}{\pi} \int_0^\pi f(\cos \theta) \cos(n\theta) d\theta, \quad \text{for } n \in \mathbb{N}, \end{aligned} \quad (1.42)$$

these are the Fourier coefficients on the circle.

Let Ψ_d be the class of continuous mappings $\psi : [0, \pi] \rightarrow \mathbb{R}$ with $\psi(0) = 1$ such that the continuous functions $C : \mathbb{S}^d \times \mathbb{S}^d \rightarrow \mathbb{R}$ defined through $C(\xi, \eta) = \psi(\theta(\xi, \eta))$ are positive definite. Gneiting (2013a) used Theorem 1.3.9 to characterize the members of class Ψ_d through the representation

$$\psi(\theta) = \sum_{n=0}^{\infty} b_{n,d} \frac{C_n^{(d-1)/2}(\cos \theta)}{C_n^{(d-1)/2}(1)}, \quad \theta \in [0, \pi], \quad (1.43)$$

where $b_{n,d} = \tilde{b}_{n,d} C_n^\lambda(1)$ with $\{b_{n,d}\}_{n=0}^\infty$ being a uniquely identified probability mass system and given by

$$b_{n,d} = \frac{2n + d - 1}{2^{3-d}\pi} \frac{(\Gamma(\frac{d-1}{2}))^2}{\Gamma(d-1)} \int_0^\pi \psi(\theta) C_n^{(d-1)/2}(\cos \theta) (\sin \theta)^{d-1} d\theta, \quad (1.44)$$

We follow [Daley and Porcu \(2014\)](#) and [Ziegel \(2014\)](#) when calling $b_{n,d}$ for the d -Schoenberg coefficients.

CHAPTER 2

Equivalence and Orthogonality of Gaussian Measures on Spheres

In this chapter, we provide necessary and sufficient conditions for the equivalence of two Gaussian measures with two different covariance models with associated d -Schoenberg sequences.

In general, any two probability measures there are three possibilities of their relationship to each other, equivalent, orthogonal or neither. The next section gives some basic to study that subject. For example, if we taken $E = \{1, 2, 3, 4, 5\}$ with two probability measures μ and ν on $\mathcal{B}(E)$ such that μ allocate probability $\frac{1}{4}$ to $\{1, 2, 3\}$ and probability zero to $\{4, 5\}$ and ν allocate probability $\frac{1}{4}$ to $\{2, 3, 4\}$ and probability zero to $\{1, 5\}$. Hence, μ and ν are equivalent for $\{2, 3, 5\}$, and are orthogonal for $\{1, 4\}$, so in general we say in this case that μ and ν are neither equivalent nor orthogonal, see (Stein, 1999a; Gelfand et al., 2010). As for Gaussian measures have only two possibilities equivalent or orthogonal, see (Ibragimov and Rozanov, 1978).

2.1 Some Basic Structures

Let E be an infinite-dimensional real separable Banach space with associated norm $\|\cdot\|$, and Borel σ -algebra $\mathcal{B}(E)$. We denote by $\mathcal{C}_b(E)$ the Banach space of all continuous, bounded real-valued functions defined on E equipped with the norm

$$\|\varphi\| = \sup_{x \in E} |\varphi(x)|,$$

and for the closed subspace of $\mathcal{C}_b(E)$ of all real-valued function which are uniformly continuous by $\mathcal{C}_b^u(E)$. A probability measure on a measurable space $(E, \mathcal{B}(E))$ is a σ -additive function μ from $\mathcal{B}(E)$ into $[0, 1]$ such that $\mu(E) = 1$. Hence, the probability measure on E usually mean a probability measure defined on the Borel σ -algebra of E . The measures on $(E, \mathcal{B}(E))$ are completely determined by their integrals with respect to all $\varphi \in \mathcal{C}_b(E)$.

Definition 2.1.1. (*Da Prato and Zabczyk, 1992*) Let μ, ν are two probability measures on $\mathcal{B}(E)$, we say that μ equal to ν (notation: $\mu = \nu$) if for all $\varphi \in \mathcal{C}_b(E)$, then

$$\int_E \varphi(x) \mu(dx) = \int_E \varphi(x) \nu(dx).$$

We propose now to recall the definition of the absolutely continuous of probability measures.

Definition 2.1.2. Let μ, ν are two probability measures on $\mathcal{B}(E)$, we say that μ is absolutely continuous with respect to ν (notation: $\mu \ll \nu$) if $\nu(A) = 0$ implies $\mu(A) = 0$ for every $A \in \mathcal{B}(E)$ (i.e., every null set of ν is also a null set of μ).

Similarly, we can define $\nu \ll \mu$. Following [Skorokhod and Yadrenko \(1973\)](#), we say that $\mu \sim \nu$ if and only if, μ is absolutely continuous with respect to ν and ν is absolutely continuous with respect to μ , i.e., for every $A \in \mathcal{B}(E)$,

$$\mu(A) = 0 \iff \nu(A) = 0.$$

Now, we will discuss the relation between probability measures and its density functions by using theorem of Radon-Nikodym, see (Rudin, 1986; Da Prato, 2006).

Theorem 2.1.3 (Radon-Nikodym). *Consider μ, ν are two probability measure on $\mathcal{B}(E)$. If μ is absolutely continuous with respect to ν , then there exists a unique non-negative measurable function φ on E , such that for every $A \in \mathcal{B}(E)$,*

$$\mu(A) := \int_A \varphi(x) \nu(dx), \quad (2.1)$$

and φ is called the density function of μ relative to ν .

By Theorem 2.1.3 we can define the equivalence relation between the probability measures. We said that $\mu \sim \nu$ if and only if $\mu \ll \nu$, and $\nu \ll \mu$. Thus there are density functions $\varphi, \psi \in \mathcal{B}(E)$ associated to μ , and ν , respectively such that,

$$\varphi(x) = \frac{d\mu}{d\nu}(x), \quad \psi(x) = \frac{d\nu}{d\mu}(x), \quad x \in E,$$

and satisfy $\varphi(x)\psi(x) = 1$ almost everywhere $x \in E$. Hence, $\varphi(x) > 0$ almost everywhere $x \in E$ with respect to μ . To study the equivalence and the orthogonality of any two probability measures, we will provide the following important definition, that is, Kullback-Leibler divergence, see (Hershey and Olsen, 2007; Heyer, 2012).

Definition 2.1.4 (Kullback-Leibler Divergence). *Let φ and ψ are probability density function on E . The Kullback-Leibler divergence (For short, KL) is defined as*

$$KL(\varphi||\psi) = \int_E \varphi(x) \log \left(\frac{\varphi(x)}{\psi(x)} \right) dx.$$

KL-divergence is commonly referred to as the relative entropy distance between probability distributions. It is so easy to show that the KL-divergence is non-negative, and equal to zero at $\varphi = \psi$ almost everywhere. Note that the KL-divergence is not a distance (metric) in the sense that not satisfies all metric axioms. For example, it does not satisfy the symmetry, i.e., $KL(\varphi||\psi) \neq KL(\psi||\varphi)$.

So, we will give another definition KL-divergence such that satisfy the symmetry, so-called "symmetrized KL-divergence".

Definition 2.1.5 (Symmetrized Kullback-Leibler Divergence). *Let φ and ψ are probability density function on E . The symmetrized Kullback-Leibler divergence is defined as*

$$D_{KL}(\varphi, \psi) = KL(\varphi||\psi) + KL(\psi||\varphi). \quad (2.2)$$

Note that, symmetrized KL-divergence is symmetric but not a metric, because it does not satisfy the triangle inequality. We can rewrite (2.2) in the following formula,

$$D_{KL}(\varphi, \psi) = \int_E [\varphi(x) - \psi(x)] \log \left(\frac{\varphi(x)}{\psi(x)} \right) dx. \quad (2.3)$$

Also, we call the symmetrized KL-divergence by "entropy distance" between φ and ψ .

Example 2.1.6. *Suppose that φ and ψ are two probability density functions of d -dimensional normal distribution with means \mathbf{m}_1 and \mathbf{m}_2 , and variances \mathbf{C}_1 and \mathbf{C}_2 , respectively. The KL-divergence from φ to ψ , is defined by*

$$KL(\varphi||\psi) = \frac{1}{2} \left[\log \left(\frac{\det(\mathbf{C}_2)}{\det(\mathbf{C}_1)} \right) + (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{C}_2^{-1} (\mathbf{m}_2 - \mathbf{m}_1) + \text{Tr}(\mathbf{C}_2^{-1} \mathbf{C}_1) - d \right].$$

Since, the probability density functions of d -dimensional Gaussian, is given by

$$\begin{aligned} \varphi(x) &= \frac{1}{\sqrt{\det(2\pi\mathbf{C}_1)}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m}_1)^T \mathbf{C}_1^{-1} (\mathbf{x} - \mathbf{m}_1) \right\}, \\ \psi(x) &= \frac{1}{\sqrt{\det(2\pi\mathbf{C}_2)}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m}_2)^T \mathbf{C}_2^{-1} (\mathbf{x} - \mathbf{m}_2) \right\}. \end{aligned}$$

Thus, we have

$$\log \left(\frac{\varphi(x)}{\psi(x)} \right) = \frac{1}{2} \left[\log \left(\frac{\det(\mathbf{C}_2)}{\det(\mathbf{C}_1)} \right) + (\mathbf{x} - \mathbf{m}_2)^T \mathbf{C}_2^{-1} (\mathbf{x} - \mathbf{m}_2) - (\mathbf{x} - \mathbf{m}_1)^T \mathbf{C}_1^{-1} (\mathbf{x} - \mathbf{m}_1) \right]$$

Hence,

$$\begin{aligned}
\text{KL}(\varphi||\psi) &= \mathbb{E} \log \left(\frac{\varphi(x)}{\psi(x)} \right) \\
&= \frac{1}{2} \left[\log \left(\frac{\det(\mathbf{C}_2)}{\det(\mathbf{C}_1)} \right) + \mathbb{E}[(\mathbf{x} - \mathbf{m}_2)^T \mathbf{C}_2^{-1} (\mathbf{x} - \mathbf{m}_2)] - \right. \\
&\quad \left. \mathbb{E}[(\mathbf{x} - \mathbf{m}_1)^T \mathbf{C}_1^{-1} (\mathbf{x} - \mathbf{m}_1)] \right] \\
&= \frac{1}{2} \left[\log \left(\frac{\det(\mathbf{C}_2)}{\det(\mathbf{C}_1)} \right) + (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{C}_2^{-1} (\mathbf{m}_2 - \mathbf{m}_1) \right. \\
&\quad \left. + \text{Tr}(\mathbf{C}_2^{-1} \mathbf{C}_1) - \text{Tr}(\mathbf{C}_1^{-1} \mathbf{C}_1) \right]
\end{aligned}$$

It is easy to find the symmetrized KL-divergence for the probability densities in the previous example, that is, given by this from

$$\begin{aligned}
D_{\text{KL}}(\varphi, \psi) &= \frac{1}{2} \left[(\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{C}_2^{-1} (\mathbf{m}_2 - \mathbf{m}_1) + (\mathbf{m}_2 - \mathbf{m}_1)^T \mathbf{C}_1^{-1} (\mathbf{m}_2 - \mathbf{m}_1) \right. \\
&\quad \left. + \text{Tr}(\mathbf{C}_2^{-1} \mathbf{C}_1) + \text{Tr}(\mathbf{C}_1^{-1} \mathbf{C}_2) - 2d \right].
\end{aligned}$$

Consider μ and ν be two probability measures. Then according to Theorem 2.1.3, we have a corresponding probability density functions φ , and ψ are called spectral density of μ and ν , respectively. By the uniqueness in "Radon-Nikodym" theorem, we can gain the equivalence and orthogonality of μ and ν by looking at the conduct of their spectral densities φ and ψ , respectively. Let us start with the definition of Hellinger integral for two probability measures on $\mathcal{B}(E)$.

Definition 2.1.7. *Let λ, μ, ν are probability measures such that both μ and ν are absolutely continuous with respect to λ . The Hellinger integral of μ and ν is defined by*

$$H(\mu, \nu) = \int_E \sqrt{\frac{d\mu}{d\lambda}} \sqrt{\frac{d\nu}{d\lambda}} d\lambda. \tag{2.4}$$

It is easy to check that the Hellinger integral is not depend on the choice of the measure λ , for instance, we can take it to be $\frac{1}{2}(\mu + \nu)$ or $(\mu + \nu)$. Also, we can rewrite the Hellinger integral of μ and ν by their spectral densities φ and ψ as

following,

$$H(\varphi, \psi) = \int_E \sqrt{\varphi(x)\psi(x)}dx.$$

Example 2.1.8. Suppose that φ and ψ are two probability density functions of d -dimensional normal distribution with means \mathbf{m}_1 and \mathbf{m}_2 , and variances \mathbf{C}_1 and \mathbf{C}_2 , respectively. The Hellinger integral of φ and ψ , is defined by

$$H(\varphi, \psi) = \frac{\sqrt[4]{\det(\mathbf{C}_1)\det(\mathbf{C}_2)}}{\sqrt{\det(\frac{1}{2}\mathbf{C}_1 + \frac{1}{2}\mathbf{C}_2)}} \exp \left\{ -\frac{1}{8}(\mathbf{m}_1 - \mathbf{m}_2)^T \left(\frac{1}{2}\mathbf{C}_1 + \frac{1}{2}\mathbf{C}_2 \right)^{-1} (\mathbf{m}_1 - \mathbf{m}_2) \right\}.$$

By Hellinger integral formula, we define the Hellinger distance as following formula,

$$D_H(\varphi, \psi) = \sqrt{1 - H(\varphi, \psi)},$$

the connection between Hellinger distance and the KL-divergence is

$$KL(\varphi||\psi) \geq 2D_H^2(\varphi, \psi) = 2(1 - H(\varphi, \psi)).$$

Proposition 2.1.9. (*Da Prato and Zabczyk, 1992*) The following properties hold for an arbitrary pair μ, ν of probability measures on $(E, \mathcal{B}(E))$.

- $0 \leq H(\mu, \nu) \leq 1$.
- $H(\mu, \nu) = 0$ if and only if μ and ν are orthogonal.

2.2 Necessary and Sufficient Condition on Spheres

Equivalence and orthogonality of probability measures are useful tools when assessing the asymptotic properties of both prediction and estimation for Gaussian fields. Denote with μ_i , $i = 1, 2$, two probability measures defined on the same measurable space $\{\Omega, \mathcal{A}\}$, where \mathcal{A} denotes the Borel σ -algebra on Ω . μ_1 and μ_2 are called equivalent (denoted by $\mu_1 \equiv \mu_2$) if $\mu_2(A) = 1$ for any $A \in \mathcal{A}$ implies

$\mu_1(A) = 1$ and vice versa. On the other hand, μ_1 and μ_2 are orthogonal (denoted by $\mu_1 \perp \mu_2$) if there exists an event A such that $\mu_2(A) = 1$ but $\mu_1(A) = 0$. For a real-valued Gaussian random field $Z = \{Z(\mathbf{x}), \mathbf{x} \in \mathbb{S}^d\}$, to define the concepts presented in Section 1.3, we restrict the event A to the σ -algebra generated by Z . We emphasize this restriction by saying that the two measures are equivalent on the paths of Z .

By Theorem 5.13 of [Marinucci and Peccati \(2011\)](#), the Gaussian random field Z admits a spectral representation

$$Z(\mathbf{x}) = \sum_{n=0}^{\infty} \sum_{\ell=1}^{h(n,d+1)} \alpha_{n\ell} Y_n^\ell(\mathbf{x}), \quad \mathbf{x} \in \mathbb{S}^d, \quad (2.5)$$

where $\{\alpha_{n\ell} \mid n \in \mathbb{N}_0, 1 \leq \ell \leq h(n, d + 1)\}$ is a sequence of Gaussian random variables defined by

$$\alpha_{n\ell} = \int_{\mathbb{S}^d} Z(\mathbf{x}) Y_n^\ell(\mathbf{x}) d\sigma_d(\mathbf{x}).$$

Theorem 1.3.9 showed that for the covariance function $K : \mathbb{S}^d \times \mathbb{S}^d \rightarrow \mathbb{R}$ associated to $Z(\mathbf{x})$, there exists a mapping $\psi : [0, \pi] \rightarrow \mathbb{R}$ such that

$$K(\mathbf{x}, \mathbf{y}) = \mathbb{E}(Z(\mathbf{x})Z(\mathbf{y})) = \psi(\theta) = \sum_{n=0}^{\infty} b_{n,d} \frac{C_n^{(d-1)/2}(\cos \theta)}{C_n^{(d-1)/2}(1)}, \quad (2.6)$$

where $\theta \in [0, \pi]$, $\{b_{n,d}\}_{n=0}^{\infty}$ is a uniquely determined sequence of positive coefficients given by (1.44) with $\sum_n b_{n,d} = 1$, and $\mathbb{E}(\cdot)$ denotes the probabilistic expectation. The d -Scheonberg coefficients is given by Equation (1.44) which is due to the fact that the sequence $(\alpha_{n\ell})$ satisfies $\mathbb{E}(\alpha_{n\ell}) = 0$ and

$$\mathbb{E}(\alpha_{n\ell} \alpha_{n'\ell'}) = \frac{\mathfrak{S}(\mathbb{S}^d) b_{n,d}}{h(n, d + 1)} \delta_{nn'} \delta_{\ell\ell'}.$$

Throughout the chapter, we shall refer to K or ψ equivalently as the covariance function of Z , wherever no confusion arises.

Now, let $Z_1(\mathbf{x})$ and $Z_2(\mathbf{x})$ be two zero-mean Gaussian fields with covariance functions $\psi_1(\theta)$ and $\psi_2(\theta)$, respectively. The corresponding covariance opera-

tors on the Hilbert space $L^2(\mathbb{S}^d)$ can be associated with a covariance function $\psi \circ \theta \in L^2(\mathbb{S}^d \times \mathbb{S}^d)$, so that

$$\mathcal{C}_i f(\mathbf{x}) = \int_{\mathbb{S}^d} \psi_i(\theta(\mathbf{x}, \mathbf{y})) f(\mathbf{y}) d\mathbf{y}, \quad \forall f \in L^2(\mathbb{S}^d), \quad i = 1, 2. \quad (2.7)$$

Assume that μ_1 and μ_2 are two Gaussian measures defined on the Borel σ -algebra $\mathcal{B}(\mathbb{S}^d)$ with corresponding covariance operators \mathcal{C}_1 and \mathcal{C}_2 , respectively. [Skorokhod and Yadrenko \(1973\)](#) provided a detailed study of the absolute continuity property for Gaussian measures and their equivalence in the Euclidean case, using the Feldman-Hajek theorem ([Skorokhod and Yadrenko, 1973](#); [Da Prato and Zabczyk, 1992](#)). We use the same approach to study the equivalence and the orthogonality of any two Gaussian measures on the sphere \mathbb{S}^d . An important ingredient for the results proven subsequently is the Feldman-Hajek theorem ([Da Prato and Zabczyk, 1992](#)), which is reported here for as reference contained exposition.

Theorem 2.2.1 (Feldman-Hajek theorem). *Let μ_1 and μ_2 be two zero-mean Gaussian measures with corresponding covariance operators \mathcal{C}_1 and \mathcal{C}_2 , respectively, as defined through (2.7). The measures μ_1 and μ_2 are equivalent if and only if the operator $D = \mathcal{C}_2^{-1/2} \mathcal{C}_1 \mathcal{C}_2^{-1/2}$ is positive definite, invertible, bounded and $D - I \in L^2(\mathbb{S}^d)$, where I is the identity operator. Moreover, the Radon-Nikodym derivative is given by the following formula*

$$\frac{d\mu_2}{d\mu_1} = \prod_{n=0}^{\infty} \prod_{m=1}^{h(n,d+1)} \sqrt{1 + \delta_n^m} \exp \left\{ \frac{-\delta_n^m}{2(1 + \delta_n^m)} \langle \mathcal{C}_2^{-1/2} X(\mathbf{x}), e_n^m \rangle^2 \right\},$$

where e_n^m are the eigenvectors of $D - I$, and δ_n^m are their corresponding eigenvalues.

Two Gaussian measures are either equivalent or orthogonal. This comes from [Ibragimov and Rozanov \(1978\)](#), [Kühn and Liese \(1979\)](#) and [Da Prato and Zabczyk \(1992\)](#), but we state it formally for the convenience of the reader.

Theorem 2.2.2. *Any two zero-mean Gaussian measures μ_1 and μ_2 on $\mathcal{B}(\mathbb{S}^d)$ are either equivalent or orthogonal.*

The next result gives a sufficient and necessary condition for the equivalence of Gaussian measures on the sphere \mathbb{S}^d .

Theorem 2.2.3. *Let $Z_1(\mathbf{x})$ and $Z_2(\mathbf{x})$ be two zero-mean Gaussian random fields on \mathbb{S}^d with corresponding d -Schoenberg coefficients $b_{n,d,1}$ and $b_{n,d,2}$, respectively. Then, the measures μ_1 and μ_2 are equivalent if and only if*

$$\sum_{n=0}^{\infty} h(n, d+1) \left(\frac{b_{n,d,1}}{b_{n,d,2}} - 1 \right)^2 < \infty. \quad (2.8)$$

Proof. Necessity. Let μ_1 and μ_2 be two equivalent Gaussian measures. Arguments in [Ibragimov and Rozanov \(1978\)](#) show that the equivalence of probability measures depends on the entropy distance between them. Hence, a necessary step in the proof is to provide a closed form for the entropy distance between μ_1 and μ_2 . According to [Theorem 2.2.1](#), we define the likelihood ratio $L = d\mu_2/d\mu_1$ between μ_2 and μ_1 on \mathbb{S}^d . We then have

$$\begin{aligned} L &= \prod_{n=0}^{\infty} \prod_{m=1}^{h(n,d+1)} \sqrt{1 + \delta_n^m} \exp \left\{ \frac{-\delta_n^m}{2(1 + \delta_n^m)} \langle \mathcal{C}_2^{-1/2} X(\mathbf{x}), e_n^m \rangle^2 \right\} \\ &= \exp \left\{ -\frac{1}{2} \sum_{n=0}^{\infty} \sum_{m=1}^{h(n,d+1)} \left[\langle \mathcal{C}_2^{-1/2} X(\mathbf{x}), e_n^m \rangle^2 \frac{\delta_n^m}{(1 + \delta_n^m)} - \ln(1 + \delta_n^m) \right] \right\}. \end{aligned} \quad (2.9)$$

The log-likelihood is defined by

$$\ln L = -\frac{1}{2} \sum_{n=0}^{\infty} \sum_{m=1}^{h(n,d+1)} \left[\langle \mathcal{C}_2^{-1/2} X(\mathbf{x}), e_n^m \rangle^2 \frac{\delta_n^m}{(1 + \delta_n^m)} - \ln(1 + \delta_n^m) \right]. \quad (2.10)$$

To find a closed form of the entropy distance between μ_1 and μ_2 , we need to calculate the expectation of [\(2.10\)](#) with respect to μ_1 and μ_2 . By assumption μ_1 and μ_2 are equivalent. Thus, we can assume, following [Skorokhod and Yadrenko \(1973\)](#), that

$$\mathbf{E}_i \left[\langle \mathcal{C}_2^{-1/2} X(\mathbf{x}), e_n^m \rangle^2 \right] = \begin{cases} 1 & \text{for } i = 1 \\ 1 + c_n & \text{for } i = 2, \end{cases}$$

where \mathbb{E}_i is the expectation with respect to μ_i , $c_n \geq 0$ and $\sum c_n^2 < \infty$. Hence, we conclude that the convergence of the series in (2.10) can be evaluated on the basis of the following argument

$$\begin{aligned} \mathbb{E}_1 \left[\langle \mathcal{C}_2^{-1/2} X(\mathbf{x}), e_n^m \rangle^2 \frac{\delta_n^m}{(1 + \delta_n^m)} - \ln(1 + \delta_n^m) \right] &= \frac{\delta_n^m}{(1 + \delta_n^m)} - \ln(1 + \delta_n^m) = \mathcal{O}((\delta_n^m)^2), \\ \mathbb{E}_2 \left[\langle \mathcal{C}_2^{-1/2} X(\mathbf{x}), e_n^m \rangle^2 \frac{\delta_n^m}{(1 + \delta_n^m)} - \ln(1 + \delta_n^m) \right] &= (1 + c_n) \frac{\delta_n^m}{(1 + \delta_n^m)} - \ln(1 + \delta_n^m) = \mathcal{O}((\delta_n^m)^2). \end{aligned}$$

Following Stein (1999a) and Georgiou (2006), we can now define the entropy distance (denoted by R), as follows

$$\begin{aligned} R &= -\mathbb{E}_1(\ln L) + \mathbb{E}_2(\ln L) \\ &= \sum_{n=0}^{\infty} h(n, d+1) \sum_{m=1}^{\infty} \mathcal{O}((\delta_n^m)^2). \end{aligned}$$

Thus, the entropy distance between μ_1 and μ_2 depends on the eigenvalues of the operator $D - I$. Since the covariance operators \mathcal{C}_1 and \mathcal{C}_2 are diagonal (by unique decomposition of positive definite operators), we have

$$\delta_n^m = \frac{b_{n,d,1}}{b_{n,d,2}} - 1.$$

From the arguments in (Ibragimov and Rozanov, 1978) we have that the measures μ_1 and μ_2 are equivalent if and only if the entropy distance between μ_1 and μ_2 is finite, i.e. the following series

$$\sum_{n=0}^{\infty} h(n, d+1) \left(\frac{b_{n,d,1}}{b_{n,d,2}} - 1 \right)^2$$

is convergent, which proves the result.

Sufficiency. Let μ_1 and μ_2 be two zero-mean Gaussian measures with corresponding covariance operators \mathcal{C}_1 and \mathcal{C}_2 , respectively, associated with covariance functions $\psi_1(\theta)$ and $\psi_2(\theta)$ with corresponding d -Schoenberg coefficients $b_{n,d,1}$ and $b_{n,d,2}$, respectively. Assume that (2.8) holds. Define the operator $D : L^2(\mathbb{S}^d) \rightarrow L^2(\mathbb{S}^d)$ by $D = \mathcal{C}_2^{-1/2} \mathcal{C}_1 \mathcal{C}_2^{-1/2}$. Since the covariance operators

\mathcal{C}_1 and \mathcal{C}_2 are diagonal, and by the positive definiteness of \mathcal{C}_1 and \mathcal{C}_2 , we have that the operator D is positive definite with eigenvalues $b_{n,d,1}/b_{n,d,2} > 0$. Since the operator D has eigenvalues greater than 0 and the operator $D - I$ is diagonal, then its eigenvalues (singular values) are greater than -1 . Then by (2.8), we conclude that $D - I$ is Hilbert-Schmidt operator. This completes the proof. \square

Since the Hilbert space $L^2(\mathbb{S}^d)$ is a separable Hilbert space, let μ_i , $i = 1, 2$ be two zero-mean Gaussian measures with covariance operators \mathcal{C}_i , $i = 1, 2$. Define the ratio between \mathcal{C}_1 and \mathcal{C}_2 as $\Lambda = \frac{\langle X(\mathbf{x}), \mathcal{C}_2 X(\mathbf{x}) \rangle}{\langle X(\mathbf{x}), \mathcal{C}_1 X(\mathbf{x}) \rangle}$, for any $X(\mathbf{x}) \in L^2(\mathbb{S}^d)$. Assume $K = \sup_{X(\mathbf{x}) \in L^2(\mathbb{S}^d)} \Lambda_{X(\mathbf{x})}$ and $k = \inf_{X(\mathbf{x}) \in L^2(\mathbb{S}^d)} \Lambda_{X(\mathbf{x})}$. Then we have following result.

Lemma 2.2.4. *Let μ_1 and μ_2 be equivalent zero-mean Gaussian measures on \mathbb{S}^d . Then, $0 < k < K < \infty$.*

Proof. According to Theorem 2.2.1, there exists an operator $D = \mathcal{C}_2^{-1/2} \mathcal{C}_1 \mathcal{C}_2^{-1/2}$ such that D is positive definite, invertible, and bounded. Thus, we can rewrite the covariance operator \mathcal{C}_1 as $\mathcal{C}_1 = \mathcal{C}_2^{1/2} D \mathcal{C}_2^{1/2}$. Hence, the ratio between \mathcal{C}_1 and \mathcal{C}_2 is given by

$$\Lambda = \frac{\langle X(\mathbf{x}), \mathcal{C}_2 X(\mathbf{x}) \rangle}{\langle X(\mathbf{x}), \mathcal{C}_1 X(\mathbf{x}) \rangle} = \frac{\|\mathcal{C}_2^{1/2} X(\mathbf{x})\|^2}{\|D^{1/2} \mathcal{C}_2^{1/2} X(\mathbf{x})\|^2}.$$

Since D is invertible and bounded, then $\|D\|^{-1} \leq \Lambda_{X(\mathbf{x})} \leq \|D^{-1}\|$, i.e., $0 < k < K < \infty$. \square

2.3 Selected Examples

We now apply the necessary and sufficient condition proposed in Section 2.2 to study the equivalence and orthogonality of zero-mean Gaussian measures associated with some particularly given covariance functions.

2.3.1 Multiquadric Covariance Function

Let $p_i \in (0, 1)$ and $\tau_i > 0$, $i = 1, 2$ be the parameters of the negative Binomial distribution on \mathbb{S}^d (Møller et al., 2015), given by

$$\beta_n(p_i, \tau_i) = \sigma_i^2 \binom{\tau_i + n - 1}{n} p_i^n (1 - p_i)^{\tau_i}, \quad n = 0, 1, \dots$$

The corresponding covariance obtained through the representation (2.6) in \mathbb{S}^∞ is

$$\psi_i(\theta) = \sigma_i^2 \left(\frac{1 - p_i}{1 - p_i \cos \theta} \right)^{\tau_i}, \quad 0 \leq \theta \leq \pi. \quad (2.11)$$

If we assume that $p_i = \frac{2\delta_i}{1+\delta_i^2}$ with $\delta_i \in (0, 1)$, then the previous representation of the covariance will take the following expression (Gneiting, 2013a),

$$\psi_i(\theta) = \psi(\theta; \delta_i, \tau_i, \sigma_i) = \sigma_i^2 \frac{(1 - \delta_i)^{2\tau_i}}{(1 + \delta_i^2 - 2\delta_i \cos \theta)^{\tau_i}}. \quad (2.12)$$

Let $d \geq 2$ and $\tau_i = \frac{d-1}{2}$. So the d -Schoenberg coefficients have the form

$$b_{n,d,i} = \sigma_i^2 \binom{d+n-2}{n} \delta_i^n (1 - \delta_i)^{d-1}, \quad i = 1, 2. \quad (2.13)$$

The following proposition discusses the orthogonality condition of any two zero-mean Gaussian measures on \mathbb{S}^2 with a Multiquadric covariance function.

Proposition 2.3.1. *Let $Z_i(\mathbf{x})$, $i = 1, 2$ be zero-mean Gaussian random fields on \mathbb{S}^2 with Multiquadric covariance functions $\psi_i(\theta) = \psi(\theta; \delta_i, \tau, \sigma_i)$, $i = 1, 2$ as in (2.12) with $\tau = \frac{d-1}{2}$. Let μ_i , $i = 1, 2$, be two different Gaussian measures with Schoenberg coefficients $b_{n,d,i}$ as in (2.13). Then μ_1 and μ_2 are equivalent if and only if $\sigma_1^2 = \sigma_2^2$ and $\delta_1 = \delta_2$.*

Proof. On \mathbb{S}^2 , $h(n, 3) = 2n + 1$ and the Schoenberg coefficients (2.13) have the form

$$b_{n,2,i} = \sigma_i^2 \delta_i^n (1 - \delta_i), \quad i = 1, 2. \quad (2.14)$$

By Proposition 2.2.3, to prove the orthogonality of μ_1 and μ_2 we need to show that the series in (2.8) is divergent. For our case, the series (2.8) is of the form

$$\sum_{n=0}^{\infty} h(n, 3) \left(\frac{b_{n,2,1}}{b_{n,2,2}} - 1 \right)^2 = \sum_{n=0}^{\infty} (2n+1) \left(\left(\frac{\sigma_1}{\sigma_2} \right)^2 \left(\frac{\delta_1}{\delta_2} \right)^n \left(\frac{1-\delta_1}{1-\delta_2} \right) - 1 \right)^2. \quad (2.15)$$

To check the convergence or divergence of this series, we use Raabe's test as follows.

Step1. Consider

$$a_n = (2n+1) \left(\left(\frac{\sigma_1}{\sigma_2} \right)^2 \left(\frac{\delta_1}{\delta_2} \right)^n \left(\frac{1-\delta_1}{1-\delta_2} \right) - 1 \right)^2,$$

$$a_{n+1} = (2n+3) \left(\left(\frac{\sigma_1}{\sigma_2} \right)^2 \left(\frac{\delta_1}{\delta_2} \right)^{n+1} \left(\frac{1-\delta_1}{1-\delta_2} \right) - 1 \right)^2.$$

Step2. Now we need to evaluate the following term

$$n \left[\frac{a_n}{a_{n+1}} - 1 \right] = n \left[\frac{(2n+1) \left(\left(\frac{\sigma_1}{\sigma_2} \right)^2 \left(\frac{\delta_1}{\delta_2} \right)^n \left(\frac{1-\delta_1}{1-\delta_2} \right) - 1 \right)^2}{(2n+3) \left(\left(\frac{\sigma_1}{\sigma_2} \right)^2 \left(\frac{\delta_1}{\delta_2} \right)^{n+1} \left(\frac{1-\delta_1}{1-\delta_2} \right) - 1 \right)^2} - 1 \right]. \quad (2.16)$$

Step3. To find the limit of (2.16), we have two different cases. The first one, if

$\delta_1 < \delta_2$, then $\left(\frac{\delta_1}{\delta_2} \right)^n$ and $\left(\frac{\delta_1}{\delta_2} \right)^{n+1}$ tend to zero as $n \rightarrow \infty$. Hence,

$$\begin{aligned} \lim_{n \rightarrow \infty} n \left[\frac{a_n}{a_{n+1}} - 1 \right] &= \lim_{n \rightarrow \infty} n \left[\frac{(2n+1) \left(\left(\frac{\sigma_1}{\sigma_2} \right)^2 \left(\frac{\delta_1}{\delta_2} \right)^n \left(\frac{1-\delta_1}{1-\delta_2} \right) - 1 \right)^2}{(2n+3) \left(\left(\frac{\sigma_1}{\sigma_2} \right)^2 \left(\frac{\delta_1}{\delta_2} \right)^{n+1} \left(\frac{1-\delta_1}{1-\delta_2} \right) - 1 \right)^2} - 1 \right] \\ &= \lim_{n \rightarrow \infty} n \left[\frac{2n+1}{2n+3} - 1 \right] = \lim_{n \rightarrow \infty} \left[\frac{-2n}{2n+3} \right] \\ &= -1 < 1. \end{aligned}$$

Then Raabe's test yields divergence, meaning that the Gaussian measures μ_1 and μ_2 are orthogonal. For the second case, if $\delta_1 > \delta_2$, then $\left(\frac{\delta_1}{\delta_2} \right)^{n+1}$

tends to infinity faster than $\left(\frac{\delta_1}{\delta_2}\right)^n$ as $n \rightarrow \infty$. Hence,

$$\begin{aligned} \lim_{n \rightarrow \infty} n \left[\frac{a_n}{a_{n+1}} - 1 \right] &= \lim_{n \rightarrow \infty} n \left[\frac{(2n+1) \left(\left(\frac{\sigma_1}{\sigma_2} \right)^2 \left(\frac{\delta_1}{\delta_2} \right)^n \left(\frac{1-\delta_1}{1-\delta_2} \right) - 1 \right)^2}{(2n+3) \left(\left(\frac{\sigma_1}{\sigma_2} \right)^2 \left(\frac{\delta_1}{\delta_2} \right)^{n+1} \left(\frac{1-\delta_1}{1-\delta_2} \right) - 1 \right)^2} - 1 \right] \\ &= \lim_{n \rightarrow \infty} (-n) = -\infty. \end{aligned}$$

Then the series (2.15) is always divergent. The proof is completed. \square

2.3.2 The Sine Power Covariance Functions

The sine power covariance function is given by

$$\psi(\theta; \alpha, \sigma) = \sigma^2 \left[1 - \left(\sin \frac{\theta}{2} \right)^\alpha \right], \quad 0 \leq \theta < 2\pi, \quad (2.17)$$

where $\alpha \in (0, 2)$, and $\sigma^2 > 0$ are parameters (Soubeyrand et al., 2008; Gneiting, 2013a). Let μ_1 and μ_2 be zero-Gaussian measures on the circle \mathbb{S}^1 with corresponding covariance functions $\psi_i = \psi(\theta; \alpha_i, \sigma_i)$, $i = 1, 2$ as in (2.17). The following proposition discusses the orthogonality condition of μ_1 and μ_2 .

Proposition 2.3.2. *Let $Z_i(\mathbf{x})$, $i = 1, 2$ be zero-mean Gaussian random fields on \mathbb{S}^1 with covariance functions $\psi_i = \psi(\theta; \alpha_i, \sigma_i)$, $i = 1, 2$ as in (2.17). Let μ_i , $i = 1, 2$ be Gaussian measures associated to $Z_i(\mathbf{x})$, $i = 1, 2$. Then, μ_1 and μ_2 are equivalent if and only if $\alpha_1 = \alpha_2$, $\sigma_1^2 = \sigma_2^2$.*

Proof. The circular covariance function given by (2.17) has the following Schoenberg coefficients (Soubeyrand et al., 2008),

$$b_{n,1,i} = \frac{\sigma_i^2}{\sqrt{2}} \Gamma_{n+1} \left(\frac{\alpha_i}{2} \right), \quad \Gamma_{n+1} \left(\frac{\alpha_i}{2} \right) = \frac{-1}{(n+1)!} \prod_{m=0}^n \left(m - \frac{\alpha_i}{2} \right), \quad i = 1, 2. \quad (2.18)$$

According to Proposition 2.2.3, we have to check the convergence of the series in (2.8) to determine if the given Gaussian measures μ_1 and μ_2 are equivalent or

orthogonal. So, for our case the series in (2.8) takes the following form

$$\sum_{n=0}^{\infty} h(n, 2) \left(\frac{b_{n,1,1}}{b_{n,1,2}} - 1 \right)^2 = \sum_{n=0}^{\infty} 2 \left(\Lambda \prod_{m=0}^n \left(\frac{2m - \alpha_1}{2m - \alpha_2} \right) - 1 \right)^2, \quad (2.19)$$

where $\Lambda = \left(\frac{\sigma_1}{\sigma_2} \right)^2$. We will use Raabe's test to check the convergence of (2.19).

Step1. Now a_n and a_{n+1} are given as follows

$$a_n = \left(\Lambda \prod_{m=0}^n \left(\frac{2m - \alpha_1}{2m - \alpha_2} \right) - 1 \right)^2, \quad a_{n+1} = \left(\Lambda \prod_{m=0}^{n+1} \left(\frac{2m - \alpha_1}{2m - \alpha_2} \right) - 1 \right)^2.$$

Step2. Next we evaluate the following term

$$n \left[\frac{a_n}{a_{n+1}} - 1 \right] = n \left[\left(\frac{\Lambda \prod_{m=0}^n \left(\frac{2m - \alpha_1}{2m - \alpha_2} \right) - 1}{\Lambda \prod_{m=0}^{n+1} \left(\frac{2m - \alpha_1}{2m - \alpha_2} \right) - 1} \right)^2 - 1 \right]. \quad (2.20)$$

Step3. To find the limit of (2.20), we have two different cases to study the limit

of (2.20). The first case, if $\alpha_1 > \alpha_2$, then $\frac{2m - \alpha_1}{2m - \alpha_2} < 1$. Moreover, we have

$$\lim_{n \rightarrow \infty} n \left[\frac{a_n}{a_{n+1}} - 1 \right] = \lim_{n \rightarrow \infty} n \left[\left(\frac{\Lambda \prod_{m=0}^n \left(\frac{2m - \alpha_1}{2m - \alpha_2} \right) - 1}{\Lambda \prod_{m=0}^{n+1} \left(\frac{2m - \alpha_1}{2m - \alpha_2} \right) - 1} \right)^2 - 1 \right] = 0 < 1.$$

Thus, the series (2.19) is divergent. The second case, if $\alpha_1 < \alpha_2$, then

$\frac{2m - \alpha_2}{2m - \alpha_1} < 1$. Moreover, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} n \left[\frac{a_n}{a_{n+1}} - 1 \right] &= \lim_{n \rightarrow \infty} n \left[\left(\frac{\Lambda \prod_{m=0}^n \left(\frac{2m - \alpha_1}{2m - \alpha_2} \right) - 1}{\Lambda \prod_{m=0}^{n+1} \left(\frac{2m - \alpha_1}{2m - \alpha_2} \right) - 1} \right)^2 - 1 \right] \\ &= \lim_{n \rightarrow \infty} n \left[\left(\frac{\Lambda - \prod_{m=0}^n \left(\frac{2m - \alpha_2}{2m - \alpha_1} \right)}{\Lambda \frac{2n+1 - \alpha_1}{2n+1 - \alpha_2} - \prod_{m=0}^n \left(\frac{2m - \alpha_2}{2m - \alpha_1} \right)} \right)^2 - 1 \right] \\ &= \lim_{n \rightarrow \infty} n \left[\left(\frac{2n+1 - \alpha_2}{2n+1 - \alpha_1} \right)^2 - 1 \right] \\ &= \lim_{n \rightarrow \infty} n \left[\frac{2(2n+1)(\alpha_1 - \alpha_2) + (\alpha_2^2 - \alpha_1^2)}{(2n+1)^2 - 2(2n+1)\alpha_1 + \alpha_1^2} \right] \end{aligned}$$

$$= \alpha_1 - \alpha_2 < 1.$$

Thus, the series (2.19) is divergent.

Since the series is always divergent, then the Gaussian measures μ_1 and μ_2 are always orthogonal. \square

2.4 A Simulation Study

We explore numerically some consequences of the theoretical results in Sections 2.2 and 2.3 on the estimation of some specific covariance models of Gaussian random fields defined on the sphere. Specifically, we focus on maximum likelihood estimation.

A consequence of Proposition 2.3.1 is that when estimating the parameters of the Multiquadric model

$$\psi_M(\theta) = \sigma_M^2 \frac{(1 - \delta)^{2\tau}}{(1 + \delta^2 - 2\delta \cos \theta)^\tau}, \quad \theta \in [0, \pi] \quad (2.21)$$

on \mathbb{S}^2 , then the parameters σ_M^2 and δ are both consistently estimable under fixed domain asymptotic

A consequence of Proposition 2.3.2 is that when estimating the parameters of the Sine power model

$$\psi_S(\theta) = \sigma_S^2 \left[1 - \left(\sin \frac{\theta}{2} \right)^\alpha \right], \quad \theta \in [0, \pi] \quad (2.22)$$

on \mathbb{S}^1 , then the parameters σ_S^2 and α are both consistently estimable under fixed domain asymptotic

In our simulation setting we work on \mathbb{S}^2 and, in addition to the Multiquadric and Sine power models, we also consider the exponential covariance model with the

geodesic distance

$$\psi_E(\theta) = \sigma_E^2 \exp\left(-\frac{\theta}{b}\right), \quad (2.23)$$

where $\theta \in [0, \pi]$ and $\sigma_E^2, b > 0$.

Results in [Zhang \(2004\)](#) guarantee that when considering the Euclidean distance on \mathbb{R}^d , $d = 1, 2, 3$, σ_E^2 and b are not consistently estimable under fixed domain asymptotic. We explore the behavior of the maximum likelihood estimation of both parameters on \mathbb{S}^2 using the geodesic distance.

We first consider some details on the computation on the great circle (GC) on \mathbb{S}^2 . For two location sites in longitude and latitude (expressed in decimal degrees), $P_i = (\text{lon}_i, \text{lat}_i)^T$ and $P_j = (\text{lon}_j, \text{lat}_j)^T$ and the radius of the sphere denoted by R , the great circle distance is given by $R\theta$ where

$$\begin{aligned} \theta &:= \theta(\mathbf{x}_i, \mathbf{x}_j) = \arccos(\langle \mathbf{x}_i, \mathbf{x}_j \rangle) \\ &= [\arccos\{\sin a_i \sin a_j + \cos a_i \cos a_j \cos(b_i - b_j)\}]. \end{aligned}$$

Here $a_i = (\text{lat}_i)\pi/180$, $a_j = (\text{lat}_j)\pi/180$, $b_i = (\text{lon}_i)\pi/180$, $b_j = (\text{lon}_j)\pi/180$, $\mathbf{x}_i = (a_i, b_i)^T$ and $\mathbf{x}_j = (a_j, b_j)^T$. Here we work, without loss of generality, on the unit sphere (i.e, $R = 1$).

In order to check for consistency of the parameters we first consider 2000 locations uniformly distributed on the surface of the unit sphere and we then consider an increasing sequence $n = 100, 200, 400, 800, 1200, 1600, 2000$ of location sites randomly chosen without replacement from the previous set points (see [Figure 3.2](#) for the cases $n = 100, 800, 2000$). For each n we simulate, through Cholesky decomposition, 1000 zero mean Gaussian random fields with covariance [\(2.21\)](#) with $\sigma_M^2 = 1$, $\delta = 0.95$, covariance [\(2.22\)](#) with $\sigma_S^2 = 1$, $\alpha = 0.74$ and covariance [\(2.23\)](#) with $\sigma_E^2 = 1$ and $b = 1/3$.

The covariances are reported in [Figure 3.2](#). Note that, under this specific parameterization, they have a common practical range approximately equal to 1. For

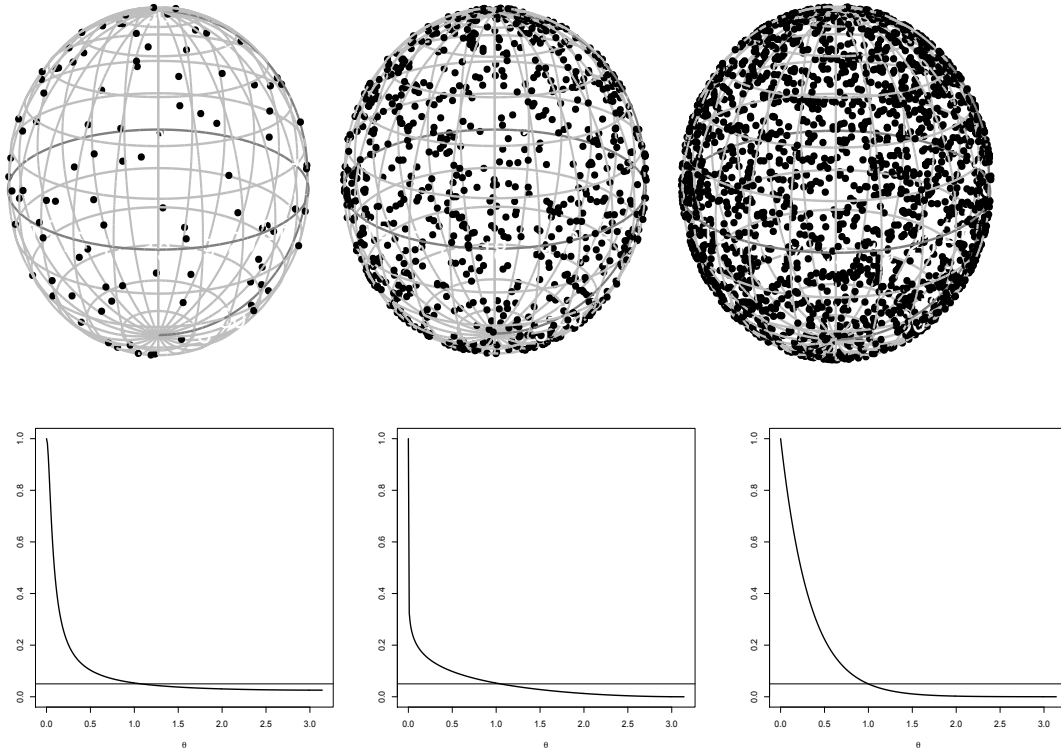


FIGURE 2.1. Top: from the left to the right, increasing sequence of location sites on the unit sphere ($n = 100, 800, 2000$) considered in the simulation study. Bottom: from left to right, the Multiquadric, Sine power and exponential correlation functions considered in the simulation study.

each n , for each model and for each simulation, we estimate through maximum likelihood the parameters $(\sigma_M^2, \delta)^T$ in the first case, $(\sigma_S^2, \alpha)^T$ in the second case, and $(\sigma_E^2, b)^T$ in the exponential case.

In order to check for consistency of the parameters we look at the sample variance of the maximum likelihood estimates when increasing $n = 100, 400, \dots, 2000$. In order to take into account the different order of magnitude of the variances of each parameter we consider a relative sample variance that is for each n , we first consider the sample variances of the estimates and then we divided them by the maximum (over n) of these sample variances. Since the sample variance is not increasing with n when $n = 100$ the relative sample variance is equal to 1, and then it does not increase with n . Figure 2.2 (left part) shows for each model how the sample relative variance of the ML estimates of the scale parameters δ , α and b decreases when the number of location sites increases. Figure 2.2 (right

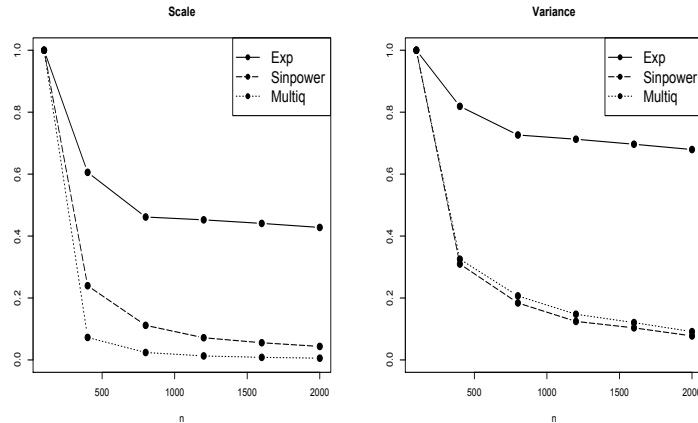


FIGURE 2.2. Relative sample variance of the maximum likelihood estimates of the scale (left) and variance (right) parameters when increasing the number of location sites on a fixed region of the unit sphere for the Exponential, Sine power and Multiquadric covariance models.

part) shows for each model how the sample variance of the maximum likelihood estimates of σ_x^2 , $x = M, S, E$ decreases when increasing the number of location sites increases.

From this example, it becomes apparent that sampling more data on \mathbb{S}^2 may not improve the jointly estimation of the scale and variance parameters when using the exponential model. This suggests that scale and variance parameters are not jointly estimable on \mathbb{S}^2 . Similar simulation results have been obtained in (Zhang, 2004) when working on \mathbb{R}^2 . For the Multiquadric covariance model, as expected from Propositions 2.3.1, there is a clear pattern of decreasing (relative) sample variance when jointly estimating the scale and variances parameters. For the Sine power model, even if Proposition 2.3.2 is valid on \mathbb{S}^1 , our simulation results suggest that orthogonality is still valid on \mathbb{S}^2 .

2.5 Conclusions

We have studied equivalence of Gaussian measures associated to random fields over spheres. Our results show that the analogue of Skorohod (1967) conditions

still hold over spheres, but this time in terms of convergence of the ratios of d -Schoenberg coefficients.

An important fact is that for two celebrated parametric classes of covariances on spheres (the Sine power and the Multiquadric) we found that Gaussian measures related to them are always orthogonal, with the important implications in terms of ML estimation that have been illustrated through Section 2.4. We found similar results for other parametric classes of covariances defined over spheres.

[Gneiting \(2013a\)](#) has shown that, under some additional restrictions on the parameters, some parametric families of covariance functions being valid on Euclidean spaces can be used with the great circle distance and thus adapted to the sphere. This is the case for the Generalized Cauchy, the Dagum, the Wendland and the Power Exponential models. The Matérn covariance ([Matérn, 1960](#)) can be adapted at the expense of a severe restriction on the smoothing parameter, which makes it not very appealing on the sphere.

Calculation of the d -Schoenberg coefficients for these models is in general unfeasible. Some exceptions can be found in [Møller et al. \(2015\)](#), where it is shown that the 1-Schoenberg coefficients associated to Wendland functions have oscillating behaviors. This makes extremely difficult to apply our results for our cases and opens future researches oriented to finding other criteria for equivalence of Gaussian measures over the sphere. This is the case for the Generalized Cauchy, the Dagum, the Wendland and the Power Exponential models. The Matérn covariance ([Matérn, 1960](#)) can be adapted at the expense of a severe restriction on the smoothing parameter, which makes it not very appealing on the sphere.

Calculation of the d -Schoenberg coefficients for these models is in general unfeasible. Some exceptions can be found in [Møller et al. \(2015\)](#), where it is shown that the 1-Schoenberg coefficients associated to Wendland functions have oscillating behaviors. This makes extremely difficult to apply our results for our cases and opens future researches oriented to finding other criteria for equivalence of Gaussian measures over spheres.

CHAPTER 3

Gneiting's Problems and the Class Ψ_d of Positive Definite Functions over Hyperspheres

In this chapter we provide with partial answers to Problems 1 and 2 proposed in [Gneiting \(2013b\)](#) on isotropic positive definite functions on spheres: find an expression for any general d -Schoenberg coefficient in terms of Fourier coefficients on the circle, find conditions in order to transfer from the Euclidean space to the sphere setting, and give a lower bound for the minimal curvature of functions vanishing beyond any given range. In the final section, we show expressions for the 2-Schoenberg coefficients of functions in the exponential and Askey's families (completing the ones given in [Møller et al. \(2015\)](#)).

This result might have led to another methodology for building new isotropic covariance models in \mathbb{S}^d (just by choosing appropriate sequences of non negative coefficients with convergent series and forming the Gegenbauer expansion of such sequence). However, the lack of links between the properties of the coefficients and the properties of the model prevents practitioners to use it nowadays. In

this direction, (Gneiting, 2013a) has given sufficient conditions on the sequence of coefficients in order to yield strictly positive definiteness (which is useful for interpolation of scattered data in spherical domains). This result motivates the search of explicit formulas for the coefficients of well-known models, and our Section 3.3 goes in that direction.

In the sequel, coefficients a_n in the Gegenbauer series expansion shall be referred to as the *n-th d-Schoenberg coefficient* (as coined in Daley and Porcu (2014)), and rephrased as $b_{n,d}$, in order to stress the dependence on the dimension d . We shall also restrict to members ψ of classes Ψ_d , sticking to notation in Gneiting (2013a), so that the series expansion is:

$$\psi(\theta) = \sum_{n=0}^{\infty} b_{n,d} \frac{C_n^\lambda(\cos \theta)}{C_n^\lambda(1)}, \quad (3.1)$$

where $b_{n,d}$ defined by (1.44).

3.1 The class Ψ_d and d -Schoenberg coefficients

Throughout, we shall be sloppy whenever using the abuse of notation θ for the geodesic distance $\theta(\xi, \eta)$ between $\xi, \eta \in \mathbb{S}^d$. We also consider the Hilbert sphere $\mathbb{S}^\infty = \{x \in \mathbb{R}^{\mathbb{N}} : \|x\| = 1\}$. We denote C_n^λ the n -th Gegenbauer polynomial of order $\lambda > 0$, uniquely identified through the intrinsic relation

$$\frac{1}{(1 + r^2 - 2r \cos \theta)^\lambda} = \sum_{n=0}^{\infty} r^n C_n^\lambda(\cos \theta), \quad \theta \in [0, \pi],$$

where $r \in (-1, 1)$. It is of fundamental importance that

$$\left| C_n^\lambda(x) \right| \leq \frac{\Gamma(n + 2\lambda)}{n! \Gamma(2\lambda)} = C_n^\lambda(1), \quad x \in [-1, 1].$$

The following trigonometric representation for Gegenbauer polynomials, see (Szegő, 1939, p. 93, Equation 4.9.22) will turn to be useful subsequently: for

any $\lambda > 0$, $\lambda \notin \mathbb{N}$ and $\theta \in [0, \pi]$,

$$C_n^\lambda(\cos \theta) = \frac{2^{2-2\lambda}\Gamma(n+2\lambda)}{\Gamma(\lambda)\Gamma(n+\lambda+1)} \sum_{\mu=0}^{\infty} \frac{(1-\lambda)_\mu(n+1)_\mu \sin(n+2\mu+1)\theta}{\mu!(n+\lambda+1)_\mu (\sin \theta)^{2\lambda-1}}, \quad (3.2)$$

where $(x)_m := x(x+1)\cdots(x+m-1)$ denotes the rising factorial operator for any non negative integer m with $(x)_0 = 1$.

The case $\lambda \in \mathbb{N}$ has been elusive so far, and we obtain it here in a technical result that will be of use in the following sections.

Lemma 3.1.1. *Let $\lambda \in \mathbb{N}$. Then, it is true that*

$$C_n^\lambda(\cos \theta) = \frac{2^{2-2\lambda}\Gamma(n+2\lambda)}{\Gamma(\lambda)\Gamma(n+\lambda+1)} \sum_{\mu=0}^{\lambda-1} \frac{(1-\lambda)_\mu(n+1)_\mu \sin(n+2\mu+1)\theta}{\mu!(n+\lambda+1)_\mu (\sin \theta)^{2\lambda-1}}. \quad (3.3)$$

Proof. Let μ be a nonnegative integer. Define

$$(1-\lambda)_\mu = (1-\lambda)(2-\lambda)\cdots(\mu-\lambda-1)(\mu-\lambda),$$

where $\lambda \in \mathbb{N}$. Thus, for all $\mu \geq \lambda$, the rising factorial will vanish. Moreover, the summation in (3.3) equals zero for $\mu \geq \lambda$. The proof is completed. \square

An alternative proof of the previous result can be provided through induction on $\lambda \in \mathbb{N}$. When $\lambda = (d-1)/2$, Gegenbauer polynomials have an explicit relationship with spherical harmonics, being orthonormal basis of $L^2(\mathbb{S}^d, d(\mathbb{S}^d))$, with $d(\mathbb{S}^d)$ is the surface measure, having total mass equal to $2\pi^{\frac{d+1}{2}}/\Gamma(\frac{d+1}{2})$.

Let Ψ_d be the class of continuous mappings $\psi : [0, \pi] \rightarrow \mathbb{R}$ with $\psi(0) = 1$ such that the continuous functions $C : \mathbb{S}^d \times \mathbb{S}^d \rightarrow \mathbb{R}$ defined through $C(\xi, \eta) = \psi(\theta(\xi, \eta))$ are positive definite. Theorem 1.3.9 characterized the positive definite functions defined on the spheres of any dimension.

Following Corollary 1 in (Gneiting, 2013a) the classes Ψ_d are nested, with the inclusion relation

$$\Psi_1 \supset \Psi_2 \supset \cdots \supset \Psi_\infty := \bigcap_{d \geq 1} \Psi_d$$

being strict, and where Ψ_∞ has as direct relation to the Hilbert sphere as previously defined.

A useful way to rephrase Schoenberg's result is the following, and we state it formally because it will turn to be useful for the subsequent exposition. [Gneiting \(2013a\)](#) and [Beatson et al. \(2013\)](#) obtain recurrent formulas in order to write coefficient $b_{n,d}$ as a linear combination of $b_{n,d-2}$ and $b_{n+2,d-2}$. By applying recursivity, each coefficient $b_{n,d}$ can be finally written, when d is odd, as a linear combination of Schoenberg coefficients in the circle, $\{b_{n+2k,1}\}_{k=0}^{\lfloor d/2 \rfloor}$, and when d is even, as a linear combination of Schoenberg coefficients in the sphere, $\{b_{n+2k,2}\}_{k=0}^{d/2}$.

Arguments in [Gneiting \(2013a\)](#) and classical Fourier inversion show that the d -Schoenberg coefficients are defined by (1.44). We recall that 1-Schoenberg coefficients are the Fourier coefficients for even functions:

$$b_{0,1} := \frac{1}{\pi} \int_0^\pi f(\cos \theta) d\theta, \quad b_{n,1} := \frac{2}{\pi} \int_0^\pi f(\cos \theta) \cos(n\theta) d\theta, \quad (n \geq 1). \quad (3.4)$$

3.2 Gneiting's problems with their solutions

3.2.1 Statements of the Problems

We now expose the problems faced in this chapter together with their partial solutions.

Problem 1. ([Gneiting, 2013b, Problem 1](#)) Let $n \geq 0$ and $k \geq 1$ be integers. Find the coefficients $a_{n,1}, \dots, a_{n,k}$ in the expansion

$$b_{n,2k+1} = \sum_{i=0}^k a_{n,i} b_{n+2i,1} \quad (3.5)$$

associated to the $(2k+1)$ -Schoenberg coefficients in terms of Fourier coefficients $b_{n,1}, \dots, b_{n+2k,1}$. Similarly, find the $(2k+2)$ -Schoenberg coefficients in terms of the 2-Schoenberg coefficients $b_{n,2}, b_{n+2,2}, \dots, b_{n+2k,2}$.

In order to state Problem 2, we follow [Gneiting \(2013a\)](#) when calling Ψ_d^c the subclass of Ψ_d having members ψ that vanish for any $\theta \geq c$, with $c \in (0, \pi]$. When $c < \pi$, then any member of Ψ_d^c has local support, otherwise it is called *globally supported*.

Problem 2. ([Gneiting, 2013b](#), Problem 3) For an integer $d \geq 1$, and for a given $c \in (0, \pi]$, find

$$a_d^c := \inf_{\psi \in \Psi_d^c} (-\psi''(0)). \quad (3.6)$$

This is a problem of applied interest when $d = 2$. In atmospheric data assimilation, locally supported isotropic correlation functions are used for the distance-dependent reduction of global scale covariance estimates in ensemble Kalman filter settings (see [Gneiting \(2013a\)](#), and the references therein). Thus, it is appealing to use a member of the class Ψ_2^c with minimal curvature at the origin.

Some comments are in order. The solution of Problem 1 requires the use of recursive formulae for the Gegenbauer polynomials and a constructive argument that will be exposed subsequently. An approach of Problem 2 relies on considering $\tilde{\Psi}_d^c$, the subclass of Ψ_d given by those members $\psi \in \Psi_d$ such that $\psi(c) = 0$. Clearly, we have

$$\Psi_d^c \subset \tilde{\Psi}_d^c \subset \Psi_d, \quad (3.7)$$

with the inclusion relation being strict. The definition of the class $\tilde{\Psi}_d^c$ in concert with Schoenberg's representation and the oscillatory nature of Gegenbauer polynomial implies that, for any member of the class Ψ_d , there exists a collection of members $\tilde{\psi}$ of the class $\Psi_d^{c_k}$, for $\{c_k\}$ being a sequence of constants with $c_k \in (0, \pi]$, such that

$$\psi(\theta) = \sum_{k=0}^{\infty} b_{k,d} \tilde{\psi}_k(\theta), \quad \theta \in [0, \pi].$$

Another relevant comment is that Theorems 2 and 3 in [Gneiting \(2013a\)](#) provide the upper bound $a_d^c \leq \frac{1}{c^2} \frac{4}{d} j_{\frac{d-2}{2}}^2$, where j_ν denotes the first positive zero of the

Bessel function J_ν . Some of these zeros are:

$$j_0 \approx 2.4048, \quad j_{0.5} \approx \pi \quad j_1 \approx 3.8317.$$

According to [Ehm et al. \(2004\)](#) the constant a_d^c in Euclidean spaces depend on Boas-Kac roots, but [Ziegel \(2014\)](#) showed that the convolution root does not always exist for positive definite functions on spheres. This makes the problem mathematically more interesting, and certainly tricky.

3.2.2 Main results

Proposition 3.2.1. *Let $d > 1$ be an integer, and let $\lambda := (d - 1)/2$. Then,*

$$b_{n,d} = \frac{\sqrt{\pi}\Gamma(n+2\lambda)}{2^{2\lambda}\Gamma(\lambda+1/2)\Gamma(n+\lambda)} \left[b_{n,1}^* - \lambda \sum_{\mu=1}^{\infty} \frac{(1-\lambda)_{\mu-1}(n+1)_{\mu-1}(n+2\mu)}{\mu!(n+\lambda+1)_{\mu}} b_{n+2\mu,1} \right]$$

for $n \geq 0$, where $b_{n,1}^* = b_{n,1}$ when $n \geq 1$ and $b_{0,1}^* = 2b_{0,1}$. If d is odd, the expression involves only a finite number of coefficients, i.e., $b_{n,1}^*, b_{n+2,1}, \dots, b_{n+2\lambda,1}$.

Proof. By plugging Equation (3.2) into (1.44), we get

$$b_{n,d} = \alpha(\lambda, n) \int_0^\pi \sum_{\mu=0}^{\infty} \beta(\lambda, n, \mu) \frac{\sin(n+2\mu+1)\theta}{(\sin\theta)^{2\lambda-1}} \psi(\theta) (\sin\theta)^{2\lambda} d\theta, \quad (3.8)$$

where $\alpha(\lambda, n) = \frac{(n+\lambda)2^{2-2\lambda}\Gamma(n+2\lambda)\Gamma(\lambda)}{\Gamma(\lambda)\Gamma(n+\lambda+1)\Gamma(\lambda+1/2)\Gamma(1/2)}$ and $\beta(\lambda, n, \mu) = \frac{(1-\lambda)_{\mu}(n+1)_{\mu}}{\mu!(n+\lambda+1)_{\mu}}$. Thus,

$$b_{n,d} = \frac{\alpha(\lambda, n)}{2} \sum_{\mu=0}^{\infty} \beta(\lambda, n, \mu) \int_0^\pi \psi(\theta) [\cos(n+2\mu)\theta - \cos(n+2\mu+2)\theta] d\theta.$$

Direct inspection on the expression above, in concert with Equation (3.4) for the Fourier 1-Schoenberg coefficients gives

$$b_{n,d} = \frac{\pi\alpha(\lambda, n)}{4} \left[\beta(\lambda, n, 0)[b_{n,1}^* - b_{n+2,1}] + \sum_{\mu=1}^{\infty} \beta(\lambda, n, \mu)[b_{n+2\mu,1} - b_{n+2\mu+2,1}] \right],$$

where $b_{n,1}^* = 2b_{0,1}$ and $b_{n,1}$ for $n \geq 1$. We can rewrite the previous representation as follows

$$b_{n,d} = \frac{\pi\alpha(\lambda, n)}{4} \left(\beta(\lambda, n, 0)b_{n,1}^* + b_{n+2,1}[\beta(\lambda, n, 1) - \beta(\lambda, n, 0)] \right. \\ \left. + b_{n+4,1}[\beta(\lambda, n, 2) - \beta(\lambda, n, 1)] + \dots \right).$$

Furthermore,

$$b_{n,d} = \frac{\pi\alpha(\lambda, n)}{4} \left(\beta(\lambda, n, 0)b_{n,1}^* + \sum_{\mu=1}^{\infty} b_{n+2\mu,1}[\beta(\lambda, n, \mu) - \beta(\lambda, n, \mu - 1)] \right), \quad (3.9)$$

where $\beta(\lambda, n, 0) = 1$ and

$$\beta(\lambda, n, \mu) - \beta(\lambda, n, \mu - 1) = \beta(\lambda, n, \mu - 1) \left[\frac{-\lambda(n + 2\mu)}{\mu(n + \lambda + \mu)} \right] \\ = \frac{-\lambda(1 - \lambda)_{\mu-1}(n + 1)_{\mu-1}(n + 2\mu)}{\mu!(n + \lambda + 1)_{\mu}}.$$

This fact completes the proof by plugging this last expression into Equation (3.9). \square

Remark. *The previous proposition applies for general $\lambda > 0$. In case that $\lambda \in \mathbb{N}$, i.e., d is odd, then the series becomes a finite linear combination, with μ ranging from 1 to $\lambda - 1$.*

We are now able to face Problem 2, where a formal statement for a partial solution is exposed in the following.

Proposition 3.2.2. *Let $d > 1$ be an integer. Then:*

- (i) $a_d^c \geq \frac{1}{1 - \cos c}$ if $c \in [\pi/2, \pi]$.
- (ii) $a_d^c \geq \frac{(d+1)(\cos c)(2 - \cos c) + 1}{(1 - \cos c)((d+1)\cos c + 1)}$ if $c \in [\arccos \sqrt{\frac{1}{d+1}}, \pi/2]$.

Proof. It is easy to check [Beatson et al. \(2013\)](#) that for any $\psi \in \Psi_d$ with associated d -Schoenberg coefficients $\{b_{n,d}\}_{n=0}^{\infty}$, it is true that $-\psi''(0) = \frac{1}{d} \sum_{n=1}^{\infty} n(n + d - 1)b_{n,d}$. Since the sequence $\{b_{n,d}\}_{n=0}^{\infty}$ is a probability mass system, functions ψ

with mass concentrated in earlier coefficients have a lower value of $-\psi''(0)$, and it is essential for the estimation of the infimum a_d^c in Equation (3.6).

The set Ψ_d^c is difficult to tackle, because locally supported functions have an infinite number of non null d -Schoenberg coefficients. In view of this, we consider Ψ_d^c as a subset of the more amenable set $\tilde{\Psi}_d^c := \{\psi \in \Psi_d : \psi(c) = 0\}$, of functions having at least one zero at the fixed value $\theta = c$. Now, denote

$$\tilde{a}_d^c := \inf_{\psi \in \tilde{\Psi}_d^c} [-\psi''(0)]. \quad (3.10)$$

Obviously, we have $a_d^c \geq \tilde{a}_d^c$ thanks to (3.7), and the latter value is attainable at a known function for a range of values of c , as we shall show. In order to get \tilde{a}_d^c we need to solve the pair of equations

$$\sum_{n=0}^{\infty} b_{n,d} = 1 \quad \text{and} \quad \sum_{n=0}^{\infty} b_{n,d} \frac{C_n^\lambda(\cos c)}{C_n^\lambda(1)} = 0 \quad (3.11)$$

subject to the restriction $\{b_{n,d}\}_{n=0}^{\infty} \subset [0, \infty)$. As already stated, we shall check the values for functions with mass concentrated into the first coefficients. The constant function (i.e. $b_{n,d} = 0$ for $n \geq 1$) is clearly out of $\tilde{\Psi}_d^c$. Thus, we check functions with $b_{n,d} = 0$ for $n \geq 2$. Using the equations in (3.11) we get the single function

$$\psi_c(\theta) = \frac{-\cos c}{1 - \cos c} + \frac{1}{1 - \cos c} \cos \theta,$$

and a sufficient condition for ψ to belong to the class $\psi \in \tilde{\Psi}_d^c$ is that $c \in [\pi/2, \pi]$, with $-\psi_c''(0) = \frac{1}{1 - \cos c}$. Hence for $c \in [\pi/2, \pi]$, $\tilde{a}_d^c = \frac{1}{1 - \cos c}$ and it is attained at function ψ_c .

For $c \in [0, \pi/2]$ we have no members of $\tilde{\Psi}_d^c$ with $b_{n,d} = 0$ for $n \geq 2$, and we shall look for functions with $b_{n,d} = 0$ for $n \geq 3$. Using again the system (3.11) we get the set of functions that can be written as

$$\psi_\beta(\theta) = -\frac{\cos c}{1 - \cos c} + \frac{(d+1)\cos c + 1}{d}\beta$$

$$+ \left(\frac{1}{1 - \cos c} - \frac{(d+1)(1 + \cos c)}{d} \beta \right) \cos \theta + \beta \frac{(d+1) \cos^2 \theta - 1}{d},$$

indexed by $\beta := b_{2,d}$. The non negativity restriction of their coefficients turns into the inequality

$$\frac{d \cos c}{(1 - \cos c)((d+1) \cos c + 1)} \leq \beta \leq \frac{d}{(d+1) \sin^2 c},$$

which leads to a non empty set of values only if $c \geq \arccos \sqrt{\frac{1}{d+1}}$, and \tilde{a}_d^c is attained at the lowest possible value of β satisfying the inequalities. \square

This strategy might lead to values of \tilde{a}_d^c for a wider range of values c , by using functions with $b_{n,d} = 0$ for $n \geq 4$, and so on, but we have not explored further this line because of the complexity of equations. Another way (yet unexplored) of improving the lower bounds is using slightly more complex auxiliary sets $\Psi_d^{(c,c')}$ of functions having at least two zeros, or even more. We could find no examples of members of this subclass.

Figure 3.1 depicts both upper and lower bounds for the range of c in dimension $d = 2$.

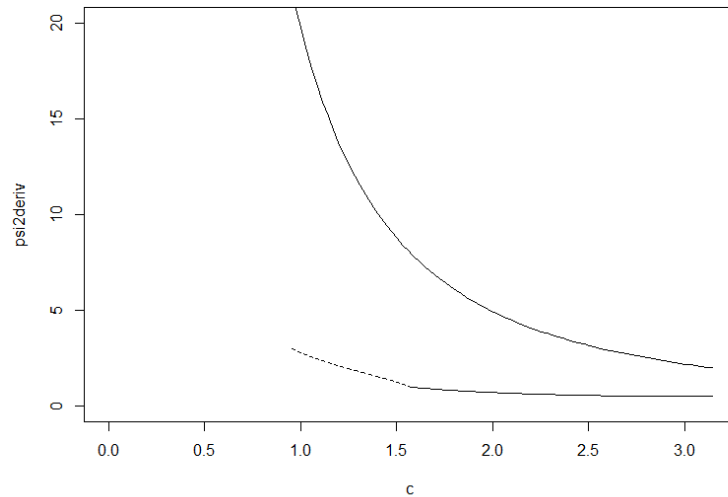


FIGURE 3.1. Upper (Ehm et al., 2004, Theorem5.1) and lower (Proposition 3.2.2) bounds for a_d^c in the range $c \in [\arccos \sqrt{\frac{1}{d+1}}, \pi]$ for $d = 2$.

3.3 On the 2-Schoenberg coefficients of some celebrated parametric families

This section inspects the problem of giving closed form expressions for the 2-Schoenberg coefficients of correlation functions in the exponential and Askey's families [Møller et al. \(2015\)](#).

A relevant remark is that what really matters is the computation of the 1- and 2-Schoenberg coefficients because all the others can then be calculated inductively by using Corollary 3 in [Gneiting \(2013a\)](#). In particular, using Theorem 4.2 in [Møller et al. \(2015\)](#) one can even get the Schoenberg's coefficients related to the representation of a given member of the class Ψ_∞ . Since the 1-Schoenberg coefficients for the exponential and Askey families have been provided in [Møller et al. \(2015\)](#), we focus here on the tricky case of the 2-Schoenberg coefficients related to these families.

First, we note that Gegenbauer polynomials simplify to Legendre polynomials P_n when dealing with \mathbb{S}^2 . Thus, classical Schoenberg's representation reduces to

$$\psi(\theta) = \sum_{n=0}^{\infty} b_{n,2} P_n(\cos \theta), \quad \theta \in [0, \pi], \quad (3.12)$$

where

$$b_{n,2} = \left(n + \frac{1}{2}\right) \int_0^\pi P_n(\cos \theta) \psi(\theta) \sin \theta d\theta,$$

for all $n \geq 0$. The following representation for Legendre polynomials turns to be useful [Dixit et al. \(2015\)](#)

$$P_n(\cos \theta) = 2^n \sum_{m=0}^n \binom{n}{m} \binom{\frac{n+m-1}{2}}{n} (\cos \theta)^m.$$

In view of the expression above, the 2-Schoenberg coefficients can be computed through

$$b_{n,2} = (2n+1)2^{n-1} \sum_{m=0}^n \binom{n}{m} \binom{\frac{n+m-1}{2}}{n} \int_0^\pi (\cos \theta)^m \sin \theta \psi(\theta) d\theta. \quad (3.13)$$

3.3.1 Exponential Family

Let us consider the exponential family of members of Ψ_∞ , given by

$$\psi_\alpha(\theta) = \exp\left(-\frac{\theta}{\alpha}\right), \quad \theta \in [0, \pi], \quad (3.14)$$

with α being a positive scaling parameter.

Proposition 3.3.1. *The 2-Schoenberg coefficients related to the members ψ_α of the class Ψ_∞ as in Equation (3.14) are given by*

$$b_{n,2}(\alpha) = \frac{2n+1}{2^{1-n}} \left\{ \sum_{m \equiv 0 \pmod{2}}^n \binom{n}{m} \binom{\frac{n+m-1}{2}}{n} \frac{(1 + e^{-\frac{\pi}{\alpha}})}{(m+1)2^m} \right. \\ \left[2^m - \sum_{k=0}^{\frac{m}{2}} \frac{1}{(2k+1)^2 \alpha^2 + 1} \binom{m+1}{\frac{m-2k}{2}} \right] + \\ \sum_{m \equiv 1 \pmod{2}}^n \binom{n}{m} \binom{\frac{n+m-1}{2}}{n} \frac{(1 - e^{-\frac{\pi}{\alpha}})}{(m+1)2^m} \\ \left. \left[2^m - \frac{1}{2} \binom{m+1}{\frac{m+1}{2}} - \sum_{k=1}^{\frac{m+1}{2}} \frac{1}{4k^2 \alpha^2 + 1} \binom{m+1}{\frac{m-2k+1}{2}} \right] \right\}. \quad (3.15)$$

Proof. We provide a proof by direct construction. We first use Equation (3.13) to obtain

$$b_{n,2}(\alpha) = (2n+1)2^{n-1} \sum_{m=0}^n \binom{n}{m} \binom{\frac{n+m-1}{2}}{n} \int_0^\pi e^{-\frac{\theta}{\alpha}} (\cos \theta)^m \sin \theta d\theta. \quad (3.16)$$

Using integration by parts, we have

$$\int_0^\pi e^{(-\frac{\theta}{\alpha})} (\cos \theta)^m \sin \theta d\theta = \left(e^{-\frac{\pi}{\alpha}} (-1)^m + 1 \right) - \frac{1}{\alpha(m+1)} \int_0^\pi e^{-\frac{\theta}{\alpha}} (\cos \theta)^{m+1} d\theta. \quad (3.17)$$

To compute the second term of (3.17), we use the explicit formulae proposed in (Jeffrey and Zwillinger, 2007, Page 228) as follows,

(a) **m is even** The integral on the right hand side of (3.17) is given by

$$\int_0^\pi e^{-\frac{\theta}{\alpha}} (\cos \theta)^{m+1} d\theta = \frac{(1 + e^{-\frac{\pi}{\alpha}})}{2^m} \sum_{k=0}^{\frac{m}{2}} \binom{m+1}{\frac{m-2k}{2}} \frac{\alpha}{(2k+1)^2 \alpha^2 + 1}. \quad (3.18)$$

(b) **m is odd** we obtain

$$\begin{aligned} \int_0^\pi e^{-\frac{\theta}{\alpha}} (\cos \theta)^{m+1} d\theta &= \binom{m+1}{\frac{m+1}{2}} \frac{\alpha (1 - e^{-\frac{\pi}{\alpha}})}{2^{m+1}} \\ &+ \frac{(1 - e^{-\frac{\pi}{\alpha}})}{2^m} \sum_{k=1}^{\frac{m+1}{2}} \binom{m+1}{\frac{m-2k+1}{2}} \frac{\alpha}{(2k\alpha)^2 + 1}. \end{aligned} \quad (3.19)$$

We can now merge (3.18) and (3.19) into (3.17) to obtain

$$\begin{aligned} \int_0^\pi e^{(-\frac{\theta}{\alpha})} (\cos \theta)^m \sin \theta d\theta &= \overbrace{\left(1 + (-1)^m e^{-\frac{\pi}{\alpha}} \right)}^{\text{general } m} - \overbrace{\frac{(1 - e^{-\frac{\pi}{\alpha}})}{(m+1)2^{m+1}} \binom{m+1}{\frac{m+1}{2}}}_{m \text{ is odd}} \\ &- \overbrace{\frac{(1 + e^{-\frac{\pi}{\alpha}})}{(m+1)2^m} \sum_{k=0}^{\frac{m}{2}} \binom{m+1}{\frac{m-2k}{2}} \frac{1}{(2k+1)^2 \alpha^2 + 1}}^{m \text{ is even}} \\ &- \overbrace{\frac{(1 - e^{-\frac{\pi}{\alpha}})}{2^m} \sum_{k=1}^{\frac{m+1}{2}} \binom{m+1}{\frac{m-2k+1}{2}} \frac{\alpha}{(2k\alpha)^2 + 1}}^{m \text{ is odd}} \end{aligned} \quad (3.20)$$

Going back by substitution into Equation (3.20) in (3.16), we obtain (3.15). This completes the proof. \square

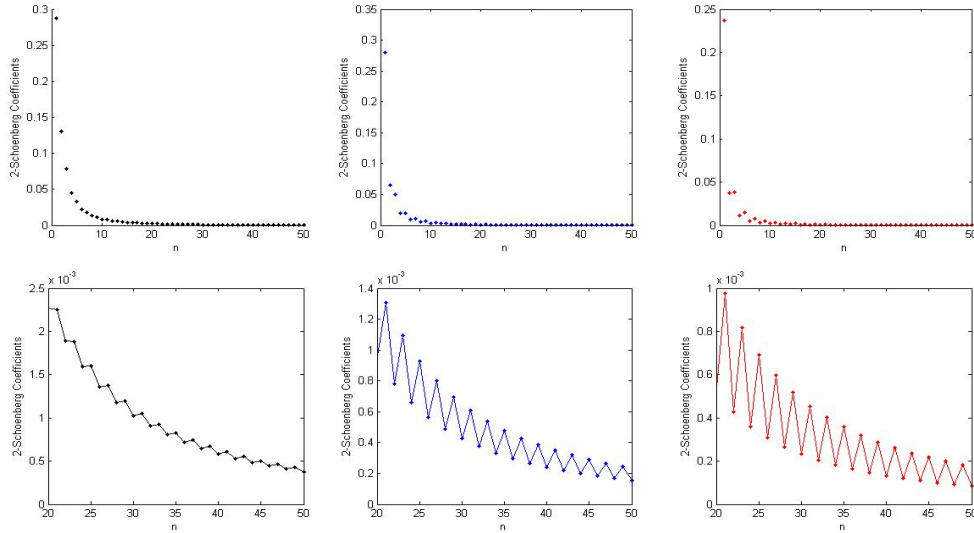


FIGURE 3.2. Top: from left to right, the 2-Schoenberg coefficients associated to the Exponential function ψ_α in Equation (3.14). Bottom: we zoom on the same coefficients from 20 to 50, to show the oscillating behavior away from the origin.

3.3.2 Askey Family

The Askey function (Askey, 1973) $\psi_{\alpha,\tau}$, is defined through

$$\psi_{\alpha,\tau}(\theta) = \left(1 - \frac{\theta}{\alpha}\right)_+^\tau \quad \text{for } \theta \in [0, \pi], \quad (3.21)$$

where $0 < \alpha$ and $\tau \geq (d+1)/2$ are sufficient conditions for $\psi_{\alpha,\tau}$ to belong to the class Ψ_d . Since we are concerned with the 2-Schoenberg's coefficients, we consider the case $\tau = 2$. A relevant remark is that when $0 < \alpha < \pi$, then the Askey function is locally supported.

Proposition 3.3.2. *The 2-Schoenberg coefficients related to the members $\psi_{\alpha,2}$ of the class Ψ_2 as in Equation (3.21) are given by*

$$b_{n,2}(\alpha) = (2n+1) 2^{n-1} \left\{ \sum_{m \equiv 0 \pmod{2}}^n \binom{n}{m} \binom{\frac{n+m-1}{2}}{n} \left[\frac{1}{m+1} + \frac{1}{(m+1)\alpha^{2m-1}} \sum_{k=0}^{\frac{m}{2}} \binom{m+1}{k} \frac{\cos(m-2k+1)\alpha - 1}{(m-2k+1)^2} \right] \right\}$$

$$\begin{aligned}
& + \sum_{m \equiv 1 \pmod{2}}^n \binom{n}{m} \binom{\frac{n+m-1}{2}}{n} \left[\frac{1}{m+1} - \frac{1}{(m+1)2^{m+1}} \binom{m+1}{\frac{m+1}{2}} \right. \\
& \left. + \frac{1}{(m+1)\alpha^2 2^{m-1}} \sum_{k=0}^{\frac{m-1}{2}} \binom{m+1}{k} \left[\frac{\cos(m-2k+1)\alpha - 1}{(m-2k+1)^2} \right] \right] \Bigg\}. \tag{3.22}
\end{aligned}$$

Proof. Again, a proof by direct construction is provided. We first use Equation (3.13) to obtain

$$\begin{aligned}
b_{n,2}(\alpha) &= \frac{(2n+1)}{2^{1-n}} \sum_{m=0}^n \binom{n}{m} \binom{\frac{n+m-1}{2}}{n} \int_0^\pi \left(1 - \frac{\theta}{\alpha}\right)_+^2 (\cos \theta)^m \sin \theta d\theta \\
&= \frac{(2n+1)}{2^{1-n}} \sum_{m=0}^n \binom{n}{m} \binom{\frac{n+m-1}{2}}{n} \int_0^\alpha \left(1 - \frac{2\theta}{\alpha} + \frac{\theta^2}{\alpha^2}\right) (\cos \theta)^m \sin \theta d\theta. \tag{3.23}
\end{aligned}$$

We now note that

$$\begin{aligned}
\int_0^\alpha \left(1 - \frac{2\theta}{\alpha} + \frac{\theta^2}{\alpha^2}\right) (\cos \theta)^m \sin \theta d\theta &= \underbrace{\int_0^\alpha (\cos \theta)^m \sin \theta d\theta}_{\mathcal{I}_1} - \frac{2}{\alpha} \underbrace{\int_0^\alpha \theta (\cos \theta)^m \sin \theta d\theta}_{\mathcal{I}_2} \\
&\quad + \frac{2}{\alpha^2} \underbrace{\int_0^\alpha \theta^2 (\cos \theta)^m \sin \theta d\theta}_{\mathcal{I}_3}. \tag{3.24}
\end{aligned}$$

The integral \mathcal{I}_1 is given by

$$\mathcal{I}_1 = \frac{1}{m+1} - \frac{(\cos \alpha)^{m+1}}{m+1}. \tag{3.25}$$

Using integration by part, we obtain that the integral \mathcal{I}_2 is given by

$$\mathcal{I}_2 = \frac{-\alpha}{m+1} (\cos \alpha)^{m+1} + \frac{1}{m+1} \int_0^\alpha (\cos \theta)^{m+1} d\theta. \tag{3.26}$$

By using the explicit formulas 3 and 4 (Jeffrey and Zwillinger, 2007, Page 153), we have that the integral \mathcal{I}_2 is given by

(a) m is even

$$\mathcal{I}_2 = \frac{-\alpha}{m+1}(\cos \alpha)^{m+1} + \frac{1}{(m+1)2^m} \sum_{k=0}^{\frac{m}{2}} \binom{m+1}{k} \frac{\sin(m-2k+1)\alpha}{m-2k+1} \quad (3.27)$$

(b) m is odd we obtain

$$\begin{aligned} \mathcal{I}_2 &= \frac{-\alpha}{m+1}(\cos \alpha)^{m+1} + \frac{\alpha}{(m+1)2^{m+1}} \binom{m+1}{\frac{m+1}{2}} \\ &+ \frac{1}{(m+1)2^m} \sum_{k=0}^{\frac{m-1}{2}} \binom{m+1}{k} \frac{\sin(m-2k+1)\alpha}{m-2k+1}. \end{aligned} \quad (3.28)$$

To compute the integral \mathcal{I}_3 , we use integration by part to get

$$\mathcal{I}_3 = \frac{-\alpha^2}{m+1}(\cos \alpha)^{m+1} + \frac{2}{m+1} \int_0^\alpha \theta(\cos \theta)^{m+1} d\theta. \quad (3.29)$$

Using the explicit formulas 6 and 7 (Jeffrey and Zwillinger, 2007, Page 215) to compute the integral in the second term of (3.29). After then substitute in (3.29) to obtain the integral \mathcal{I}_3 as follows

(a) m is even

$$\begin{aligned} \mathcal{I}_3 &= \frac{-\alpha^2}{m+1}(\cos \alpha)^{m+1} + \frac{1}{(m+1)2^{m-1}} \sum_{k=0}^{\frac{m}{2}} \binom{m+1}{k} \binom{m+1}{k} \\ &\times \left[\frac{\alpha \sin(m-2k+1)\alpha}{m-2k+1} + \frac{\cos(m-2k+1)\alpha - 1}{(m-2+k)^2} \right]. \end{aligned} \quad (3.30)$$

(b) m is odd then

$$\begin{aligned} \mathcal{I}_3 &= \frac{-\alpha^2}{m+1}(\cos \alpha)^{m+1} + \frac{\alpha^2}{(m+1)2^{m+1}} \binom{m+1}{\frac{m+1}{2}} \\ &+ \frac{1}{(m+1)2^{m-1}} \sum_{k=0}^{\frac{m-1}{2}} \binom{m+1}{k} \sum_{k=0}^{\frac{m-1}{2}} \binom{m+1}{k} \\ &\times \left[\frac{\alpha \sin(m-2k+1)\alpha}{m-2k+1} + \frac{\cos(m-2k+1)\alpha - 1}{(m-2+k)^2} \right] \end{aligned} \quad (3.31)$$

Then, from (3.25), (3.27), (3.28), (3.30) and (3.31) in (3.24), we have

(a) when m is even

$$\mathcal{I}_1 - \frac{2\mathcal{I}_2}{\alpha} + \frac{\mathcal{I}_3}{\alpha^2} = \frac{1}{m+1} + \frac{2^{1-m}}{(m+1)\alpha^2} \sum_{k=0}^{\frac{m}{2}} \binom{m+1}{k} \frac{\cos(m-2k+1)\alpha - 1}{(m-2k+1)^2} \quad (3.32)$$

(b) when m is odd we obtain

$$\begin{aligned} \mathcal{I}_1 - \frac{2\mathcal{I}_2}{\alpha} + \frac{\mathcal{I}_3}{\alpha^2} &= \frac{1}{m+1} + \frac{1}{(m+1)2^{m+1}} \binom{m+1}{\frac{m+1}{2}} \\ &+ \frac{2^{1-m}}{(m+1)\alpha^2} \sum_{k=0}^{\frac{m-1}{2}} \binom{m+1}{k} \left[\frac{\cos(m-2k+1)\alpha - 1}{(m-2k+1)^2} \right]. \end{aligned} \quad (3.33)$$

Going back by substitution into Equation (3.32) and (3.33) in (3.23), we obtain (3.22). This completes the proof. \square

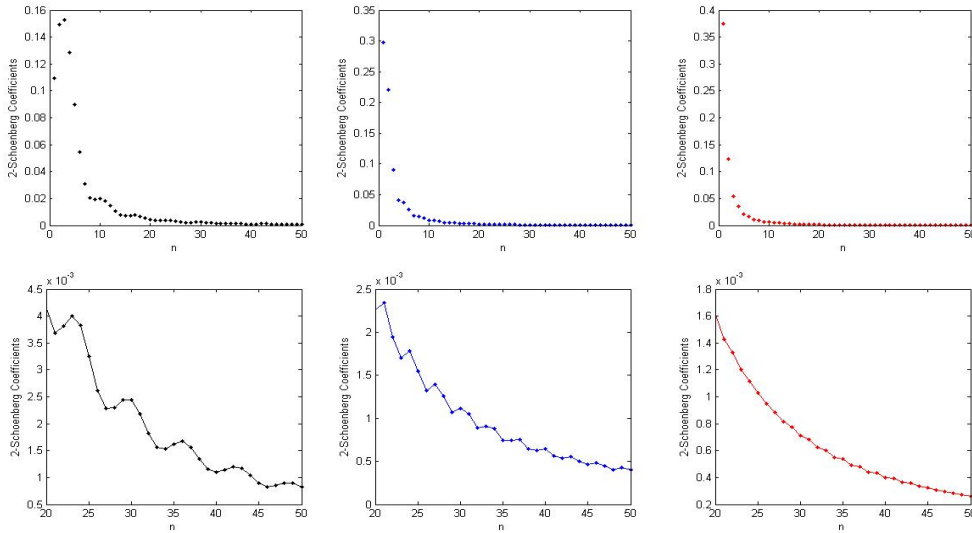


FIGURE 3.3. Top: from the left to the right, the 2-Schoenberg coefficients of the Askey function $\psi_{\alpha,2}$ defined in Equation (3.21). Bottom: we zoom on the same coefficients from 20 to 50, to show the oscillating behavior away from the origin.

CHAPTER 4

A family of Markov Processes in Maximal Compact Subgroups of a Semisimple Lie Groups

4.1 Introduction

Let \mathbb{E} be a locally compact Hausdorff space whose topology has a countable basis (usually \mathbb{R}^d), \mathcal{E} the Borel σ -algebra on \mathbb{R}^d , \mathcal{B} the Borel σ -algebra on \mathbb{R} , and $\mathbf{M} := \{\mu : 0 \leq \mu(A) < \infty \text{ for } A \in \mathcal{B}\}$ the set of nonnegative finite measures. Let $\mathbf{M}_{\sharp} := \{\mu : \mu(A) \in \mathbb{N} \text{ for } A \in \mathcal{B}\}$ be the set of counting measures, and \mathcal{M} the σ -algebra on \mathbf{M} generated by the coordinate mappings $\mu \mapsto \mu(f) = \int f d\mu$ for $f \in \mathcal{C}_k$, where \mathcal{C}_k is the set of continuous functions on \mathbb{E} with compact support. Finally, if $(\Omega, \mathcal{F}, \mathbb{P})$ denotes a probability space, a point process on \mathbb{E} is a measurable function $N : (\Omega, \mathcal{F}) \rightarrow (\mathbf{M}_{\sharp}, \mathcal{M})$.

A one-dimensional point process is intended to describe events that occur randomly over the positive real line. It can be represented as a sequence of nonnegative random variables $0 = T_0 < T_1 < T_2 < \dots$, where the generic T_n is the n -th

instant of occurrence of an event. Under the usual assumption of non-explosion, according to which $T_\infty = \lim \uparrow T_n = \infty$, the process can be represented via its associated counting process N_t where

$$N_t = n \quad \text{if } t \in [T_n, T_{n+1}), \quad n \geq 0 \quad \text{or, equivalently, } N_t = \sum_{n \geq 1} \mathbf{1}_{\{T_n \leq t\}} \quad (4.1)$$

counts the number of events up to, and including, time t (see [Daley and Vere-Jones \(2002\)](#), [Daley and Vere-Jones \(2007\)](#)). The non-explosion condition becomes $N_t < \infty$ for all $t \geq 0$. Both T_n and N_t are defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration \mathcal{F}_t to which N_t is adapted (i.e. a measurable function for all $t \geq 0$). A point process N_t is called a Poisson point process if $N_0 = 0$, N_t is a process with independent increments, and $N_t - N_s$ is a Poisson random variable of a given parameter $\Lambda_{s,t}$.

Usually one assumes $\Lambda_{s,t} = \int_s^t \lambda_u du$ for a deterministic function λ_t dubbed *intensity* of the Poisson point process N_t . If \mathcal{F}_t is the filtration \mathcal{F}_t^N generated by N_t , and $\lambda_t \equiv 1$, then N_t is called a standard Poisson process. It is also easily seen that if N_t is a Poisson process with intensity $\lambda_t = \lambda$, then $T_{n+1} - T_n$ are i.i.d. exponential random variables with parameter λ . Wiener processes are similar: the increments of a Wiener process are normally distributed, and independent if the intervals are disjoint. It is the basic building block for processes with continuous trajectories, while the Poisson process remains for processes with jumping trajectories. On the contrary, the Wiener process is itself a martingale, while the Poisson process becomes a martingale only if one subtracts from N_t the process given by its mean. Indeed,

$$M_t := N_t - \int_0^t \lambda_s ds \quad (4.2)$$

is a \mathcal{F}_t -martingale by the \mathcal{F}_0 -measurability of λ_t , assuming in addition that $\mathbf{E} \left\{ \int_s^t \lambda_u du \right\} < \infty$. As a consequence, another characterization of a Poisson process is

$$\mathbf{E} \{ N_t - N_s | \mathcal{F}_s \} = \mathbf{E} \left\{ \int_s^t \lambda_u du \middle| \mathcal{F}_s \right\} \quad (4.3)$$

which implies that $\mathbf{E}\{N_t\} < \infty$ and that M_t in (4.2) is an \mathcal{F}_t -martingale. A generalization of (4.3) admits the form

$$\mathbf{E}\left\{\int_0^\infty C_s dN_s\right\} = \mathbf{E}\left\{\int_0^\infty C_s \lambda_s ds\right\} \quad (4.4)$$

that has to be valid for all nonnegative and \mathcal{F}_t -predictable processes C_t , characterizing a Poisson process with intensity λ_t (see [Brémaud \(1981\)](#)).

Our aim is to define a family of marked point processes in noncompact semisimple Lie groups. This will help building a family of Markov processes in a maximal compact subgroup of semisimple Lie groups. This class has lots of potential applications, and it is a building block for other types of point processes.

The plan of this chapter is the following. Section 2 shows how to build marked point processes and their multivariate version for noncompact semisimple Lie groups. Then Section 3 develops Lévy processes in semisimple Lie groups that help to define Markov processes in a maximal compact subgroup of semisimple Lie groups.

4.2 Marked Point Processes and the Multivariate Case

Let $(T_n)_{n \geq 1}$ be a (univariate) point process and $(Y_n)_{n \geq 1}$ a sequence of random variables with values in $\{1, 2, \dots, K\}$, all defined on the same space $(\Omega, \mathcal{F}, \mathbb{P})$. For each $k = 1, \dots, K$ we may then consider the counting process

$$N_t(k) := \sum_{n \geq 1} \mathbf{1}_{\{T_n \leq t\}} \mathbf{1}_{\{Y_n = k\}}.$$

Each $N_t(k)$ is a univariate point process and the various $N_t(k)$'s have no common jumps. As in the univariate case, we have two equivalent representations, either as the double sequence $(T_n, Y_n)_{n \geq 1}$, or as the K -vector process $N_t = (N_t(1), \dots, N_t(K))$, being a multivariate point process, or more precisely a

K -variate point process. Considering N_t as a K -vector process, using (4.2) we can define M_t a K -vector martingale with λ_t a K -vector intensity process whose components are the individual intensities of the components $N_t(k)$ of N_t .

Considering the representation (T_n, Y_n) , we may interpret T_n as the n -th occurrence of some phenomenon, and Y_n as an attribute or mark of this phenomenon. We may then speak of (T_n, Y_n) as a marked point process, or a space-time point process, and extend its definition to allow Y_n to take values in a general measurable mark space $(\mathbb{E}, \mathcal{E})$. We synthesize this idea in the following definition.

Definition 4.2.1. *An \mathbb{E} -marked point process is a double sequence $(T_n, Y_n)_{n \geq 1}$ where T_n is a (univariate) point process, and Y_n is a sequence of \mathbb{E} -valued random variables.*

Obviously, the univariate and multivariate point processes are special cases of a marked point process (see Björk et al. (1997)). Generalizing the representation of a multivariate point process in the form of the K -vector process N_t , we associate to each $A \in \mathcal{E}$ the counting process

$$N_t(A) := \sum_{n \geq 1} \mathbb{1}_{\{T_n \leq t\}} \mathbb{1}_{\{Y_n \in A\}}$$

and we write simply $N_t = N_t(\mathbb{E})$. Considering the filtration

$$\mathcal{F}_t^N := \sigma \{N_s(A); s \leq t, A \in \mathcal{E}\}$$

we define the associated (random) counting measure

$$p((0, t], A) = N_t(A), \quad t \geq 0, \quad A \in \mathcal{E} \quad (4.5)$$

which is σ -finite under the assumption of non-explosion of T_n . This measure allows to obtain more concise expressions via integrals of the form

$$\int_0^t \int_{\mathbb{E}} H(s, y) p(ds, dy) = \sum_{n \geq 1} H(T_n, Y_n) \mathbb{1}_{\{T_n \leq t\}} = \sum_{n=1}^{N_t} H(T_n, Y_n). \quad (4.6)$$

Again, we may represent an \mathbb{E} -marked point process equivalently as the double sequence (T_n, Y_n) or as the counting measure $p(ds, dy)$.

In order to introduce now the intensity process in this more general setup, let us assume that for each $A \in \mathcal{E}$, the point process $N_t(A)$ admits the intensity $\lambda_t(A)$. This leads to a measure-valued intensity $\lambda_t(dy)$ so that, generalizing (4.4), one has

$$\mathbf{E} \left\{ \int_0^\infty \int_{\mathbb{E}} H(s, y) p(ds, dy) \right\} = \mathbf{E} \left\{ \int_0^\infty \int_{\mathbb{E}} H(s, y) \lambda_s(dy) ds \right\} \quad (4.7)$$

that has to be valid for all nonnegative \mathcal{F}_t -predictable \mathbb{E} -marked processes H (given a filtration \mathcal{F}_t on Ω). \mathcal{F}_t -predictability here means measurability with respect to $\mathcal{P}(\mathcal{F}_t) \otimes \mathcal{E}$, where $\mathcal{P}(\mathcal{F}_t)$ is the predictable σ -field on $(0, \infty) \times \Omega$. We also have the generalization of (4.2) in the form

$$q(ds, dy) = p(ds, dy) - \lambda_s(dy) ds \quad (4.8)$$

where $q(ds, dy)$ is a (signed) measure-valued martingale in the sense that the integral $\int_0^t \int_{\mathbb{E}} H(s, y) q(ds, dy)$ is a $(\mathbb{P}, \mathcal{F}_t)$ -martingale for each \mathcal{F}_t -predictable \mathbb{E} -marked process H , satisfying appropriate integrability conditions. The most common form of intensity is given by

$$\lambda_t(dy) = \lambda_t m_t(dy) \quad (4.9)$$

where λ_t is nonnegative \mathcal{F}_t -predictable, and represents the intensity of the Poisson process $N_t(\mathbb{E})$, while $m_t(dy)$ is a probability measure on \mathbb{E} (typically, the Y_n will be i.i.d., independent of $N_t(\mathbb{E})$). The pair $(\lambda_t, m_t(dy))$ is called the $(\mathbb{P}, \mathcal{F}_t)$ -local characteristics of $p(ds, dy)$.

Notice finally that, as in the univariate case, we may let $\lambda_t(dy)$ depend on some driving \mathcal{F}_0 -measurable random process Z_t , leading to a doubly stochastic marked point process. If, in the representation (4.9), λ_t is a deterministic time function, the marked point process is called a marked Poisson process.

We shall recall a corresponding result for point-process martingales that we formulate in the most general case of a marked point process. Following Theorem 8 in (Brémaud, 1981, chapter VIII), the \mathcal{F}_t -martingale M_t admits the following representation

$$M_t = M_0 + \int_0^t \int_{\mathbb{E}} H(s, y) q(ds, dy). \quad (4.10)$$

So, in the case of a multivariate (and univariate) point process, the representation (4.10) becomes

$$M_t = M_0 + \sum_{k=1}^K \int_0^t H_s(k) (dN_s(k) - \lambda_s(k) ds) \quad (4.11)$$

where $(H_t(1), \dots, H_t(K))$ is \mathcal{F}_t -predictable with $H_t(k)$ integrable with respect to $\lambda_t(k)$ for all k .

With the definition of a marked point process and with the integrals in the form of (4.6), we may now consider processes of the general form

$$X_t = X_0 + \int_0^t \alpha_s ds + \int_0^t \beta_s d\omega_s + \int_0^t \int_{\mathbb{E}} \gamma(s, y) p(ds, dy) \quad (4.12)$$

that are called jump-diffusion processes, and where the coefficients satisfy the implicit integrability conditions, β_t is adapted and $\gamma(t, y)$ is predictable as defined previously (see Ahn and Thompson (1988)). As usual, we may rewrite (4.12) in differential form and consider, more specifically, differential equations of the type

$$dX_t = X_{t-} \left(\alpha_t dt + \beta_t d\omega_t + \int_{\mathbb{E}} \gamma(t, y) p(dt, dy) \right) \quad (4.13)$$

where we write X_{t-} with $t-$ because of the predictability requirement in the last coefficient and where $\gamma(t, y) > -1$. The last term in (4.13) can also be written as

$$\int_{\mathbb{E}} \gamma(t, y) p(dt, dy) = \gamma(t, Y_t) dN_t \quad (4.14)$$

(see (4.6)), where $N_t = N_t(\mathbb{E}) = p((0, t], \mathbb{E})$ is the total number of jumps and Y_t denotes the piecewise constant, left-continuous time interpolation of the sequence

Y_n . Notice also that, in the case of a multivariate (in particular univariate) point process, this last term in (4.13) takes the form

$$\int_{\mathbb{E}} \gamma(t, y) p(dt, dy) = \sum_{k=1}^K \gamma_t(k) dN_t(k). \quad (4.15)$$

We shall not discuss in detail equations of the form (4.13), in particular the uniqueness of their solutions, but we limit ourselves to show that a solution to (4.13) is given by the following exponential formula

$$X_t = X_0 \exp \left[\int_0^t \left(\alpha_s - \frac{1}{2} \beta_s^2 \right) ds + \int_0^t \beta_s d\omega_s \right] \prod_{n=1}^{N_t} (1 + \gamma(T_n, Y_n)). \quad (4.16)$$

Note that the diffusion part in this expression follows from the usual Itô's formulae, the jump part follows from the so-called exponential formula of Stieltjes-Lebesgue Calculus (see Theorem T4 of Appendix A4 in [Brémaud (1981)]), but it can also be obtained from the generalized Itô's formulae as we show now. For this purpose let a process X_t satisfy the general equation (4.12). Given a $C^{1,2}$ -function $F(t, X)$, we have that the generalized Itô's formula, in the specific case of (4.13), becomes

$$\begin{aligned} dF(t, X_t) &= F_t(\cdot) dt + F_X(\cdot) X_t \alpha_t dt + \frac{1}{2} F_{XX}(\cdot) X_t^2 \beta_t^2 dt + F_X(\cdot) X_t \beta_t d\omega_t \\ &+ [F(t, X_{t-}(1 + \gamma(t, Y_t))) - F(t, X_{t-})] dN_t \end{aligned} \quad (4.17)$$

and where again, $N_t = N_t(\mathbb{E}) = p((0, t], \mathbb{E})$, (\cdot) stands for (t, X_t) , and the subscript in F denotes partial derivatives. Notice that, if (4.17) is written in integral form, we have the following two equivalent representations for the last term on the right hand side

$$\int_0^t [F(s, X_{s-}(1 + \gamma(s, Y_s))) - F(s, X_{s-})] dN_s = \sum_{n=1}^{N_t} [F(T_n, X_{T_n}) - F(T_n, X_{T_n^-})]$$

where the right hand side remains the same also in the more general case of Itô's formulae.

We shall now use the generalized Itô's formulae (4.17) to obtain the solution (4.16) of equation (4.13). Choosing $F(t, X) = \log X$, from (4.17) and (4.13) we have

$$dF = \alpha_t dt - \frac{1}{2} \beta_t^2 dt + \beta_t d\omega_t + \log(1 + \gamma(t, Y_t)) dN_t$$

from which

$$\log X_t = \log X_0 + \int_0^t \left(\alpha_s - \frac{1}{2} \beta_s^2 \right) ds + \int_0^t \beta_s d\omega_s + \int_0^t \log(1 + \gamma(s, Y_s)) dN_s \quad (4.18)$$

i.e. we obtain (4.16) by taking the exponential on both sides in (4.18).

Our aim in this section is to obtain a closed form for a marked point processes (later called L_t) is the more general multivariate case. For this goal, we first need to rephrase Girsanov's measure transformation theorem coming in Theorem 4.2.2, and we also need to recall (Brémaud, 1981, chp. VIII, T10) in this case coming in Theorem 4.2.3.

Theorem 4.2.2 (Girsanov's measure transformation). *Let $T > 0$ be fixed, $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ a filtered probability space with $\mathcal{F} = \cup_t \mathcal{F}_t$, and θ_t a square integrable predictable process. Define $(L_t)_{t \in [0, T]}$ by*

$$dL_t = L_t \theta_t d\omega_t, \quad L_0 = 1 \quad (4.19)$$

and suppose that, for all t , $\mathbf{E}^{\mathbb{P}} \{L_t\} = 1$. Then there exists a probability measure Q on \mathcal{F} , equivalent to \mathbb{P} , with $dQ = L_T d\mathbb{P}$ such that

$$d\omega_t = \theta_t dt + d\omega_t^Q, \quad L_0 = 1 \quad (4.20)$$

where ω_t^Q is a Q -Wiener process.

Girsanov's measure transformation allows to change the drift in a diffusion equation. In fact, suppose that, under \mathbb{P} , we have $dX_t = a_t X_t dt + \sigma_t X_t d\omega_t$, and that we would like to change to a measure $Q \sim \mathbb{P}$ (\sim meaning equivalent to), under which the same X_t satisfies $dX_t = r_t X_t dt + \sigma_t X_t d\omega_t^Q$. In this case just take $\theta_t = \sigma_t^{-1}(r_t - a_t)$.

If we have both the Wiener process ω_t and a marked point process represented by a counting measure $p(dt, dy)$, a Girsanov-type measure transformation allows, in addition to the translation of the Wiener process, to perform also a change in the intensity of the point process part. Then we can rephrase (Brémaud, 1981, VIII, T10) in the following terms.

Theorem 4.2.3. *Let $[0, T]$ be a finite time interval, $p(dt, dy)$ an \mathbb{E} -marked point process with $(\mathbb{P}, \mathcal{F}_t)$ -local characteristic $(\lambda_t, m_t(dy))$. Let $\psi_t \geq 0$ be \mathcal{F}_t -predictable and $h_t(y) \geq 0$ an \mathcal{F}_t -predictable \mathbb{E} -indexed process such that, \mathbb{P} -a.s. and for all $t \in [0, T]$,*

$$\int_0^t \psi_s \lambda_s ds < \infty; \quad \int_{\mathbb{E}} h_t(y) m_t(dy) = 1.$$

Define $L_t = L_t^{(1)} \cdot L_t^{(2)}$ where $L_t^{(1)}$ satisfies (4.19) and $L_t^{(2)}$ satisfies

$$dL_t^{(2)} = \int_{\mathbb{E}} (\psi_t h_t(y) - 1) L_{t-}^{(2)} q(dt, dy) \quad (4.21)$$

with $q(dt, dy) = p(dt, dy) - \lambda_t m_t(dy) dt$ the martingale measure associated to $p(dt, dy)$. If $\mathbf{E}^{\mathbb{P}} L_t^{(2)} = 1$ for all t , then all the statements of Girsanov's transformation hold true. In addition, $p(dt, dy)$ has the (Q, \mathcal{F}_t) -local characteristics $(\psi_t \lambda_t, h_t(y) m_t(dy))$.

Notice that, using Girsanov's transformation and (4.21), the Radon-Nikodým derivative of L_t is

$$\begin{aligned} dL_t &= D \left(L_t^{(1)} \cdot L_t^{(2)} \right) = L_{t-}^{(1)} dL_t^{(2)} + L_t^{(2)} dL_t^{(1)} \\ &= L_t \theta_t \omega_t + L_{t-} \int_{\mathbb{E}} (\psi_t h_t(y) - 1) q(dt, dy), \quad L_0 = 1 \end{aligned} \quad (4.22)$$

Using the exponential formula (4.16), we have that a solution of (4.22) is given by

$$L_t = \exp \left\{ -\frac{1}{2} \int_0^t \theta_s^2 ds + \int_0^t \theta_s d\omega_s + \int_0^t \int_{\mathbb{E}} (1 - \psi_s h_s(y)) \lambda_s m_s(dy) ds \right\} \prod_{n=1}^{N_t} (\psi_{T_n} h_{T_n}(Y_n)) \quad (4.23)$$

In the case of a multivariate (in particular univariate) point process $(N_t(1), \dots, N_t(K))$ with $(\mathbb{P}, \mathcal{F}_t)$ -intensities $(\lambda_t(1), \dots, \lambda_t(K))$, consider an \mathcal{F}_t -predictable process $(\psi_t(1), \dots, \psi_t(K))$ such that, \mathbb{P} -a.s., and for $t \in [0, T]$, $\sum_{k=1}^K \int_0^t \psi_s(k) \lambda_s(k) ds < \infty$. Define $L_t^{(2)}$ by

$$dL_t^{(2)} = \sum_{k=1}^K (\psi_t(k) - 1) L_{t-}^{(2)} (dN_t(k) - \lambda_t(k) dt). \quad (4.24)$$

Extending (4.21) - (4.22) - (4.23) to this multivariate case, we have

$$L_t = \exp \left\{ -\frac{1}{2} \int_0^t \theta_s^2 ds + \int_0^t \theta_s d\omega_s \right\} \prod_{k=1}^K \left[\exp \left\{ \int_0^t (1 - \psi_s(k)) \lambda_s(k) ds \right\} \prod_{n=1}^{N_t(k)} \psi_{T_n}(k) \right] \quad (4.25)$$

Then, under Q , the intensities become $(\psi_t(1)\lambda_t(1), \dots, \psi_t(K)\lambda_t(K))$, and a condition to have $\mathbf{E}^{\mathbb{P}} L_t^{(2)} = 1$ can be found in (Brémaud, 1981, VIII, T11). We note that the values of L_t are stochastic, since, for our case, $N_t \sim \text{Poisson}(\lambda t)$. This is an important result because it has many applications, such as in asset price models.

4.3 Lévy Processes in Semisimple Lie Groups

Consider G a noncompact semisimple Lie group, and K the maximal compact Lie subgroup of G . If we assume that $G = GL_n(\mathbb{R})$ with e the identity matrix in $GL_n(\mathbb{R})$, then we can assume $K = O(n)$. We denote the dimension of G by d . Its Lie algebra \mathfrak{g} will be identified with the left-invariant vector field at the identity e . For any $X \in \mathfrak{g}$ we denote by X^ℓ the corresponding left-invariant vector field. In the case of matrix Lie groups, this vector field is derivative in the direction of X^ℓ , that is, for a function $f : G \rightarrow \mathbb{R}$ we define

$$\frac{\partial f}{\partial X^\ell}(l) = \left. \frac{d}{dt} \right|_{t=0} f(\exp^{tX} l), \quad l \in G,$$

if such a derivative exists.

Throughout this section we will work with a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, where \mathcal{F}_t is a σ -field contained in \mathcal{F} . We assume that $\mathcal{F} = \lim_{t \uparrow \infty} \mathcal{F}_t$ and, whenever necessary, that all \mathbb{P} -null sets belong to \mathcal{F}_t for any $t \in [0, \infty)$, and that the filtration \mathcal{F}_t is right-continuous, i.e. $\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s$.

Now our aim is building a family of Markov point processes (i.e., left invariance Lévy processes) via the stochastic processes given by (4.25). Next we recall the definition of the Lévy process to the identification of how to build that process.

Definition 4.3.1. *Let L_t be a stochastic process with values in G and let $t \in [0, \infty)$, then:*

- (1) $L_s^{-1}L_t$ (resp. $L_tL_s^{-1}$) is called the right (resp. left) increment of the process L_t for $s < t$.
- (2) L_t is said to have independent right (resp. left) increments if for any $0 < t_1 < t_2 < \dots < t_n$, the processes $L_0, L_0^{-1}L_{t_1}, L_{t_1}^{-1}L_{t_2}, \dots, L_{t_{n-1}}^{-1}L_{t_n}$ (resp. $L_0, L_{t_1}L_0^{-1}, L_{t_2}L_{t_1}^{-1}, \dots, L_{t_n}L_{t_{n-1}}^{-1}$) are independent.
- (3) L_t , with independent right (resp. left) increments, is said to have stationary right (resp. left) increments if $L_s^{-1}L_t \stackrel{d}{=} L_0^{-1}L_{t-s}$ (resp. $L_tL_s^{-1} \stackrel{d}{=} L_{t-s}L_0^{-1}$) for any $s < t$.
- (4) L_t is called Càdlàg if almost all of its paths are right continuous on $[0, \infty)$ and have left limits on $(0, \infty)$.
- (5) L_t is called left (resp. right) Lévy process in G if it is a Càdlàg process with independent and stationary right (resp. left) increments.

Since the form (4.25) satisfies the five conditions in Definition (4.3.1) by considering $L_0 = e$, it is a Lévy process generated by a marked point process, i.e. this structure depends on a point process. In general, the transition probability on a semigroup of this process is given by

$$\mathbb{P}_t(f(l)) = \mathbf{E}[f(lL_t)]$$

for any non-negative Borel function f on G . Then the distribution μ_t of L_t is a weakly continuous convolution semigroup of a probability measure on G , and satisfies

$$\mathbb{P}_t(f(l)) = \int_G f(lh) d\mu_t(h), \quad t \in [0, \infty).$$

Let $\mathcal{C}_0(G)$ be the Banach space (with respect to the supremum norm) of functions on G that vanish at infinity. Just as in the Euclidean case, one obtains a Feller semigroup of operators $(T(t), t \geq 0)$ on $\mathcal{C}_0(G)$ by the prescription

$$(T(t)f)(l) = \mathbf{E}(f(lL_t)),$$

for each $t \geq 0, l \in G, f \in \mathcal{C}_0(G)$, and its infinitesimal generator can be defined by

$$\mathcal{L}f(l) := \lim_{t \rightarrow \infty} \left(\frac{1}{t} \right) \{ \mathbf{E}[f(lL_t)] - f(l) \}$$

for any continuous function f on G vanishing at infinity, such that the limit exists under the supremum norm.

We fix a basis $\{X_1, \dots, X_n\}$ for \mathfrak{g} and define a dense subspace $\mathcal{C}_2(G)$ of $\mathcal{C}_0(G)$ as follows

$$\mathcal{C}_2(G) := \left\{ f \in \mathcal{C}_0(G) : X_i^\ell(f) \in \mathcal{C}_0(G) \text{ and } X_i^\ell X_j^\ell(f) \in \mathcal{C}_0(G) \text{ for all } 1 \leq i, j \leq n \right\},$$

where X^ℓ denotes the left invariant vector field associated to $X \in \mathfrak{g}$.

By recalling and applying Theorem 5.1 in [Hunt \(1956\)](#) for the generator \mathcal{L} associated to the Lévy process L_t , we have

Theorem 4.3.2. *Since L_t is a Lévy process in G with infinitesimal generator \mathcal{L} then,*

$$(i) \mathcal{C}_2(G) \subseteq \text{Dom}(\mathcal{L}).$$

(ii) For each $l \in G, f \in \mathcal{C}_2(G)$

$$\begin{aligned} \mathcal{L}f(l) &= \sum_{i=1}^n b_i X_i^\ell f(l) + \frac{1}{2} \sum_{i,j=1}^n c_{ij} X_i^\ell X_j^\ell f(l) \\ &+ \int_{G-\{e\}} \left(f(L_t \sigma) - f(L_t) - \sum_{i=1}^n y_i(\sigma) X_i^\ell f(L_t) \right) \nu(d\sigma), \end{aligned} \quad (4.26)$$

where, $b = (b_1, \dots, b_n) \in \mathbb{R}^n, c = (c_{ij})$ is non-negative definite, symmetric $n \times n$ real-valued matrix and ν is a Lévy measure of L_t on $G - \{e\}$.

Let $(L_t)_{t \in [0, T]}$ be a continuous Lévy process in G whose generator is given by (4.26) with $\nu = 0$, i.e.,

$$\mathcal{L}(L_t) = \sum_{i=1}^n b_i X_i^\ell f(L_t) + \frac{1}{2} \sum_{i,j=1}^n c_{ij} X_i^\ell X_j^\ell f(L_t). \quad (4.27)$$

Following [Hunt \(1956\)](#) and [Ramaswami \(1974\)](#), a Lévy process L_t in G can be characterized as a Markov process with generator given by (4.27) for any smooth function on G with compact support. The main result of this section is stated as the following proposition.

Proposition 4.3.3. *If L_t is a Lévy process in a noncompact semisimple Lie group G , then there exists a Markov process in the maximal compact Lie subgroup of G .*

Proof. Assume that L_t is a Lévy process in G , and K is its maximal compact Lie subgroup. Recall that for $L_t \in G$, we can write $L_t = n_t a_t k_t$ by the Iwasawa decomposition on $G = NAK$. the map $(n, a, g) \mapsto l = nag$ is a diffeomorphism from $N \times A \times K$ onto G , where A is connected Abelian subgroup of G , N is connected Lie subgroup of G and K is the right G -space defined by $K \equiv (NA) \backslash G$. Since K is the right G -space, then $k_t = (NA)L_t$ is a left invariance Lévy process, moreover is a Markov process. This complete the proof. \square

By applying Proposition (4.3.3) to (4.25), we obtain a closed form expression for generating general families of Markov processes k_t . By providing particular forms to ϕ_s and λ_s we can derive examples cases that can be applied to a wide variety of particular situations.

4.4 Conclusions

We have defined a new family of Markov process in the maximal compact Lie group of a noncompact semisimple Lie group. In particular, we have built a marked point process given by (4.25) through the associated infinitesimal generator of L_t . Our result can be further used to drive particular cases by providing ϕ_s and λ_s .

CHAPTER 5

Conclusions and Future Work

We here highlight the most relevant conclusions that the current work has brought to light.

5.1 General Conclusions

- We have studied the equivalence and orthogonality of any two Gaussian measures on the sphere by given the necessary and sufficient condition for the equivalence of Gaussian measures on \mathbb{S}^d .
- We have proposed the equivalence and orthogonality of the Gaussian measures associated to the multiquadric and sine power covariance functions, in addition to their simulation study.
- For any dimension $d \geq 2$, we show the relation between the d -Schoenberg coefficients associated to a given positive definite functions on sphere and the Fourier coefficients on the circle. This point is a generalization of the first problem proposed in [Gneiting \(2013b\)](#).

- For an integer $d \geq 1$, and for a given $c \in (\arccos(\sqrt{\frac{1}{d+1}}), \pi]$, we solved the curvature problem for the class Ψ_d at zero, that is,

$$a_d^c := \inf_{\psi \in \Psi_d^c} (-\psi''(0)).$$

This point is the solution of an open problem proposed in [Gneiting \(2013b\)](#).

- We have given explicit formulas for the 2-Schoenberg coefficients for the exponential and Askey's covariance family on \mathbb{S}^2 .
- We have proposed a new family of Markov processes in the maximal compact Lie group of a noncompact semisimple Lie group defined by a marked point process.

5.2 Ongoing and Future Work

We now write some open ideas that will be discussed and developed in the future.

5.2.1 Equivalence of Gaussian Measures of Multivariate Random Fields on Sphere

In Chapter 2, we proposed the necessary and sufficient condition for any two Gaussian measures of scalar-valued random field to be equivalence on the spheres.

Let $\mathbf{Z} = (Z_1, \dots, Z_d)^T$ be a zero-mean Gaussian vector-valued random field on \mathbb{S}^d , with the matrix covariance operator $\mathbf{K}_{\mathbf{Z}}$ of \mathbf{Z} , given by

$$\mathbf{K}_{\mathbf{Z}} = \begin{bmatrix} K(Z_1, Z_2) & \cdots & K(Z_1, Z_d) \\ \vdots & \ddots & \vdots \\ K(Z_d, Z_1) & \cdots & K(Z_d, Z_d) \end{bmatrix}, \quad (5.1)$$

where $K(Z_i, Z_j) = \mathbb{E}(Z_i Z_j)$, $i, j = 1 \dots, d$ and defined by (2.6).

Problem 5.2.1. *Let \mathbf{Z}_1 and \mathbf{Z}_2 be two zero-mean Gaussian vector-random fields on \mathbb{S}^d with corresponding matrix covariance operator $\mathbf{K}_{\mathbf{Z}_1}$ and $\mathbf{K}_{\mathbf{Z}_2}$, respectively. Under which condition the measures μ_1 and μ_2 associated to \mathbf{Z}_1 and \mathbf{Z}_2 , respectively, are equivalent?*

5.2.2 When a Function from the Class Ψ_d Belongs to Ψ_{d+1} ?

In Chapter 3, we defined the class Ψ_d and its members are continuous function has a spectral representation (2.6). Following Trübner and Ziegel (2016) the classes Ψ_d are nondecreasing in d and satisfy the inclusion

$$\bigcap_{d=1}^{\infty} \Psi_d = \Psi_{\infty} \subseteq \dots \Psi_2 \subseteq \Psi_1 \quad (5.2)$$

Corollary 4 in Gneiting (2013a) gives the necessary and sufficient condition for a function in the class Ψ_d to belong to the class Ψ_{d+2} .

Problem 5.2.2. *If $d \geq 2$, under which condition the function in the class Ψ_d belongs to the class Ψ_{d+1} ?*

The difference between the result in (Gneiting, 2013a, corollary 4) and that problem is there the walk dimension with step 2, but here the step is 1. Also, Golinskii et al. (2015) studied that problem on the Euclidean Space \mathbb{R}^d . Problem 5.2.2 is important to solve the second problem in Gneiting (2013b).

5.2.3 Turán Constant of \mathbb{S}^d

Let the class Ψ_d defined as in Chapter 3. The Turán constant of \mathbb{S}^d is defined by $\mathcal{T} = \sup \{ \int_{\mathbb{S}^d} \psi : \psi \in \Psi_d \}$. See Arestov and Berdysheva (2001); Gorbachev and Manoshina (2004) for the history of the problem.

Problem 5.2.3. *For a positive definite function ψ on the sphere \mathbb{S}^d with $\psi(0) = 1$, how large can \mathcal{T} be?*

Révész (2011) discussed the results of the problem on the Euclidean case and for balls in Euclidean space.

Bibliography

- Adler, R. J. and Taylor, J. E. (2007), *Random Fields and Geometry*, Springer.
- Ahn, C. M. and Thompson, H. E. (1988), ‘Jump-diffusion processes and the term structure of interest rates’, *The journal of Finance* **43**(1), 155–174.
- Aliprantis, C. and Border, K. (2007), *Infinite Dimensional Analysis: A Hitchhiker’s Guide*, Springer.
- Alonso-Malaver, C., Porcu, E. and Giraldo, R. (2015), ‘Multivariate and multiradial schoenberg measures with their dimension walks’, *Journal of Multivariate Analysis* **133**, 251–265.
- Arestov, V. V. and Berdysheva, E. (2001), ‘Turán’s problem for positive definite functions with supports in a hexagon’, *Trudy Instituta Matematiki i Mekhaniki UrO RAN* **7**(1), 21–29.
- Aronszajn, N. (1950), ‘Theory of reproducing kernels’, *Transactions of the American mathematical society* **68**(3), 337–404.
- Askey, R. (1973), Radial characteristics functions., Technical report, DTIC Document.
- Atkinson, K. E. and Han, W. (2012), *Spherical Harmonics and Approximations on the Unit Sphere: an Introduction*, Springer-Verlag.

- Baker, C. R. (1973), ‘On equivalence of probability measures’, *The Annals of Probability* **1**(4), 690–698, DOI [10.1214/aop/1176996895](https://doi.org/10.1214/aop/1176996895).
- Baldi, P. and Rossi, M. (2014), ‘Representation of gaussian isotropic spin random fields’, *Stochastic Processes and their Applications* **124**(5), 1910–1941.
- Baldi, P. and Trapani, S. (2013), ‘Fourier coefficients of invariant random fields on homogeneous spaces of compact groups’, *arXiv preprint arXiv:1304.5142*.
- Baldi, P., Trapani, S. et al. (2015), Fourier coefficients of invariant random fields on homogeneous spaces of compact lie groups, in ‘Annales de l’Institut Henri Poincaré, Probabilités et Statistiques’, Vol. 51, Institut Henri Poincaré, pp. 648–671.
- Barnard, R. W. and Moore, E. H. (1935), ‘General analysis. part 1’, *Memoirs of the American Philosophical Society*.
- Beatson, R. K., Castell, W. Z. and Xu, Y. (2013), ‘A polya criterion for (strict) positive-definiteness on the sphere’, *IMA Journal of Numerical Analysis* **34**(2), 550–568, DOI [10.1093/imanum/drt008](https://doi.org/10.1093/imanum/drt008).
- Bendikov, A. and Saloff-Coste, L. (2001), ‘On the absolute continuity of gaussian measures on locally compact groups’, *Journal of Theoretical Probability* **14**(3), 887–898.
- Berg, C. and Porcu, E. (2016), ‘From Schoenberg coefficients to Schoenberg functions’, *Constructive Approximation* pp. 1–25.
- Berg, C., Christensen, J. P. R. and Ressel, P. (1984), *Harmonic Analysis on Semigroups*, Springer-Verlag.
- Berg, C., Peron, A. P. and Porcu, E. (2016), ‘Orthogonal expansions related to compact Gelfand pairs’, *arXiv preprint arXiv:1612.03718*.
- Berlinet, A. and Thomas-Agnan, C. (2011), *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Springer Science & Business Media.

-
- Bevilacqua, M., Faouzi, T., Furrer, R. and Porcu, E. (2016), ‘Estimation and prediction using generalized wendland covariance functions under fixed domain asymptotics’, *arXiv:1607.06921*.
- Björk, T., Kabanov, Y. and Runggaldier, W. (1997), ‘Bond market structure in the presence of marked point processes’, *Mathematical Finance* **7**(2), 211–239.
- Bochner, S. (1948), *Vorlesungen über Fouriersche Integrale*, Chelsea Publishing Company.
- Bolin, D. and Wallin, J. (2016), ‘Spatially adaptive covariance tapering’, *Spatial Statistics*.
- Brémaud, P. (1981), ‘Point processes and queues: martingale dynamics’.
- Buehner, M. and Charron, M. (2007), ‘Spectral and spatial localization of background-error correlations for data assimilation’, *Quarterly Journal of the Royal Meteorological Society* **133**(624), 615–630.
- Chen, D., Menegatto, V. A. and Sun, X. (2003), ‘A necessary and sufficient condition for strictly positive definite functions on spheres’, *Proceedings of the American Mathematical Society* pp. 2733–2740.
- Cheng, D. (2015), ‘Excursion probabilities of isotropic and locally isotropic gaussian random fields on manifolds’, *arXiv preprint arXiv:1504.08047*.
- Cheng, D., Xiao, Y. et al. (2016), ‘Excursion probability of gaussian random fields on sphere’, *Bernoulli* **22**(2), 1113–1130.
- Cooperstein, B. (2010), *Advanced Linear Algebra*, Taylor & Francis.
- Corwin, L. and Greenleaf, F. P. (1990), *Representations of Nilpotent Lie Groups and Their Applications: Volume 1, Part 1, Basic Theory and Examples*, Vol. 18, Cambridge University Press.
- Da Prato, G. (2006), *An Introduction to Infinite-Dimensional Analysis*, Universitext, Springer Berlin Heidelberg.

- Da Prato, G. and Zabczyk, J. (1992), *Stochastic Equations in Infinite Dimensions*, Cambridge University Press.
- Dai, F. and Xu, Y. (2013), *Approximation Theory and Harmonic Analysis on Spheres and Balls*, Springer.
- Daley, D. J. and Porcu, E. (2014), ‘Dimension walks and Schoenberg spectral measures’, *Proceedings of the American Mathematical Society Proc. Amer. Math. Soc.* **142**(5), 1813–1824, DOI [10.1090/s0002-9939-2014-11894-6](https://doi.org/10.1090/s0002-9939-2014-11894-6).
- Daley, D. J. and Vere-Jones, D. (2002), *An Introduction to the Theory of Point Processes: Volume I: General Theory and Structure*, Springer Science & Business Media.
- Daley, D. J. and Vere-Jones, D. (2007), *An Introduction to the Theory of Point Processes: volume II: General Theory and Structure*, Springer Science & Business Media.
- Daniel, Z. (2007), *Table of Integrals, Series, and Products*, Academic Press.
- de Vicente, J. I., Gandy, S. and Sánchez-Ruiz, J. (2007), ‘Information entropy of gegenbauer polynomials of integer parameter’, *Journal of Physics A: Mathematical and Theoretical* **40**(29), 8345.
- Dixit, A., Jiu, L., Moll, V. H. and Vignat, C. (2015), ‘The finite fourier transform of classical polynomials’, *Journal of the Australian Mathematical Society* **98**(02), 145–160.
- Dokmanic, I. and Petrinovic, D. (2010), ‘Convolution on the-sphere with application to pdf modeling’, *IEEE Transactions on Signal Processing* **58**(3), 1157–1170.
- Ehm, W., Gneiting, T. and Richards, D. (2004), ‘Convolution roots of radial positive definite functions with compact support’, *Transactions of the American Mathematical Society* **356**(11), 4655–4685.

-
- Estrade, A., Fariñas, A. and Porcu, E. (2016), ‘Characterization theorems for covariance functions on the n - dimensional sphere across time’. preprint, Available at <https://hal.archives-ouvertes.fr/hal-01417668>.
- Fekete, M. (1935), ‘On generalized fourier series with non-negative coefficients’, *Proceedings of the London Mathematical Society*, DOI [10.1112/plms/s2-39.1.321](https://doi.org/10.1112/plms/s2-39.1.321).
- Fourier, J. (1822), *Théorie Analytique de La Chaleur*, Paris: F. Didot.Paris: F. Didot.
- Gelfand, A. E., Diggle, P., Guttorp, P. and Fuentes, M. (2010), *Handbook of Spatial Statistics*, CRC Press.
- Genton, M. G., Kleiber, W. et al. (2015), ‘Cross-covariance functions for multivariate geostatistics’, *Statistical Science* **30**(2), 147–163.
- Georgiou, T. T. (2006), ‘Distances between power spectral densities’, *arXiv preprint*.
- Georgiou, T. T. (2007), ‘Distances and riemannian metrics for spectral density functions’, *IEEE Transactions on Signal Processing* **55**(8), 3995–4003.
- Gikhman, I. I. and Skorokhod, A. V. (1966), ‘On the densities of probability measures in function spaces’, *Russian Mathematical Surveys* **21**(6), 83–156, DOI [10.1070/rm1966v021n06abeh001197](https://doi.org/10.1070/rm1966v021n06abeh001197).
- Gneiting, T. (1999), ‘On the derivatives of radial positive definite functions’, *Journal of Mathematical Analysis and Applications* **236**(1), 86–93.
- Gneiting, T. (2002), ‘Compactly supported correlation functions’, *Journal of Multivariate Analysis* **83**(2), 493–508.
- Gneiting, T. (2013a), ‘Strictly and non-strictly positive definite functions on spheres’, *Bernoulli* **19**(4), 1327–1349, DOI [10.3150/12-bejsp06](https://doi.org/10.3150/12-bejsp06).

- Gneiting, T. (2013b), ‘Strictly and non-strictly positive definite functions on spheres: online supplement.’ Available at https://projecteuclid.org/download/suppdf_1/euclid.bj/1377612854.
- Golinskii, L., Malamud, M. and Oridoroga, L. (2015), ‘Radial positive definite functions and schoenberg matrices with negative eigenvalues’, *arXiv preprint arXiv:1502.07179*.
- Gorbachev, D. V. and Manoshina, A. (2004), ‘Turán extremal problem for periodic functions with small support and its applications’, *Mathematical Notes* **76**(5-6), 640–652.
- Gradshteyn, I. S. and Ryzhik, I. M. (2007), *Table of Integrals, Series, and Products*, seventh edn, Elsevier/Academic Press, Amsterdam.
- Grafakos, L. (2008), *Classical Fourier Analysis*, Graduate Texts in Mathematics, Springer New York.
- Guinness, J. and Fuentes, M. (2013), ‘Covariance functions for mean square differentiable processes on spheres’, *Preprint*.
- Guinness, J. and Fuentes, M. (2016), ‘Isotropic covariance functions on spheres: Some properties and modeling considerations’, *Journal of Multivariate Analysis* **143**, 143–152.
- Hamill, T. M., Whitaker, J. S. and Snyder, C. (2001), ‘Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter’, *Monthly Weather Review* **129**(11), 2776–2790.
- Hershey, J. R. and Olsen, P. A. (2007), Approximating the kullback leibler divergence between gaussian mixture models, Vol. 4, IEEE, pp. IV–317.
- Heyer, H. (2012), *Probability Measures on Locally Compact Groups*, Vol. 94, Springer Science & Business Media.
- Hunt, G. A. (1956), ‘Semi-groups of measures on lie groups’, *Transactions of the American Mathematical Society* **81**(2), 264–293.

-
- Ibragimov, I. and Rozanov, Y. (1978), *Gaussian Random Processes*, Springer-Verlag.
- Jean, D. (1980), *Special Functions and Linear Representations of Lie Groups*, Published for Conference Board of the Mathematical Sciences by the American Mathematical Society.
- Jeffrey, A. and Zwillinger, D. (2007), *Table of Integrals, Series, and Products*, Academic Press.
- Jiang, X., Ning, L. and Georgiou, T. T. (2012), ‘Distances and riemannian metrics for multivariate spectral densities’, *IEEE Transactions on Automatic Control* **57**(7), 1723–1735.
- Johnson, C. (1970), ‘Positive definite matrices’, *The American Mathematical Monthly* **77**(3), 259–264.
- Jorgensen, P., Pedersen, S. and Tian, F. (2016), *Extensions of Positive Definite Functions: Applications and Their Harmonic Analysis*, Lecture Notes in Mathematics, Springer International Publishing.
- Kailath, T. (1986), *Modern Signal Processing*, CRC Press.
- Kaufman, C. G. and Shaby, B. A. (n.d.), ‘The role of the range parameter for estimation and prediction in geostatistics’, DOI [10.1093/biomet/ass079](https://doi.org/10.1093/biomet/ass079).
- Kaufman, C. G., Schervish, M. J. and Nychka, D. W. (2008), ‘Covariance tapering for likelihood-based estimation in large spatial data sets’, *Journal of the American Statistical Association* **103**(484), 1545–1555.
- Kennedy, R. A. and Sadeghi, P. (2013), *Hilbert Space Methods in Signal Processing*, Cambridge University Press.
- Kirillov, A. A. (2004), *Lectures on the Orbit Method*, Vol. 64, American Mathematical Society Providence.

- Krasnitskii, S. (1994), ‘On necessary conditions of equivalence of gaussian measures corresponding to homogeneous random fields’, *Ukrainian Mathematical Journal* **46**(6), 765–778.
- Kühn, T. and Liese, F. (1979), ‘A short proof of the hájek-feldman theorem’, *Theory of Probability & Its Applications* **23**(2), 429–431, DOI [10.1137/1123052](https://doi.org/10.1137/1123052).
- Kuo, H.-H. (1975), *Gaussian Measures in Banach Spaces*, Springer-Verlag.
- Lang, A. and Schwab, C. (2015), ‘Isotropic gaussian random fields on the sphere: Regularity, fast simulation and stochastic partial differential equations’, *Ann. Appl. Probab.* **25**(6), 3047–3094, DOI [10.1214/14-AAP1067](https://doi.org/10.1214/14-AAP1067).
- Lang, S. (1993), *Real and Functional Analysis*, Springer-Verlag.
- Lévy, P. (1961), ‘Quelques problèmes non résolus de la théorie des fonctions caractéristiques’, *Annali di Matematica Pura ed Applicata, Series 4* **53**(1), 315–331, DOI [10.1007/bf02417802](https://doi.org/10.1007/bf02417802).
- Lloyd, J. W. (2003), Learning comprehensible theories from structured data, *in* ‘Advanced lectures on machine learning’, Springer, pp. 203–225.
- Ma, C. (2012), ‘Stationary and isotropic vector random fields on spheres’, *Mathematical Geosciences* **44**(6), 765–778.
- Ma, C. (2015), ‘Isotropic covariance matrix functions on all spheres’, *Mathematical Geosciences* **47**(6), 699–717.
- Malyarenko, A. (2012), *Invariant Random Fields on Spaces with a Group Action*, Springer Science & Business Media.
- Marinucci, D. and Peccati, G. (2011), *Random Fields on the Sphere*, Cambridge University Press.
- Marks, R. J. (2009), *Handbook of Fourier Analysis & its Applications*, Oxford University Press.

- Matérn, B. (1960), *Spatial Variation : Stochastic Models and their Application to some Problems in Forest Surveys and other Sampling Investigations*, Statens Skogsforskningsinstitut. Diss. Sthlm Univ.
- Matheron, G. (1963), ‘Principles of geostatistics’, *Economic geology* **58**(8), 1246–1266.
- Mathias, M. (1923), ‘Über positive fourier-integrale’, *Mathematische Zeitschrift* **16**(1), 103–125.
- Ménard, R. (2000), ‘Tracer assimilation’, *Inverse Methods in Global Biogeochemical Cycles* pp. 67–79.
- Menegatto, V. A. (1994), ‘Strictly positive definite kernels on the hilbert sphere’, *Applicable Analysis* **55**(1-2), 91–101.
- Møller, J., Nielsen, M., Porcu, E. and Rubak, E. (2015), ‘Determinantal point process models on the sphere’, *Centre for Stochastic Geometry and Advanced Bioimaging*.
- Morimoto, M. (1998), *Analytic Functionals on the Sphere*, American Mathematical Society.
- Parthasarathy, K. (1976), *Probability Measures on Metric Spaces*, AMS Chelsea Publishing Series, Academic Press.
- Peron, A. P. and Porcu, E. (2016), ‘Positive definite functions on complex spheres, and their walks through dimensions’, *Submitted*.
- Pesenson, I. (1999), ‘A reconstruction formula for band limited functions in $\mathbf{L}^2(\mathbb{R}^d)$ ’, *Proceedings of the American Mathematical Society* **127**(12), 3593–3600.
- Pigoli, D. (2012), Distance-based statistics for covariance operators in functional data analysis, in ‘46TH SCIENTIFIC MEETING OF THE ITALIAN STATISTICAL SOCIETY’.

- Pigoli, D., Menafoglio, A. and Secchi, P. (2016), ‘Kriging prediction for manifold-valued random fields’, *Journal of Multivariate Analysis* **145**, 117–131, DOI [10.1016/j.jmva.2015.12.006](https://doi.org/10.1016/j.jmva.2015.12.006).
- Porcu, E., Bevilacqua, M. and Genton, M. G. (2016), ‘Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere’, *Journal of the American Statistical Association* **111**(514), 888–898.
- Pothoven, K. (2013), *Real and Functional Analysis*, Springer Science & Business Media.
- Preston, C. (1976), *Random Fields*, Lecture Notes in Mathematics, 1 edn, Springer.
- Ramaswami, S. (1974), ‘Semigroups of measures on lie groups’, *J. Indian Math. Soc* **38**, 175–189.
- Rao, C. R. and Varadarajan, V. S. (1963), ‘Discrimination of gaussian processes’, *Sankhyā: The Indian Journal of Statistics, Series A* pp. 303–330.
- Rasmussen, C. E. (2006), *Gaussian Processes for Machine Learning*, Citeseer.
- Ressel, P. and Ricker, W. (1998), ‘Semigroup representations, positive definite functions and abelian C^* -algebras’, *Proceedings of the American Mathematical Society* **126**(10), 2949–2955.
- Révész, S. G. (2011), ‘Turán’s extremal problem on locally compact abelian groups’, *Analysis Mathematica* **37**(1), 15.
- Rudin, W. (1986), *Real and Complex Analysis (3rd)*, New York: McGraw-Hill Inc.
- Rudin, W. (2011), *Fourier Analysis on Groups*, John Wiley & Sons.
- Ruiz-Medina, M. D. and Porcu, E. (2014), ‘Equivalence of gaussian measures of multivariate random fields’, *Stochastic Environmental Research and Risk Assessment* **29**(2), 325–334, DOI [10.1007/s00477-014-0926-z](https://doi.org/10.1007/s00477-014-0926-z).

-
- Sasvári, Z. (2013), *Multivariate Characteristic and Correlation Functions*, Vol. 50, Walter de Gruyter.
- Schoenberg, I. J. (1938), ‘Metric spaces and completely monotone functions’, *Ann. Math.* **39**(4), 811–841.
- Schoenberg, I. J. (1942), ‘Positive definite functions on spheres’, *Duke Math. J.* **9**(1), 96–108, DOI [10.1215/s0012-7094-42-00908-6](https://doi.org/10.1215/s0012-7094-42-00908-6).
- Sejdinovic, D. and Gretton, A. (2012), ‘What is an RKHS?’.
- Skorohod, A. (1967), ‘On the densities of probability measures in functional spaces’, *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66) Vol. II: Contributions to Probability Theory* **1**, 163–182.
- Skorokhod, A. V. and Yadrenko, M. I. (1973), ‘On absolute continuity of measures corresponding to homogeneous gaussian fields’, *Theory of Probability & Its Applications* **18**(1), 27–40, DOI [10.1137/1118002](https://doi.org/10.1137/1118002).
- Soubeyrand, S., Enjalbert, J. and Sache, I. (2008), ‘Accounting for roughness of circular processes: Using gaussian random processes to model the anisotropic spread of airborne plant disease’, *Theoretical Population Biology* **73**(1), 92–103, DOI [10.1016/j.tpb.2007.09.005](https://doi.org/10.1016/j.tpb.2007.09.005).
- Spodarev, E. (2013), *Stochastic Geometry, Spatial Statistics and Random Fields: Asymptotic Methods*, Vol. 2068, Springer.
- Stein, M. (1990), ‘Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure’, *Ann. Statist.* **18**(2), 850–872, DOI [10.1214/aos/1176347629](https://doi.org/10.1214/aos/1176347629).
- Stein, M. L. (1988), ‘Asymptotically efficient prediction of a random field with a misspecified covariance function’, *The Annals of Statistics* **16**(1), 55–63.
- Stein, M. L. (1993), ‘A simple condition for asymptotic optimality of linear predictions of random fields’, *Statistics & Probability Letters* **17**(5), 399–404, DOI [10.1016/0167-7152\(93\)90261-g](https://doi.org/10.1016/0167-7152(93)90261-g).

- Stein, M. L. (1999a), *Interpolation of Spatial Data: some Theory for Kriging*, Springer.
- Stein, M. L. (1999b), ‘Predicting random fields with increasing dense observations’, *The Annals of Applied Probability* **9**(1), 242–273, DOI [10.1214/aoap/1029962604](https://doi.org/10.1214/aoap/1029962604).
- Stein, M. L. (2004), ‘Equivalence of gaussian measures for some nonstationary random fields’, *Journal of Statistical Planning and Inference* **123**(1), 1–11.
- Steinwart, I. and Christmann, A. (2008), *Support Vector Machines*, Springer Science & Business Media.
- Sugiura, M. (1990), *Unitary Representations and Harmonic Analysis: an Introduction*, Vol. 44, Elsevier.
- Szegő, G. (1939), *Orthogonal Polynomials*, American Mathematical Society.
- Taylor, J. E. and Adler, R. J. (2003), ‘Euler characteristics for gaussian fields on manifolds’, *Annals of Probability* pp. 533–563.
- Trübner, M. and Ziegel, J. F. (2016), ‘Derivatives of isotropic positive definite functions on spheres’, *arXiv preprint arXiv:1603.06727*.
- Ullah, A. (1996), ‘Entropy, divergence and distance measures with econometric applications’, *Journal of Statistical Planning and Inference* **49**(1), 137–162, DOI [10.1016/0378-3758\(95\)00034-8](https://doi.org/10.1016/0378-3758(95)00034-8).
- Voit, M. (1997), ‘Rate of convergence to gaussian measures on n-spheres and jacobi hypergroups’, *The Annals of Probability* pp. 457–477.
- Wang, Y. G., Gia, Q. T. L., Sloan, I. H. and Womerlsey, R. S. (2015), ‘Needlet approximation for isotropic random fields on the sphere’, *arXiv preprint arXiv:1512.07790*.
- Wang, Y. G., Le Gia, Q. T., Sloan, I. H. and Womersley, R. S. (2016), ‘Fully discrete needlet approximation on the sphere’, *Applied and Computational Harmonic Analysis*, DOI [10.1016/j.acha.2016.01.003](https://doi.org/10.1016/j.acha.2016.01.003).

-
- Wendland, H. (1995), ‘Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree’, *Advances in computational Mathematics* **4**(1), 389–396.
- Yadrenko, M. I. and Balakrishnan, A. V. (1983), *Spectral Theory of Random Fields*, Optimization Software, Publications Division.
- Zhang, H. (2004), ‘Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics’, *Journal of the American Statistical Association* **99**(465), 250–261, DOI [10.1198/016214504000000241](https://doi.org/10.1198/016214504000000241).
- Ziegel, J. (2014), ‘Convolution roots and differentiability of isotropic positive definite functions on spheres’, *Proceedings of the American Mathematical Society* **142**(6), 2063–2077, DOI [10.1090/s0002-9939-2014-11989-7](https://doi.org/10.1090/s0002-9939-2014-11989-7).



Your Questions
& Comments
Are Always
Welcome

