Imperial College of Science, Technology and Medicine
Department of Mathematics

Aspects of Positive Definiteness and Gaussian Processes on Planet Earth

Tasmin L. Symons

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Mathematics of Imperial College, August 2019

Abstract

This thesis studies characterisations and properties of spatial and spatio-temporal Gaussian processes defined over the sphere $\mathbb{S}^d$ (or in the spatio-temporal case the product of the sphere and the real line $\mathbb{S}^d \times \mathbb{R}$. Such processes are of importance in global weather and climate science, where the geometry is necessarily spherical, but, especially in the dynamic setting, they are less well-studied than their Euclidean counterparts.

Beginning with Brownian motion, we first look at characterising Gaussian randomness on $\mathbb{S}^d$ and $\mathbb{S}^d \times \mathbb{R}$, and how it compares with the Euclidean setting – we show that the characterisation theorems of Gaussian processes on spaces of types spanning $\mathbb{R}^d$, $\mathbb{S}^d$ and $\mathbb{S}^d \times \mathbb{R}$ can be phrased as consequences of a powerful general theorem of harmonic analysis. We go on to find the answer to a recent question posed about dimension-hopping operators for positive-definite (i.e. covariance) functions on $\mathbb{S}^d \times \mathbb{R}$, and consider how we could go about constructing dimension-hopping operators with the semi-group property on the sphere.

Later, we address the theory of the path properties of these processes, extending a finite-dimensional result the the infinite-dimensional case and showing that a remarkably elegant approach for processes on $\mathbb{R}^d$ carries over to our setting. We finish by finding the analogue of the powerful Ciesielski isomorphism for continuous functions on the two-sphere.

Foreword

Much of the work in this thesis has already been published, or submitted for publication.

1. Chapter 1 summarises, as background and motivation for what follows, the contents of:

   N. H. Bingham, Aleksandar Mijatović and **Tasmin L. Symons**, Brownian Manifolds, Negative Type and Geotemporal Covariances, *Communications on Stochastic Analysis* **10**(4) (2016), 421 – 432. `DOI: 10.31390/cosa.10.4.03`

2. Chapter 2 expands on my work in:

   N. H. Bingham and **Tasmin L. Symons**, Probability, Statistics and Planet Earth I: The Bochner-Godement theorem and Geotemporal Covariances. Submitted to *Probability Surveys*.

3. The results in Chapter 3, Sections 1 and 2, have been published in:

   N. H. Bingham and **Tasmin L. Symons**, Dimension walks on $\mathbb{S}^d \times \mathbb{R}$. *Statistics and Probability Letters* **47** (2019), 12 – 17. `DOI:10.1016/j.spl.2018.11.014`

4. The work in Chapter 4 has been written for publication in:

   N. H. Bingham and **Tasmin L. Symons**, Aspects of Gaussian processes on spheres. Submitted to *Stochastic Processes and their Applications* (Larry Shepp Memorial Special Issue).

Declaration

I declare that all the work included in this thesis is, to the best of my knowledge, original unless otherwise attributed.

Tasmin L. Symons; August 2019.

Acknowledgements

Firstly, I gratefully acknowledge the financial support of EPSRC and their Centre for Doctoral Training in the Mathematics of Planet Earth (MPECDT).

MPECDT has been much more to me than financial support, though. Thank you to the staff: I can only hope to do you all proud in my future career. And thank you to my fellow students – particularly Jemima Tabeart – for all the love and laughs. I have made friends for life.

I thank my parents, brother, and friends (Dan, Charlie, Owen, Cat, Sophie, Mike, Jack, Eleanor, the list continues) for the often timely reminders that there is more to life than mathematics.

I applaud Adam Ostaszewski, Emma McCoy, Aleksandar Mijatović, Tom Cass, Persi Diaconis, Clare O'Neill and many, many, others for their astute and intelligent contributions.

Singular amongst men in his generosity, intellect, and frankly quite staggering amount of patience is my supervisor, Professor Nicholas H. Bingham. Saying thank you barely begins to cover it, Nick.

Finally – most importantly – thank you to Imperial College London, for being *home* for eight years.

"..a fact is the most stubborn thing in the world."

Mihail Bulgakov, *The Master & Margarita*.

Contents

Introduction

Described as a 'beautiful marriage' by Gelfand and Schliep [GS16], the combination of Gaussian processes with spatial statistics is 'the most valuable tool in the toolkit for geostatistical modeling.' The literature agrees – Gaussian random fields over (subsets of) $\mathbb{R}^2$ have become ubiquitous in spatial statistics, with applications ranging from disease mapping [Bat+19] to understanding patterns in terrorism [Pyt+19].

In light of the scientific challenges posed by monitoring our changing planet, our interest is driven by a desire to model environmental phenomena – our weather, oceans, and climate. Gaussian processes are useful here in a variety of ways, including building statistical emulators of more expensive dynamics-based models [PAF18]. The governing physical processes here are supremely complex. They are also intrinsically global. The mathematical objects of interest, then, are Gaussian processes defined on the sphere $\mathbb{S}^2$, or, when the process of interest evolves with time, $\mathbb{S}^2 \times \mathbb{R}$.

It is the study of these processes which drives this thesis. Much of the theory of Gaussian processes on a d-dimensional sphere is classical, but the area is active and growing apace. The theory of Gaussian processes which evolve 'geo-temporally' – in space on the sphere and in linear time – is still very much in its infancy. This thesis aims to contribute to both theories, and highlight how they complement one another.

The plan of this thesis is as follows. In Chapter 1 we consider the prototypical Gaussian process, Brownian motion, parametrised by our parameter spaces of interest – $\mathbb{S}^d$ and $\mathbb{S}^d \times \mathbb{R}$. The theory here is largely geometric in nature, as is the question of whether a given manifold is *Brownian* (i.e. supports a Brownian motion). The sphere $\mathbb{S}^d$ is Brownian, a result well-known and due to Lévy [Lév59]. Whether $\mathbb{S}^d \times \mathbb{R}$ is Brownian, however, depends entirely on the manner in which $\times$ is interpreted – if we consider it as a product manifold then Brownian motion, surprisingly, does not exist.

Motivated by this, in Chapter 2 we consider characterising more general Gaussian processes on $\mathbb{S}^d$ and $\mathbb{S}^d \times \mathbb{R}$. As (mean zero) Gaussian processes are defined entirely by their covariance, this is equivalent to characterising positive definite functions on $\mathbb{S}^d$ and $\mathbb{S}^d \times \mathbb{R}$. We find an alternative formulation of a recent result of Berg and Porcu [BP17] characterising the positive definite functions on $\mathbb{S}^d \times \mathbb{R}$. Our formulation follows elegantly from the classical Bochner-Godement theorem. We also discuss the challenges of characterising *strict* positive definiteness on $\mathbb{S}^d \times \mathbb{R}$.

Chapter 3 concerns *walks on dimensions*, dimension-hopping operators which preserve positive definiteness. These are useful in defining new parametric families of covariances for applications. We extend a pair of papers by Beatson and zu Castell [BC17; BC16] to the geo-temporal case, and explore an alternative dimension-hopping operator based on properties of the ultraspherical polynomials.

In Chapter 4 we discuss the fascinating path properties of Gaussian random fields on $\mathbb{S}^d$ and $\mathbb{S}^d \times \mathbb{R}$. This is a rapidly growing area of research, and, after obtaining Belyaev's dichotomy for fields on the sphere, we survey two approaches to quantifying the regularity of a spherical process's paths: one concerns asymptotic estimates [Mal13], the other integrability conditions [LS15]. We extend a theorem of Malyrenko's to the Hilbert sphere $\mathbb{S}^\infty$, and show that the Schilling's [Sch00] remarkable approach to proving the Kolmogorov-Chentsov theorem is valid for processes parametrised by $\mathbb{S}^d$.

Finally, in Chapter 5 we find a Ciesielski isomorphism – an isomorphism between a function space and a sequence space – for functions in $C(\mathbb{S}^2)$. In the classical setting the Ciesielski isomorphism has become an invaluable tool, but the Faber-Schauder construction used to prove the isomorphism [Sem82] requires some adapting to the spherical setting. This is addressed here.

We finish with a brief summary and some ideas for future directions of research.

Chapter 1

Brownian Manifolds

A stochastic process, $X = \{X_t\}$ is a mathematical model for a random phenomenon evolving temporally, with time t. In many applications it is preferable for the relevant parameter to be a point in space – a *spatial process*, or a *random field*; sometimes we need both time and space – a *spatio-temporal* process. Our main interest here is the case when space is a sphere (a model for the Earth). In this case we will speak of a *geo-temporal process*.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be our probability space, and $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ be a measurable space with measure μ (more general measure spaces are permissible, but in this thesis we will focus on real-valued random fields).

Definition 1.1. *A stochastic process parametrised by an index set M on $(\Omega, \mathcal{F}, \mathbb{P})$ (we will usually say 'a stochastic process on M', leaving the probability space to one side) is a family $\{X_t : t \in M\}$ of random variables.*

That is, for each $t \in M$, X_t is a measurable mapping from $(\Omega, \mathcal{F}, \mathbb{P})$ *to* $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. *For fixed $\omega \in \Omega$ the set $\{X_t(\omega) : t \in M\}$ is the* path *of the field.*

The simplest manifold that might be used for the space variable is Euclidean space, of dimension d say, $\mathbb{R}^d$. In the (particularly classical) literature this set-up is sometimes referred to as multi-dimensional (d-dimensional) time.

Throughout this thesis we concern ourselves with *Gaussian random fields* on the sphere $\mathbb{S}^d$ and on the sphere-cross-line $\mathbb{S}^d \times \mathbb{R}$. As the prototypical Gaussian process is Brownian motion, an exploration of Brownian motion in these two scenarios is of value. The work in this chapter has been published in colloboration with N. H. Bingham and Aleksandar Mijatović as [BMS16] and is presented here with their kind permission as background and motivation for the work on more general Gaussian processes to follow in Chapters 2 – 5.

1.1 Brownian manifolds and negative type

1.1.1 Lévy's Brownian motion with multi-dimensional time

A sensible definition of Brownian motion – the stochastic process describing the motion of particles suspended in fluid – is as the centred Gaussian process

$B = (B_t : t \in \mathbb{R})$ with *incremental variance*

$$i(s,t) := \mathbb{E}[(B_t - B_s)^2] = |t - s| \tag{1.1}$$

and $B_0 = 0$. It is helpful, here, to regard B as a map $t \mapsto B_t$ from $\mathbb{R}$ to the Hilbert space $H = L^2(\Omega, \mathcal{F}, P)$, so we have

$$\|B_t - B_s\|^2 = |t - s|, \tag{1.2}$$

the left-hand side being the incremental variance. The covariance is then given by the inner product

$$c(t,s) := \langle B_t, B_s \rangle = \frac{1}{2}(|t| + |s| - |t - s|). \tag{1.3}$$

The linearity of the inner product shows that increments $B_{t_n} - B_{t_{n-1}}$, $B_{t_{n-1}} - B_{t_{n-2}}, \ldots, B_{t_1} - B_{t_0}$ are uncorrelated and thus are independent, by the Gaussianity of B.

Defining the Brownian covariance as an inner product allows us to extend to multi-dimensional processes. Lévy [Lév48] showed that one can define Brownian motion with multi-dimensional time (or "multi-parameter Brownian motion", or, in two dimensions, a "Brownian sheet") as the real-valued centred Gaussian process $B = (B_t : t \in \mathbb{R}^d)$ with incremental variance given by the $d-$dimension equivalent of (1.1). For later treatments, see also [Lév59; Lév66].

We have

$$i(s,t) = c(s,s) + c(t,t) - 2c(s,t), \tag{1.4}$$

and as $c(s,s) = E[B_s^2] = i(s,0)$,

$$c(s,t) = \frac{1}{2}(i(s,0) + i(t,0) - i(s,t)). \tag{1.5}$$

Thus either of c, i determines the other; i is more convenient here.

Lévy also showed that Brownian motion can be defined so as to be parametrised by the sphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$, in addition to $\mathbb{R}^d$ as above. Now the incremental variance is given by the geodesic distance d on the sphere (from the North Pole o, which plays the role of the origin above):

$$i(s,t) = \|B_s - B_t\|^2 = d(s,t). \tag{1.6}$$

Thus $\sqrt{d}(s,t) = \|B_t - B_s\|$; one calls $\sqrt{d}$ a *Hilbertian distance.*

A word on terminology: our incremental variance is also known by several other names: the *variogram* (a term due to Matheron, arising in mining), the structure function (Yaglom), mean-squared difference (Jowett), etc.; see e.g. Cressie ([Cre93, §2.3.1]).

1.1.2 Brownian and non-Brownian manifolds

For M a manifold with geodesic distance d, or more generally with (M, d) a metric space, one can proceed as above and call $B = (B_x : x \in M)$ a Brownian motion *parametrised* by M if the B_x are centred Gaussian, the incremental variance is the geodesic distance,

$$\mathbb{E}[(B_x - B_y)^2] = d(x, y), \tag{1.7}$$

and the finite-dimensional distributions are Gaussian (that is, linear combinations $\sum c_i B_{t_i}$ are Gaussian). Then (1.6) above is satisfied with d the geodesic distance on M. Call such a manifold, or metric space, *Brownian*.

Euclidean space and spheres are Brownian, by Lévy's results above. Further examples are given by the real or complex hyperbolic spaces, a result due to Faraut and Harzallah ([FH74, Prop. 7.3.]) (and implicit in Gangolli [Gan67]). By contrast, quaternionic hyperbolic spaces are not Brownian ([Far73, Cor. IV.2] or [FH74]), and nor is the octonion (Cayley) projective plane.

The question of whether a space M is Brownian is thus purely geometric, as it depends on whether a map B exists satisfying (1.6).

1.1.3 Spaces and kernels of negative type

Definition 1.2. *A metric space* (M,d) *is of negative type if*

$$\sum\nolimits_{i,j=1}^{n} d(x_i,x_j)u_iu_j \leq 0 \tag{1.8}$$

for all $n = 2,3,\cdots$, *all points* $x_i \in M$ *and all real* u_i *with* $\sum u_i = 0$ *(the term conditionally negative definite is also used, reflecting the condition* $\sum u_i = 0$*). Call* M *of* strictly negative type *if the sum above is negative for all such* u_i *not all zero.*

Such spaces are important in a variety of contexts, and have been studied at length in the books by Blumenthal [Blu70] and Deza and Laurent [DL97].

A function $k : X \times X \to \mathbb{R}$ mapping a pair of inputs to a real number is often referred to as a kernel – the term comes from the theory of integral operators, where kernel functions act as weights $(T_k f)(x) := \int k(x,y)f(y)d\mu(y)$. We say a kernel is of *negative type* if

$$\sum\nolimits_{i,j=1}^{n} k(x_i,x_j)u_iu_j \leq 0 \tag{1.9}$$

for all $n = 2,3,\cdots$, all points $x_i \in M$ and all real u_i with $\sum u_i = 0$, and of *positive type* (or positive definite) if

$$\sum\nolimits_{i,j=1}^{n} k(x_i,x_j)u_iu_j \geq 0 \tag{1.10}$$

for all $n = 2, 3, \cdots$ and all points $x_i \in M$; similarly for *strictly positive type.*

Incremental variances i are of negative type: using (1.4)

$$\sum_{i,j=1}^{n} i(x_i, x_j)u_i u_j = \sum_{i,j=1}^{n} \left[c(x_i, x_i) + c(x_j, x_j) - 2c(x_i, x_j)\right] u_i u_j \tag{1.11}$$

$$= \sum_{i=1}^{n} c(x_i, x_i)u_i \sum_{j=1}^{n} u_j + \sum_{j=1}^{n} c(x_j, x_j)u_j \sum_{i=1}^{n} u_i$$

$$- 2\sum_{i,j=1}^{n} c(x_i, x_j)u_i u_j \tag{1.12}$$

$$= - 2\sum_{i,j=1}^{n} c(x_i, x_j)u_i u_j \leq 0, \tag{1.13}$$

since $\sum u_i = 0$ and covariances c are of positive type.

1.1.4 Schoenberg's theorems

It was shown by Schoenberg [Sch37; Sch38] that a metric space (M, d) is of negative type if and only if there is a map $\phi : M \to H$ for some Hilbert space H with

$$d(x, y) = \|\phi(x) - \phi(y)\|^2. \tag{1.14}$$

Thus, when $H = L^2(\Omega, \mathcal{F}, P)$ as before, M is Brownian if and only if it is of negative type, and then Brownian motion B on (parametrised by) M is the map ϕ above. Then $\phi : (M, \sqrt{d}) \to H$ is called the *Brownian embedding* (or just, embedding). See Lyons [Lyo13] for a short proof of Schoenberg's theorem.

The other classical result of Schoenberg relevant here [Sch38] is that a kernel k is of negative type iff e^{-tk} is of positive type for every $t \geq 0$. This suggests the Lévy-Khintchine formula, and was part of Gangolli's motivation for his theory of *Lévy-Schoenberg kernels* [Gan67].

1.1.5 Geo-temporal covariances

We now turn to less classical matters. When modelling phenomena on the globe, we need both a space coordinate (on the sphere) and a time coordinate (on the line or half-line); thus the space $M = \mathbb{S} \times \mathbb{R}$ (or $M = \mathbb{S} \times \mathbb{R}_+$) is needed. The most basic Gaussian process one might wish to model on M is Brownian motion. But the product can be taken in several different senses, and it turns out that the question of *existence* of Brownian motion depends on which kind of product we take. Recall that by Lévy's results, Brownian motion exists on both $\mathbb{S}$ and $\mathbb{R}$ (or $\mathbb{R}_+$), since both are of negative type.

First, take the product of metric spaces, under Hamming distance ("city-block metric"), under which distances s add:

$$s := s_1 + s_2, \tag{1.15}$$

in the obvious notation. From the definition of negative type, this property is preserved under such products; see e.g. [Blu70, §3.2]. So Brownian motion on the sphere cross line exists, with the product taken in this sense.

Next, one can take the product under the ordinary cartesian (or pythagorean) rule:

$$s^2 := {s_1}^2 + {s_2}^2. \tag{1.16}$$

Here again, Brownian motion exists. McKean [McK63] gives a thorough study of the white-noise case (from which the Brownian case follows by integration), starting from the work of Chentsov [Che57] on white noise in this setting. McKean's construction moves between Euclidean space $\mathbb{R}^{d+1}$ and 'sphere cross half-line', $\mathbb{S}^d \times \mathbb{R}_+$.

By contrast, if one takes the cartesian product of two Riemannian manifolds, distance is given by the differential cartesian rule [Lee18, p. 20]:

$$ds^2 := ds_1^2 + ds_2^2, \tag{1.17}$$

again in the obvious notation. It turns out that $M = \mathbb{S} \times \mathbb{R}$ is no longer of negative type – *so is no longer Brownian* – viewed as a manifold in this way. The same holds for any product of manifolds with at least one spherical factor – or even a factor with two pairs of antipodal points. This is purely geometric, rather than probabilistic – see [HKM02] for background and details.

Thus Brownian motion exists on $M = \mathbb{S}^d \times \mathbb{R}$, regarded as a product of metric spaces in both the above senses, though not of Riemannian manifolds. This provides a route to geo-temporal modelling – but separates the effects of space and time.

1.1.6 Remarks

1. Testing for independence

Ideas closely related to the above have found applications in statistics, in the work of Székely, Rizzo and Bakirov [SRB07] and Székely and Rizzo [SR09]. See in particular the extensive commentary to the invited paper [SR09]. Their work introduces the concept of *distance correlation* (later simplified by Lyons, below) to develop non-parametric tests of independence for high-dimensional random vectors without needing to invert large matrices at (prohibitively) large computational cost. Distance covariance, defined for two random vectors X and Y with characteristic functions f_X, f_Y and joint characteristic function $f_{X,Y}$ by $\mathcal{V}(X,Y) = ||f_{X,Y}(t,s) - f_X(t)f_X(s)||^2$ for a suitable choice of norm $||.||$ (see [SR09, §2.2]), is shown to be a natural extension of Pearson's correlation, with the crucial distinction that zero distance correlation *implies independence.*

The test statistic developed by Székely and Rizzo depends only on the distance between observations: given a bivariate sample $((X_1,Y_1),\cdots,(X_n,Y_n))$, where each coordinate has finite mean, it turns out that one can test for independence of the X- and Y-coordinates, consistently against all alternatives (again, with finite means) by test statistics involving only *distances* between observations.

2. Distance covariance

The theory of distance covariance in metric spaces has been re-worked and simplified by Russell Lyons [Lyo13; Lyo14]. The crux is, as above, that the distance covariance of (X, Y) is zero if and only if X and Y are independent. It turns out that this does not hold for general metric spaces, but does so exactly for those of *strong negative type*, a class that includes Euclidean spaces.

3. Other approaches

The first person to use white-noise integrals for Lévy's Brownian motion was Chentsov [Che57], an approach later taken up by Lévy himself [Lév66], and McKean [McK63].

4. Hypergroups

The theory of hypergroups is by now well established, but too extensive for us to discuss here. We refer to the standard work on the subject by Bloom and Heyer [BH94]. Hypergroups make contact with the work studied here, for instance through our main example, the symmetric spaces of rank one; these have constant curvature κ. For the spherical case $\kappa > 0$, the relevant hypergroup here is the Bingham (or Bingham-Gegenbauer) hypergroup [Bin72b], [BH94, p. 3.4.23]. For the Euclidean case $\kappa = 0$, it is the Kingman (or Kingman-Bessel) hypergroup [BH94, p. 3.4.30]. For the hyperbolic case $\kappa < 0$, it is the hyperbolic hypergroup [BH94, p. 3.5.68].

5. Markov property

In one dimension, the Markov property is expressed by present time being a *splitting time*: past and future are *conditionally independent given the present*. In higher dimensions, the geometry is necessarily more complicated.

In the plane, for example, one might have values within and without a contour conditionally independent given values on the contour.

6. Fractional processes

Brownian motion is too smooth for some purposes, and may be usefully generalised to fractional Brownian motion, which has a parameter that governs the degree of smoothness. Such fractional Gaussian fields have been studied in contexts related to ours by Gelbaum [Gel14].

7. Higher dimensions

It is of interest to see what happens to the n-dimensional spheres and hyperbolic spaces considered here as the dimension $n \to \infty$. There has been much studied in recent decades, due largely to Olshanski, Okounkov and Vershik. For background and details, see several recent papers by Jacques Faraut, e.g. [Far12].

1.2 Preliminaries from Riemannian geometry and harmonic analysis

We have established that $\mathbb{S}^d \times \mathbb{R}$, viewed as a manifold, does not support Brownian motion. This raises the question: what kinds of Gaussian randomness can exist in space-time, when the space is a sphere? Answering this boils down to the theory of spherical functions on symmetric spaces and, pending the answer in Chapter 2, we sketch the theory we will need from geometry and harmonic analysis below.

1.2.1 Spaces of constant curvature

We refer the reader to [Tu07] for an introduction to general manifolds, and satisfy ourselves here by specialising to the three familiar examples of Riemannian manifolds of constant curvature κ:

1. $\kappa = 0$: Euclidean space $\mathbb{R}^d$;
2. $\kappa > 0$: spheres $\mathbb{S}^d := \{x \in \mathbb{R}^{d+1} : ||x|| = 1\}$;
3. $\kappa < 0$: hyperbolic space $\mathbb{H}^d$.

As noted above, one can extend Lévy's results on Brownian motion on $\mathbb{R}^d$ and $\mathbb{S}^d$ to $\mathbb{H}^d$: we summarise this by saying that $\mathbb{S}^d, \mathbb{R}^d, \mathbb{H}^d$ are *Brownian manifolds*: they can be index spaces for Brownian motion.

These three families are the main examples of Riemannian *symmetric spaces* (below) of *rank one*. By contrast, the other examples are not Brownian: see the comment on the Kazhdan property below.

1.2.2 Symmetric spaces

A *symmetric space* (Helgason [Hel1,2,3,4], Wolf [Wol1,2]) is a Riemannian manifold M whose curvature tensor is invariant under parallel translation.

These are the spaces where at each point x the *geodesic symmetry* exists: this fixes x and reverses the (direction of) geodesics through x, an involutive automorphism [Wol2, Ch. 11]. Then M is a Riemannian homogeneous space $M = G/K$, with G a closed subgroup of the isometry group of M containing the transvections, and K the isotropy subgroup of G fixing the base-point x; (G, K) is called a Riemannian symmetric pair. The Banach algebra $L_1(K\backslash G/K)$ of (Haar) integrable functions on G bi-invariant under K is *commutative*. Such pairs are called *Gelfand pairs*, and such Banach algebras are called *commutative spaces* [Wol07]. We remark that the product of two symmetric spaces is itself symmetric.

1.2.3 Spherical functions

For harmonic analysis on symmetric spaces, one needs (cf. the Fourier transform in Euclidean space and the Gelfand transform for Banach algebras) spherical measures, spherical functions and the spherical transform [Wol07, Ch. 8, 9]. For (G, K) a Gelfand pair (with G at least locally compact, which covers all our cases of interest), a spherical measure m is a K-bi-invariant multiplicative linear functional on $C_c(K\backslash G/K)$; a *spherical function* is a continuous function $\phi : G \to \mathbb{C}$ with the measure $m_\phi(f) := \int_G f(x)\phi(x^{-1})d\mu_G(x)$ spherical. The *spherical transform* for (G, K) is the map

$$f \mapsto \hat{f}(\phi) := m_\phi(f) = \int_G f(x)\phi(x^{-1})d\mu_G(x). \tag{1.18}$$

The *positive definite* spherical functions ϕ on (G, K) are in bijection with the irreducible unitary representations π of G with a K-fixed unit vector u via

$$\phi(g) = \langle u, \pi(g)u \rangle \tag{1.19}$$

(the *Gelfand-Naimark-Segal* construction). These form the *spherical dual*, Λ, which will be of central importance to the results in Chapter 2.

Spherical functions have the following properties:

1. ϕ is uniformly continuous on G;
2. $\phi(x)\phi(y) = \int_K \phi(xky)dk$, for $x, y \in G$;
3. $\phi(1) = 1$;
4. ϕ is positive definite on G;
5. $f * \phi$ is proportional to ϕ for all $f \in C_c(K\backslash G/K)$.

When the space is compact the spherical functions form a countable set – in the case of the $d-$sphere this the set of ultraspherical polynomials of index $\lambda = (d-1)/2$. In turn, the addition formula for the Gegenbauer polynomials reflects (and is most easily proved via) the action of the group $SO(d)$ on the sphere $\mathbb{S}^d = SO(d+1)/SO(d)$ as a coset space.

The mathematics here is dominated by orthogonal polynomials as a direct consequence of the geometry of the sphere: it arises as a quotient of compact

Lie groups; harmonic analysis on these rests on the Peter-Weyl theorem, for which one has the Schur orthogonality relations.

1.2.4 The Kazhdan property

Before moving on to more general Gaussian processes, let us briefly comment on the Brownian motion in general symmetric spaces. In this case, when $M = G/K$ is a symmetric space, the geometrical property of being Brownian has an algebraic interpretation. Kazhdan defined a locally compact group to have Property (T), now called the *Kazhdan property*, if the unit representation is isolated in the space of unitary representations. Such *Kazhdan groups* have been much studied; see the book by Bekka, de la Harpe and Valette [BHV08] for a very thorough survey of the development of this theory. In the rank-one case, the spherical dual can be identified with a set $\Lambda \subset \mathbb{R}$, where if M is compact Λ is a discrete set. If M is Euclidean, or is real or complex hyperbolic space, Λ can be identified with $[0, \infty)$ so, M is Brownian but is not Kazhdan (0 corresponding to the unit representation). On the other hand if M is quaternionic hyperbolic space, or the octonion (Cayley) projective plane, $\Lambda = \{0\} \cup [\lambda_0, \infty)$, where $\lambda_0 > 0$, so here M is Kazhdan but not Brownian (see Kostant [Kos69]; cf. [FH74]).

Chapter 2

Geo-temporal Covariances

2.1 Gaussian Processes on the Sphere

We now turn away from Brownian motion to more general Gaussian processes on $\mathbb{S}^d$ and $\mathbb{S}^d \times \mathbb{R}$. Characterising isotropic Gaussian randomness on these spaces boils down to characterising the general form of *positive-definite* functions on these spaces. For the sphere this result is classical, and due to Schoenberg [Sch42]. In the geo-temporal case the result is much more recent, dating to Berg and Porcu in 2017 [BP17]. In this chapter we survey the background of positive definite functions on spheres, and proceed to find an alternative formulation of Berg and Porcu's result with a much simpler proof: it is simply a consquence of the Bochner-Godement theorem applied with some care to the product space $\mathbb{S}^d \times \mathbb{R}$.

Take a Gaussian process X *on (defined on*, or *parametrised by) I*: $X = \{X_t, t \in I\}$, on a probability space $(\Omega, \mathcal{F}, P)$ taking values $X_t \in \mathbb{R}$ (or a more general metric space). In the most general setting, no structure is needed on the index set I (see e.g. [AT07, §1.2]). All we need is a mean function μ on I (we centre the process so $\mu = 0$ below unless otherwise stated), and a *covariance function* c on $I \times I$,

$$c(s,t) := \mathrm{cov}(X_s, X_t),$$

which is *positive definite* ('non-negative definite'), or of *positive type*:

$$\sum_{i,j=1}^{n} c(t_i,t_j)u_i u_j \geq 0, \qquad n \in \mathbb{N}, t_i \in I, u_i \in \mathbb{R}. \tag{2.1}$$

Unless otherwise stated throughout this thesis we assume our covariance functions are continuous and *isotropic* – the covariance function $c(s,t)$ depends on the distance $d(s,t)$ (with the appropriate metric). We now define the classes of continuous isotropic positive definite functions $\mathcal{P}(I)$:

Definition 2.1. *The class of continuous functions $f : I \times I \to \mathbb{R}$ such that f is positive definite is denoted $\mathcal{P}(I)$.*

We abuse notation and redefine $\mathcal{P}(\mathbb{S}^d)$ as the class of the functions $f \in C[-1,1]$ such that $C(x,y) = f(\cos d(x,y))$ is positive definite on $\mathbb{S}^d$, and similarly $\mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ is the classes of continuous functions $f : [-1,1] \times \mathbb{R} \to \mathbb{R}$ such that $f(\cos d(x,y), t)$ is positive definite on $\mathbb{S}^d \times \mathbb{R}$. To streamline the exposition we refer to such functions f as 'positive definite on the sphere(-cross-line)', but the reader should keep in mind that it is actually f composed with cos which is truly positive definite on the sphere, and similarly for the sphere-cross-line.

2.2 The Bochner-Godement theorem

In its modern formulation, this very useful result is as follows:

Theorem 2.1 (Bochner-Godement [Wol07, Th. 9.3.4])**.** *The general isotropic positive definite function f on a symmetric space M is given (to within scale c) by a mixture of positive definite spherical functions ϕ_λ over the spherical dual Λ by a probability measure μ:*

$$f(x) = c \int_\Lambda \phi_\lambda(x) \mu(d\lambda), \tag{2.2}$$

or, fixing x and suppressing it in the notation,

$$f = c \int_\Lambda \phi(\lambda) \mu(d\lambda). \tag{2.3}$$

For background and details, see [Wol07, Ch. 9], [FH74, Th. 3.1] – alternatively, [AB76] provides a concise and readable summary of what we need here. The crucial insight is that the Bochner-Godement theorem, stated as here, hugely simplifies the proofs of Schoenberg's theorem and its descendants.

The case when the symmetric space M is a product $M = M_1 \times M_2$ of symmetric spaces M_i is important below. The spherical dual Λ of M is then accordingly the cartesian product of the Λ_i:

$$M = M_1 \times M_2; \qquad \Lambda = \Lambda_1 \times \Lambda_2. \tag{2.4}$$

Then (2.3) becomes, in the obvious notation,

$$\psi = c \int_\Lambda \phi(\lambda) d\mu(\lambda) = c \int_{\Lambda_1 \times \Lambda_2} \phi_1(\lambda_1)\phi_2(\lambda_2) d\mu(\lambda_1, \lambda_2). \tag{2.5}$$

The first of the five results is the prototype:

Theorem 2.2 (Bochner [Boc33])**.** *The general stationary positive definite function $f : \mathbb{R} \to \mathbb{R}$ is a multiple of a characteristic function (Fourier-Stieltjes transform of a probability measure F):*

$$f(t) = c \int_{\mathbb{R}} e^{itx} dF(x). \tag{2.6}$$

This is the Euclidean case $\mathbb{R} = G\backslash K$, $G = \mathbb{R}, K = \{e\}$. In the case where the covariance function $c(t,s)$ is not just stationary but *radial*, i.e. $c(t,s) = f(||t-s||)$ where $|| \cdot ||$ is Euclidean distance, the radialisation of the Fourier transform in Bochner's theorem simplifies to the following early result of Schoenberg's.

Theorem 2.3 (Schoenberg [Sch38])**.** *The general radial positive definite* *function $f : \mathbb{R}^+ \to \mathbb{R}$ has integral representation*

$$f(t) = \int_0^\infty \Gamma\left(\frac{d-1}{2}\right)\left(\frac{2}{rt}\right)^{\frac{d-1}{2}} J_{(d-1)/2}(rt)\, dF(r), \tag{2.7}$$

where F is a probability measure on $\mathbb{R}^+$ and $J_{(d-1)/2}$ is a Bessel function of the first kind (cf. Bochner and Chandrasekharan [BC49, §II.y]).

These modified Bessel functions arise whenever one radialises a characteristic function, as in Kingman's random walks with spherical symmetry [Kin63].

Now we turn to the manifolds of interest in (global) geo-statistical applications – the sphere $\mathbb{S}^d$ and related spaces. Here the spherical dual Λ consists of families of normalised ultraspherical (also called Gegenbauer) polynomials [AB76, §3]. Before normalisation the Gegenbauer polynomials (C_n^λ) are defined using the generating function

$$(1 - 2xr + r^2)^{-\lambda} = \sum_{n=0}^{\infty} C_n^\lambda(x) r^n, \quad |r| < 1, x \in \mathbb{C}, \tag{2.8}$$

or, equivalently,

$$\frac{1 - r^2}{(1 - 2xr + r^2)^{\lambda+1}} = \sum_{n=0}^{\infty} C_n^\lambda(x)\lambda^{-1}(n + \lambda) r^n, \quad |r| < 1, x \in \mathbb{C}. \tag{2.9}$$

Comparing this to the familiar generating function for the Chebychev polynomials

$$\frac{1 - r^2}{1 - 2xr + r^2} = T_0(x) + 2\sum_{n=1}^{\infty} T_n(x) t^n \tag{2.10}$$

the case $\lambda = 0$ follows from taking the limit:

$$\lim_{\lambda \to 0} \frac{n+\lambda}{\lambda} C_n^\lambda(x) = \begin{cases} 1 & n = 0, \\ 2T_n(x) & n = 1, 2, \ldots \end{cases} \tag{2.11}$$

Thus, the family $\{C_n^0(x)/C_n^0(1) : n = 0, 1, \ldots\}$ is precisely the family of Chebychev polynomials $T_n(x) = \cos(n \arccos x)$.

The ultraspherical polynomials form an important subclass of the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ when $\alpha = \beta$:

$$C_n^\lambda(x) = \frac{(2\lambda)_n}{(\lambda + 1/2)_n} P_n^{(\lambda-1/2,\lambda-1/2)}(x), \tag{2.12}$$

where $(a)_n := a(a+1)(a+2)...(a+n-1) = \Gamma(a+n)/\Gamma(a)$ is the Pochammer function. For background on Jacobi polynomials see e.g. [AAR99, Ch. 6].

We normalise the ultraspherical polynomials so they take value 1 at $x = 1$:

$$W_n^\lambda(x) := \frac{C_n^\lambda(x)}{C_n^\lambda(1)}. \tag{2.13}$$

Then the families $\{W_n^\lambda : \lambda > -1/2\}$ are *orthogonal polynomials* on $[-1, 1]$ with respect to the probability measure

$$G_\lambda(dx) := \frac{\Gamma(\lambda + 1)}{\sqrt{\pi}\Gamma(\lambda + 1/2)} \cdot (1 - x^2)^{\lambda - 1/2} dx, \tag{2.14}$$

i.e.

$$\int_{-1}^{1} W_n^\lambda(x) W_m^\lambda(x) G_\lambda(dx) = \frac{\delta_{m,n}}{\omega_n^\lambda}, \tag{2.15}$$

where

$$\omega_n^\lambda := \frac{(n+\lambda)}{\lambda} \cdot \frac{\Gamma(n+2\lambda)}{\Gamma(2\lambda)}. \tag{2.16}$$

Note that we can rewrite $G_\lambda(dx)$ as

$$G_\lambda(dx) = \frac{\sigma_{d-1}}{\sigma_d} \cdot (1-x^2)^{\lambda-1/2} dx, \tag{2.17}$$

where σ_d is the surface area of the $d-$sphere $\mathbb{S}^d$ and setting $\lambda = (d-1)/2$

$$\sigma_d = \omega(\mathbb{S}^d) = \frac{2\pi^{(d+1)/2}}{\Gamma((d+1)/2)} = \frac{2\pi^{\lambda+1}}{\Gamma(\lambda+1)}. \tag{2.18}$$

The ultraspherical polynomials are closely related to spherical harmonics. Spherical harmonics of degree n are the restriction to the sphere $\mathbb{S}^d$ of real-valued harmonic homogeneous polynomials in $\mathbb{R}^{d+1}$ of degree n. A polynomial is *harmonic* if it is a polynomial solution of the Laplace equation

$$\sum_{i=1}^{d+1} \frac{\partial^2 u}{\partial x_i^2} = 0, \tag{2.19}$$

and *homogeneous* if it defines a homogeneous function (i.e. if all the non-zero terms are of the same degree). For example, let $d+1=2$, and n a positive integer. Then the homogeneous polynomial $u(x,y) = x^2 - y^2$ satisfies (2.19). Restricting to the circle $\mathbb{S}^1$ we have $x = \cos\theta$, $y = \sin\theta$ and $u(\theta) = \cos 2\theta$, the Chebychev polynomial $T_2(x)$.

Allowing the zero-function to be a harmonic homogeneous polynomial of degree 0 the space $\mathcal{H}_n(d)$ of spherical harmonics of degree n on the sphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ form a finite-dimensional vector space of dimension (see [AAR99, 9.3])

$$\begin{aligned}\dim(\mathcal{H}_n(d)) &= (2n+d-1)\frac{(n+d-2)!}{n!(d-1)!} \qquad (2.20)\\ &= (2n+d-1)\frac{\Gamma(n+d-1)}{n!\Gamma(d)}\\ &= (2n+2\lambda)\frac{\Gamma(n+2\lambda)}{n!\Gamma(2\lambda+1)}\\ &= \frac{1}{n!}\frac{n+\lambda}{\lambda}\frac{\Gamma(n+2\lambda)}{\Gamma(2\lambda)} = \frac{\omega_n^\lambda}{n!}. \qquad (2.21)\end{aligned}$$

Then the orthogonality relation (2.15) can be written as

$$\int_{-1}^{1} W_n^\lambda(x)W_m^\lambda(x)(1-x^2)^{\lambda-1/2}dx = \frac{\sigma_d}{\sigma_{d-1}}\frac{1}{n!\dim(\mathcal{H}_n(d))}\delta_{m,n}. \qquad (2.22)$$

With the ultraspherical polynomials thus introduced we can formulate *Schoenberg's theorem*, the classical result characterising positive-definiteness on the sphere.

Theorem 2.4 (Bochner, Schoenberg [Sch37]). *For $\mathbb{S}^d$, the general isotropic pd function $f \in \mathcal{P}(\mathbb{S}^d)$ is (to within scale c) a mixture of (normalised) ultraspherical polynomials $W_n^\lambda(x)$ with $\lambda = (d-1)/2$ – that is, $f \in \mathcal{P}(\mathbb{S}^d)$ if and* *only if it has the Fourier-Gegenbauer expansion:*

$$f(x) = c\sum_{n=0}^{\infty} a_n W_n^{\lambda}(x),\ a_n \geq 0,\ \sum_{n=0}^{\infty} a_n = 1, \tag{2.23}$$

where $x = \cos d(x,y)$*,* $x, y \in \mathbb{S}^d$*. The sequence* (a_n) *is a mixing law whose components are the* Schoenberg coefficients *of* f*.*

Proof. This is immediate from the Bochner-Godement theorem. For the $d-$dimensional sphere, the spherical dual $\Lambda(\mathbb{S}^d)$ is given by the (discrete) family of ultraspherical polynomials of order $\lambda = (d-1)/2$, $\{W_n^{\lambda} \circ \cos : n = 0, 1, \ldots\}$ [Bin73; AB76]. Since in this case the spherical dual is discrete, the integral of the Bochner-Godement theorem is reduced to a sum over n, with the spectral measure μ now a probability sequence (a_n).

In particular, writing $x = \cos\theta$, the general positive definite function on $\mathbb{S}^d$ is given by

$$f(x) = \sum_{n=0}^{\infty} W_n^{\lambda}(x) a_n$$

for some sequence (a_n) such that $a_n \geq 0$ for all n, and $\sum_{n=0}^{\infty} a_n = 1$. $\square$

The classes $\mathcal{P}(\mathbb{S}^d)$ nest:

$$\mathcal{P}(\mathbb{S}^1) \supset \mathcal{P}(\mathbb{S}^2) \supset \ldots \supset \mathcal{P}(\mathbb{S}^{\infty}). \tag{2.24}$$

To see this, note that, by definition, if $f \in \mathcal{P}(\mathbb{S}^{d+1})$ the $n \times n$ matrix $A = (A_{ij})$, $A_{ij} = f(d(x_i, x_j))$, is positive definite for all finite collections

$x_1, \ldots, x_n \in \mathbb{S}^{d+1}$. Since $\mathbb{S}^1 \subset \mathbb{S}^2 \subset \ldots$, this holds for all $x_1, \ldots, x_n \in \mathbb{S}^d \subset \mathbb{S}^{d+1}$, thus $f \in \mathcal{P}(\mathbb{S}^d)$.

We define $\mathcal{P}(\mathbb{S}^\infty) := \cap \mathcal{P}(\mathbb{S}^d)$ to be the set of functions positive definite on *all spheres* and, equivalently, positive definite on the infinite-dimensional *Hilbert sphere*. In this limiting case the ultraspherical polynomials W_n^∞ are simply the monomials x^n, and $f \in \mathcal{P}(\mathbb{S}^\infty)$ if and only if

$$f(x) = c \sum_{n=0}^{\infty} a_n x^n, \tag{2.25}$$

with $a_n \geq 0$ and $\sum_{n=0}^{\infty} a_n = 1$ [Sch42].

The Fourier-Gegenbauer form of (2.23) allows one to find the sequence of Schoenberg coefficients (a_n) from a given function f via the Fourier-Gegenbauer transform:

$$a_n = \int_{-1}^{1} f(x) W_n^\lambda(x)(1-x)^{\lambda - 1/2} dx. \tag{2.26}$$

These coefficients encode entirely the covariance function, and thus the features of the underlying Gaussian field. This will be exploited in depth in later chapters, particularly with regard to *path properties*, where the continuity and differentiability properties of the field are defined by the rate of decay of the sequence (a_n).

A result generalising the Bochner-Schoenberg theorem to any compact symmetric space of rank one is due to Askey and Bingham in the 1970s.

Theorem 2.5 (Askey and Bingham [AB76]). *For a compact symmetric space* *of rank one, M, the general isotropic positive-definite function $f \in \mathcal{P}(M)$ is (to within scale factor c) a mixture*

$$f(x) = c \sum_{n=0}^{\infty} a_n \phi_n(x),\ a_n \geq 0,\ \sum_{n=0}^{\infty} a_n = 1 \tag{2.27}$$

of the (countably many) spherical functions ϕ_n.

Theorem 2.5 completes the picture for non-product spaces. Two recent results are on products $M = M_1 \times M_2$ of symmetric spaces. (2.5), combined with the Bochner-Schoenberg theorem, makes the first, on products of spheres, immediate from the Bochner-Godement theorem [GMP16].

Theorem 2.6 (Guella, Menegatto and Peron [GMP16])**.** *The general isotropic pd function on $\mathbb{S}^{d_1} \times \mathbb{S}^{d_2}$ is*

$$c \sum_{m,n=0}^{\infty} a_{mn} W_m^{\lambda_1}(x_1) W_n^{\lambda_2}(x_2),\ \ a_{mn} \geq 0,\ \ \sum_{m,n=0}^{\infty} a_{mn} = 1, \tag{2.28}$$

where $x_i = \cos d(x_i, y_i)$, $x_i, y_i \in \mathbb{S}^{d_i}$ and $\lambda_i = \frac{1}{2}(d_i - 1)$.

The second, the case of 'sphere cross line' $M = \mathbb{S}^d \times \mathbb{R}$, was answered by Berg and Porcu [BP17] in 2017.

Theorem 2.7a (Berg and Porcu [BP17]) *The class $\mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ of isotropic stationary sphere-cross-line covariances f coincides with the class of mixtures of products of Gegenbauer polynomials $W_n^{\lambda}(x)$ and positive definite functions*

$a_n(t) \in \mathcal{P}(\mathbb{R})$:

$$f(x,t) = c\sum_{n=0}^{\infty} a_n(t) W_n^{\lambda}(x), \;\; a_n(0) \geq 0, \;\; \sum_{n=0}^{\infty} a_n(0) = 1, \tag{2.29}$$

with x as in Theorem 2.4. The coefficient functions are given by the Fourier-Gegenbauer transforms

$$a_n(t) = \int_{-1}^{1} f(x,t) W_n^{\lambda}(x)(1-x^2)^{\lambda-1/2} dx. \tag{2.30}$$

We offer here a new, alternative formulation of the Berg-Porcu theorem which shows it to be, as Schoenberg's theorem is, a direct corollary of the Bochner-Godement theorem. Both formulations are illuminating: the Berg-Porcu formulation of Theorem 2.7a is convenient in practice (indeed, we shall use it in Chapter 3). Given a function $f \in \mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ one can use (2.30) to find the Fourier-Gegenbauer coefficients, although we note that as in the purely spatial case the integral in (2.30) rarely has a closed form.

On the other hand, the formulation below displays more explicitly the underlying analysis: Theorem 2.7a combines the mixing law (a_n) from Schoenberg's theorem with the temporally varying component to obtain a sequence of positive-definite functions, whereas here we separate them with the main user-benefit being a quick-and-easy proof, given below.

Theorem 2.7b. *The class of isotropic stationary sphere-cross-line covariances* $\mathcal{P}(\mathbb{S}^d \times \mathbb{R}) := \{f \in C([-1,1]\times\mathbb{R}) : f(\cos\theta, t)$ is positive definite for $\theta \in [0,\pi], t \in \mathbb{R}\}$ *coincides with the class of mixtures of products of Gegenbauer* *polynomials $W_n^\lambda(x)$ and characteristic functions $\phi_n(t)$ on the line:*

$$f(x,t) = c\sum_{n=0}^{\infty} a_n W_n^\lambda(x)\phi_n(t), \qquad a_n \geq 0,\ \sum_{n=0}^{\infty} a_n = 1. \tag{2.31}$$

Proof. This is the special case of the Bochner-Godement theorem for $M = \mathbb{S}^d \times \mathbb{R}$, with spherical dual $\Lambda = \Lambda_1 \times \Lambda_2 = \mathbb{N}_0 \times \mathbb{R}$.

In the product formulation of Bochner-Godement (2.5) take $\lambda = (\lambda_1, \lambda_2) = (n, x)$ with $n \in \mathbb{N}_0$, $x \in \mathbb{R}$ – then $\phi_1(n) = W_n^\lambda$ by Schoenberg's theorem, and $\phi_2(x) = e^{ix\cdot} = \{t \mapsto e^{itx}, t \in \mathbb{R}\}$, by Bochner's theorem. Thus

$$\psi = c\int_\Lambda \phi(\lambda)\mu(d\lambda) = \int_{\mathbb{N}_0\times\mathbb{R}} W_n^\lambda e^{ix\cdot}\mu(d(n,x)). \tag{2.32}$$

To perform the integration, we can use the language of either probability theory or measure theory, depending on the reader's taste. For the first: since μ is a probability measure, the integral is an expectation of the random variable $\lambda = (\lambda_1, \lambda_2)$ with law μ. So

$$f = c\mathbb{E}[\phi(\lambda)] = c\mathbb{E}\left[\phi_1(\lambda_1)\phi_2(\lambda_2)\right] \tag{2.33}$$

$$= c\mathbb{E}\left[\mathbb{E}\left[\phi_1(\lambda_1)\phi_2(\lambda_2)|\lambda_1\right]\right], \tag{2.34}$$

by the Tower property of conditional expectation [Wil91, §9.7] (also sometimes called the Conditional Mean Formula [Wil01, p. 390] or the chain rule [Kal97, p. 105]). Condition on $\lambda_1 = n$. The first factor in the conditional expectation is W_n^λ, as in the Bochner-Schoenberg theorem. If the conditional law of $\lambda_2|(\lambda_1 = n)$ is μ_n, with characteristic function $\phi_n(t)$, then the second factor is the integral of the character e^{ixt} with respect to μ_n, namely $\phi_n(t)$. The remaining expectation is thus of $W_n^\lambda \phi_n(t)$ over the law $a = (a_n)$ of λ_1, giving (2.31) as required.

For the second, use disintegration of measures (a generalisation of Fubini's theorem to non-product measures: see e.g. Kallenberg [Kal97, Th. 6.4]), integrating μ on Λ first over the x-variable above for fixed n. This gives a probability measure, μ_n. The remaining integration (of $W_n^\lambda \phi_n$)) is a summation over n with weights a_n, again giving (2.31). $\square$

Note that, since $\phi_n(0) = 1$, the mixing law (a_n) in (2.31) may be determined via the Fourier-Gegenbauer transform

$$a_n = \int_{-1}^{1} f(x,1) W_n^\lambda(x)(1-x^2)^{\lambda-1/2} dx. \tag{2.35}$$

2.3 Chordal vs. Geodesic distance

As we have seen, the positive definite functions $f : \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \to \mathbb{R}$ are the Fourier(-Stieltjes) transforms of (positive, finite) measures, by Bochner's theorem. If, further, $f(x,y)$ is a radial function purely of the (Euclidean) distance $t = ||x-y||$ between its two arguments we can use Theorem 2.3.

A natural question here is whether one can restrict these radial functions from $\mathbb{R}^{d+1}$ to isotropic functions on the sphere $\mathbb{S}^d$, whilst retaining positive definiteness. A recipe due to Yadrenko [Yad83] allows for the construction of a positive definite function $\hat{f}$ on $\mathbb{S}^d$ from a positive definite function f on $\mathbb{R}^{d+1}$ via the transformation

$$\hat{f}(\theta) = f\left(2\sin\frac{\theta}{2}\right), \qquad \theta \in [0,\pi].$$

Although this construction is convenient, one predictably pays a price for ignoring the geometry of the sphere. Gneiting [Gne13] highlighted this approach's severe restrictions – there is a lower bound of $\inf_{t>0} 1/t \sin t \approx -0.21$ on $\hat{f}$ when $d = 2$ (the geostatistical case). Moreover, Gneiting argues that, since $\sin\theta \approx \theta$ when θ is small, this construction is doomed to conflict with spherical geometry. It is preferable, therefore, to construct our positive definite functions on the sphere itself, using *geodesic* (or great circle) distance in preference to the Euclidean distance above.

2.4 Strict Positive Definiteness

Schoenberg's theorem characterises the class $\mathcal{P}(\mathbb{S}^d)$ of isotropic positive-definite functions on the sphere – that is, functions f such that for any $n \in N$ and any collections $u_1, u_2, \ldots u_n \in \mathbb{R}$ and $x_1, x_2, \ldots x_n \in \mathbb{S}^d$

$$\sum_{i,j=1}^{n} f(d(x_i, x_j))u_i u_j \geq 0. \tag{2.36}$$

Note that this is equivalent to the $n \times n$ matrix $A := f(d(x_i, x_j))$ being *non-negative definite.*

Definition 2.2. *If for all sets of n distinct points on $\mathbb{S}^d$ the matrices $A = f(d(x_i, x_j))$ are positive definite (in the sense of matrices, i.e. non-singular, of full rank, and so on) we say the function f is* strictly positive definite, *and write $f \in \mathcal{P}^+(\mathbb{S}^d)$.*

It is of interest to find necessary and sufficient conditions for membership of the class $\mathcal{P}^+(\mathbb{S}^d)$. We focus on approaches phrased in terms of conditions on the Schoenberg coefficients a_n.

The simplest condition is a sufficient condition found by Xu and Cheney.

Theorem 2.7 (Xu and Cheney [XC92])**.** *Let $f \in \mathcal{P}(\mathbb{S}^d)$ with Schoenberg coefficients $(a_k)_{k\in\mathbb{N}}$. Let $x_1, x_2, \ldots x_n$ be distinct points on $\mathbb{S}^d$. If $a_k > 0$ for $0 \le k < n$, then the matrix $A_{i,j} = f(d(x_i, x_j))$ is positive definite.*

Proof. We show Xu and Cheney's proof in the case $d = 1$, as a model for the general case. In fact, Xu and Cheney's result is stronger when $d = 1$, with positivity being enforced only for the first $\lfloor \frac{n}{2} \rfloor$ coefficients [XC92, Th. 1]. Let $A \in \mathbb{R}^{n\times n}$ be the matrix given by $f(d(x_i, x_j))$. Assume, for a contradiction, that there exists some non-zero $u = (u_1, u_2, \ldots u_n)$, $u_i \in \mathbb{R}$ such that $u^T A u = 0$.

Recall that, when $d = 1$ $\lambda = 0$ and so the ultraspherical polynomials in the Schoenberg expansion of $f \in \mathcal{P}(\mathbb{S}^1)$ are the Chebyshev polynomials $W_n^0(x) = \cos(n \arccos x)$. Writing $x_i = (\cos\theta_i, \sin\theta_i)$, so the arc length $d(x_i, x_j) =$

$\theta_i - \theta_j$ we have

$$A_{ij} = f(d(x_i, x_j)) = \sum_{k=0}^{\infty} a_k \cos(kd(x_i, x_j)) \tag{2.37}$$

$$= \sum_{k=0}^{\infty} a_k \cos(k\theta_i - \theta_j) \tag{2.38}$$

$$= \sum_{k=0}^{\infty} a_k \cos k\theta_i \cos k\theta_j + \sum_{k=0}^{\infty} a_k \sin k\theta_i \sin k\theta_j, \tag{2.39}$$

yielding $A = B + C$, where

$$B_{ij} = \sum_{k=0}^{\infty} a_k \cos k\theta_i \cos k\theta_j, \quad C_{ij} = \sum_{k=0}^{\infty} a_k \sin k\theta_i \sin k\theta_j. \tag{2.40}$$

These are sums of non-negative definite matrices of rank 1, and are thus non-negative definite themselves. Thus, by assumption, $u^T Bu = 0$ and $u^T Cu = 0$ and

$$0 = \sum_{i,j=1}^{n} u_i u_j B_{ij} = \sum_{i,j=1}^{n} u_i u_j \sum_{k=0}^{\infty} a_k \cos k\theta_i \cos k\theta_j \tag{2.41}$$

$$= \sum_{k=0}^{\infty} a_k \sum_{i,j=1}^{n} u_i u_j \cos k\theta_i \cos k\theta_j. \tag{2.42}$$

Now let $r = \lfloor \frac{n}{2} \rfloor$. Then, since $a_0, a_1, \ldots a_r$ are strictly positive, $\sum_{i=1}^{n} u_i \cos k\theta_i = 0$ for $k = 0, 1, \ldots, r$. Clearly, then, the operator $\mathcal{L}(f) := \sum_{i=1}^{n} u_i f(\theta_i)$ annihilates $\cos k\theta$, $k = 0, 1, \ldots r$. The same argument applied to the matrix C shows that $\mathcal{L}$ also annihilates $\sin k\theta$, $k = 0, 1, \ldots, r$.

Trigonometric polynomials have the interpolation property – there exists a trigonometric polynomial p of degree at most r satisfying $p(\theta_i) = u_i$, $i =$

$1, \dots, n$ (if $n \le 2r+1$). Thus,

$$0 = \mathcal{L}(p) = \sum_{i=1}^{n} u_i p(\theta_i) = \sum_{i=1}^{n} u_i^2 > 0, \tag{2.43}$$

the desired contradiction. So, $u_i = 0$ for all i. The proof of the general case $d > 1$ follows along similar lines and we refer the reader to Xu and Cheney [XC92] for full details. □

This yields the following corollary, a sufficient condition for a positive definite function f to be strictly positive definite.

Corollary 2.1 (Xu and Cheney [XC92])**.** *If $f(x) = \sum a_n W_n^\lambda(x) \in \mathcal{P}(\mathbb{S}^d)$ and $a_n > 0$ for all n, then $f \in \mathcal{P}^+(\mathbb{S}^d)$.*

Weaker sufficient conditions, and necessary ones, have been found for membership of $\mathcal{P}^+(\mathbb{S}^d)$ since 1992. A well-known result is the following, by Chen, Menegatto and Sun [CMS03].

Theorem 2.8 (Chen, Menegatto and Sun)**.** *Let $d \ge 2$. A function $f \in \mathcal{P}(\mathbb{S}^d)$ is strictly positive definite, $f \in \mathcal{P}^+(\mathbb{S}^d)$, if and only if $a_n > 0$ for infinitely many odd n, and infinitely many even n.*

We omit the proof, which relies on careful selection of a polar origin, and then representing points $x_i \in \mathbb{S}^d$ in polar form using points $x_i' \in \mathbb{S}^{d-1}$, allowing the use of a representation theorem for the ultraspherical polynomials. Note that this technique fails when $d = 1$ – indeed, the sufficient condition in Theorem

2.8 fails in this case and the question of characterising fully the members of $\mathcal{P}(\mathbb{S}^1)$ remains open.

Questions of membership of $\mathcal{P}^+(\mathbb{S}^d \times \mathbb{R})$ are more complicated still, with even a sufficient condition still elusive – Xu and Cheney's approach fails (observe that in the geo-temporal case the double sum in (2.42) cannot be split, as it will contain a factor of $\phi_k(t_i - t_j)$, and it does not seem possible to adapt Chen, Menegatto and Sun's technique to include a temporal factor.

2.5 Complements

The theory here is rich and historied, and we collect here some remarks expanding upon more general mathematical and historical aspects of the above, for interest.

2.5.1 The Gelfand-Naimark-Segal (GNS) construction

For G a topological group, K a closed subgroup, π a unitary representation of G in a Hilbert space H, say $u \in H$ is *cyclic* if $\{\pi(g)u : g \in G\}$ generates a dense subspace of H. The Gelfand-Naimark-Segal (GNS) construction [GN43; Seg47], which we touched upon slightly when discussing spherical functions in Chapter 1, says that for $\phi \in \mathcal{P}(K\backslash G/K)$, there exists a unitary representation (π_ϕ, H_ϕ) of G and a K-invariant cyclic unit vector u with

$$\phi(g) = (u, \pi(g)u) \qquad g \in G, \tag{2.44}$$

and (π_ϕ, H_ϕ, u) is unique up to isomorphism.

The GNS construction permeates the modern treatment of spherical functions on symmetric spaces; see e.g. [Hel62, Ch. X.4]. It is easily seen that any ϕ as in (2.44) is positive definite; these are the spherical functions that arise in, say, symmetric spaces of compact type.

In (2.44) above, ϕ is an extreme point in the convex set $\mathcal{P}(K\backslash G/K)$ if and only if the unitary representation (π_ϕ, H_ϕ) is irreducible [Far08, Prop 1.4]. This shows very clearly the role of convexity and the Krein-Milman theorem [Rud91, Th. 3.21] – which states that a compact convex subset of a locally convex vector space is the closed convex hull of its extremal points – here. In particular, the classes $\mathcal{P}(\mathbb{S}^d)$ are convex and the Bochner-Godement expansion (2.3) may be intepreted as Choquet representations.

2.5.2 Gelfand pairs

For G a locally compact group with compact subgroup K, (G, K) is called a *Gelfand pair* if the convolution algebra of K-biinvariant compactly supported continuous measures on G is commutative; equivalently, if for any locally convex irreducible representation π of G, its space of K-invariant vectors is at most one-dimensional. For symmetric spaces $M = (G, K)$, (G, K) is a Gelfand pair; see e.g. [Hel62, Th. 4.1].

2.5.3 Pólya criteria

Verifying (strict) positive definiteness is a highly non-trivial task – checking whether infinitely many of the Schoenberg coefficients of a given function are positive is almost always a completely intractable problem. Similarly in Euclidean space, checking whether a given function is positive definite using Bochner's theorem is often impossible. To circumvent this, we need easy-to-check sufficient conditions for positive definiteness. The most ubiquitous of these is Pólya's criterion [Gne01; Pól18; Pól49; Luk70].

Theorem 2.9 (Pólya's criterion [Pól18; Luk70])**.** *If f is even, continuous and convex on $[0, \infty)$, and $\lim_{t\to\infty} f(t) = 0$, then $f \in \mathcal{P}(\mathbb{R})$.*

A Pólya condition for spheres was found recently by Beatson [BCX14], relying on a conjecture later proved by Xu [Xu18].

Theorem 2.10 (Pólya's criterion for spheres [BCX14])**.** *Let $d \geq 2$. If f is such that*

1. $f \in C^\lambda[-1, 1]$;
2. $\operatorname{supp} f \subset [0, 1)$;
3. *the derivative from the right of* $(f \circ \cos)^{(\lambda+1)}(0)$ *exists;*

4. $(-1)^\lambda(f \circ \cos)^{(\lambda)}$ *is convex,*

then $f \in \mathcal{P}(\mathbb{S}^d)$*. If, further,* $f \circ \cos$ *restricted to* $(0, \pi)$ *does not reduce to a linear polynomial, then* $f \in \mathcal{P}^+(\mathbb{S}^d)$*.*

The (rather unnatural looking) condition (*4*) is a technical requirement in [BCX14, Th. 1.3], ensuring tat the Fourier-Gegenbauer coefficients in their construction of members of $\mathcal{P}^+(\mathbb{S}^d)$ are non-negative.

2.5.4 Anisotropy

We have assumed (and will continue to assume throughout this thesis) the spatial component of the process is *isotropic* – the covariance between two points x and y depends solely on the distance between them. This convenient assumption facilitates all the theory exposited above and is key to much of what follows in subsequent chapters – and is pervasive in spatial statistics: see [Ma17] for a discussion of isotropy's impact on covariance modelling over $\mathbb{R}^d$.

2.5.5 Lagrangian frameworks

In fluid dynamics one has a choice of one's frame of reference – one may either sit on the bank and watch the water pass (Eulerian), or one may follow the body of water in a boat (Lagrangian). In geostatistical cases, when the symmetry of an isotropic model is an inappropriate assumption, a Lagrangian covariance structure may be more suitable, or, as suggested by Gneiting [Gne02b], a convex combination of a symmetric covariance and a Lagrangian one.

The basic idea of a Lagrangian covariance is to take a stationary spatial random field with covariance $C_s(\mathbf{h})$ and follow it along as it moves with (random) velocity $\mathbf{V}$. The covariance of this field then has the form

$$C(\mathbf{h}, t) = \mathbb{E}_{\mathbf{V}} \left(C_s(\mathbf{h} - t\mathbf{V}) \right).$$

Gneiting et al. [GGG07] discuss possible choices for $\mathbf{V}$, in particular noting that prevailing winds may be modelled by the simplest case $\mathbf{V} = \mathbf{v}$, a constant. They also raise the interesting possibility of a dynamic velocity $\mathbf{V}(t)$, which would yield nonstationary covariances in a very natural way. This framework has recently been extended to the $\mathbb{S}^d \times \mathbb{R}$ case by Alegria and Porcu [AP17], who also discuss the *dimple effect* (where the field is more strongly correlated in time than in space) for transport models.

2.5.6 Multivariate applications

The results discussed above extend to the multivariate setting. In this case, one needs to consider *cross-covariance* functions. Recent work [Ale+19] extends the traditional Euclidean framework to the $\mathbb{S}^d \times \mathbb{R}$ setting discussed here, offering some parametric families of matrix-valued covariances on sphere cross line.

Chapter 3

Dimension Walks

For a given function, checking positive definiteness can be a cumbersome task, even with the assistance of the powerful theorems of Chapter 2. This motivates the question: given a positive definite function, what operations can we apply to it which preserve positive definiteness? These positivity-preserving operators are useful in the development of new families of (strictly) positive definite functions for use as covariance models in applications.

Two such positivity-preserving operators are the *montée* and *descente* operators introduced by Mathéron [Mat70]. These operators map positive definite functions to new functions *positive definite with respect to a space of different dimension.* Wendland [Wen95] coined the phrase 'walks through dimensions' to describe this dimension-hopping property. Somewhat counterintuitively, the montée ("up") operator walks *down* dimensions, whilst descente increases the dimension. In fact, the names refer to the effect of the operators on *smoothness* – montée, by integrating, increases the smoothness of the positive definite function it is applied to, whereas descente does the opposite.

For functions $f \in \mathcal{P}(\mathbb{R}^d)$ the montée, denoted I, and descente, D, operators are defined by

$$(If)(t) := \frac{\int_t^\infty uf(u)\,\mathrm{d}u}{\int_0^\infty uf(u)\,\mathrm{d}u}, \qquad t \geq 0 \tag{3.1}$$

and

$$(Df)(t) := \begin{cases} 1, & t = 0 \\ f'(t)/tf''(0), & t > 0 \end{cases} \tag{3.2}$$

Note that, for f absolutely continuous, $(DI)f = (ID)f = f$. The term 'walks through dimensions' becomes clear with the following pair of theorems [Mat70; Wen95]:

Theorem 3.1. *Let $f \in \mathcal{P}(\mathbb{R}^d), d \geq 3$. If $uf(u)$ is integrable over $[0, \infty)$, then $If \in \mathcal{P}(\mathbb{R}^{d-2})$.*

Theorem 3.2. *Let $f \in \mathcal{P}(\mathbb{R}^d)$. If f'' exists, then $Df \in \mathcal{P}(\mathbb{R}^{d+2})$.*

The utility of the above theorems was demonstrated by Wendland [Wen95] in his construction of a compactly supported correlation model in Euclidean space, later adapted by Gneiting [Gne99; Gne02a]. Compactly supported correlation functions are of particular use in meteorological forecast verification, where error correlations vanish beyond a certain length-scale (typically a few thousand kilometres) [GC99]. Moreover, using a compactly supported model reduces computational cost, allowing for sparse matrix techniques to be utilised.

The primary obstacle to the development of valid compactly supported positive definite functions is the difficulty of parametrising their smoothness properties. As we shall see in Chapter 4, the behaviour of the covariance function near the origin is determined by the smoothness of the underlying random field – this is usually phrased in terms of mean-square differentiability: a random field with covariance function c is k-times mean-square differentiable if and only if $c^{(2k)}(0)$ exists. Families of covariance models where this can be parametrised are the most appealing to practitioners, and dimension-hopping operators are key here.

Consider the truncated power function

$$p_{\nu,0}(t) := \begin{cases} (1-t)^{\nu}, & t \in [0,1], \\ 0, & t \geq 1. \end{cases} \tag{3.3}$$

This is in $\mathcal{P}(\mathbb{R}^d)$ if and only if $\nu \geq (d+1)/2$. Wendland showed, using Theorems 3.1 and 3.2, that repeated application of the montée operator

$$p_{\nu,k}(t) = I^k p_{\nu,0}(t), \qquad k \in \mathbb{N} \tag{3.4}$$

gives covariance functions in $\mathcal{P}(\mathbb{R}^d)$ if and only if $\nu \geq (d+1)/2+k$. Since $p_{\nu,k}$ is $2k$-times differentiable at zero [Wen95], this gives a parametric family of compactly covariance functions from which one can choose any (even) degree of differentiability at the origin.

3.1 Montée and descente on $\mathbb{S}^d$

The meteorological applications of montée and descente in Euclidean space motivate the development of analogous operators for functions in $\mathcal{P}(\mathbb{S}^d)$, to facilitate the construction of locally supported families of covariance models with nice parametrisation properties. These analogues were found by Beatson and zu Castell [BC17] and applied to generalise the Wendland construction above. Spherical montée and descente operators have also been introduced for the Hilbert sphere $\mathbb{S}^\infty$ [Zie14] and the complex sphere [MPP17].

Definition 3.1 ([BC17])**.** *For $f \in \mathcal{P}(\mathbb{S}^d)$ absolutely continuous and integrable on $[-1,1]$, define the the montée operator by*

$$(\mathcal{I}f)(x) := \int_{-1}^{x} f(t)\,\mathrm{d}t, \tag{3.5}$$

and the descente operator by

$$(\mathcal{D}f)(x) := f'(x). \tag{3.6}$$

Clearly for f absolutely continuous $(\mathcal{D}\mathcal{I})f = (\mathcal{I}\mathcal{D})f = f$ (for almost all $x \in [-1,1]$).

As one would expect from Schoenberg's theorem, the properties of the montée and descente operators on spheres are inherited from the differentiability and integrability properties of Gegenbauer polynomials. From Szegő [Sze39, §4.7.14]:

$$\mathcal{D}W_n^\lambda(x) = \begin{cases} 2\lambda W_{n-1}^{\lambda+1}(x), & \lambda > 0 \\ 2W_{n-1}^1(x), & \lambda = 0. \end{cases} \tag{3.7}$$

Following [BC17], define a new index μ by

$$\mu_\lambda := \begin{cases} l, & \lambda > 0 \\ 1, & \lambda = 0, \end{cases}$$

allowing (3.7) to be more concisely written as

$$\mathcal{D}W_n^\lambda(x) = 2\mu_\lambda W_{n-1}^{\lambda+1}(x), \quad \lambda \geq 0. \tag{3.8}$$

Then, writing (3.8) in terms of the montée operator $\mathcal{I}$,

$$\mathcal{I}W_{n-1}^{\lambda+1}(x) = \frac{1}{2\mu_\lambda}\left(W_n^\lambda(x) - W_n^\lambda(-1)\right), \quad \lambda \geq 0. \tag{3.9}$$

These then give the following dimension walks on spheres:

Theorem 3.3 ([BC17])**.** *Let $d \in \mathbb{N}$. If $f \in \mathcal{P}(\mathbb{S}^{d+2})$, then there exists a constant C such that $C + \mathcal{I}f \in \mathcal{P}(\mathbb{S}^d)$.*

Theorem 3.4 ([BC17])**.** *Let $d \in \mathbb{N}$. If $f \in \mathcal{P}(\mathbb{S}^d)$ has a continuous derivative f', then $\mathcal{D}f \in \mathcal{P}(\mathbb{S}^{d+2})$.*

Theorem 3.3 shows that the montée operator for spheres walks down dimensions, producing (as one expects from integration) a smoother function. Theorem 3.4 gives a rougher function on a space of higher dimension, unless the derivative fails to exist or be continuous. As in the Euclidean case, the montée operator can be used to construct families of increasingly smooth (strictly) positive definite functions [BC17, §3].

3.2 Montée and descente on $\mathbb{S}^d \times \mathbb{R}$

The possibility of walks on dimensions for members of $\mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ was raised by Porcu et al. in a recent survey [PAF18]:

> "*The literature on walks through dimensions is related to operators that allow one, for a given positive definite function on the n-dimensional sphere, to obtain new classes of positive definite functions on n'-dimensional spheres, with $n \neq n'$. The application of such operators has consequences on the differentiability at the origin of the involved functions. Walks on spheres have been proposed by Beatson et al. [BC17; BC16], Ziegel [Zie14] and Massa et al. [MPP17]. This last work extends the previous work to the case of complex spheres. It would be timely to obtain walks through dimensions for the members of the classes $\mathcal{P}(\mathbb{S}^n \times \mathbb{R})$, for n a positive integer.*"

Here we show the results of Beatson and zu Castell [BC17; BC16] for one-and two- step operators on the sphere extend to the spatio-temporal setting, answering the question posed above. The results in this section, and the subsequent section detailing one-step walks, have been published in Statistics and Probability Letters [BS19].

For the spatio-temporal montée and descente operators, defined below, we abuse notation and reuse $\mathcal{I}$ and $\mathcal{D}$ from the purely spatial setting.

Definition 3.2. *For $f \in \mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ integrable define the spatio-temporal montée operator by*

$$(\mathcal{I}f)(x,t) := \int_{-1}^{x} f(u,t)du. \tag{3.10}$$

For $f \in \mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ with continuous partial derivative in its first variable, the spatio-temporal descente operator $\mathcal{D}$ is defined as

$$(\mathcal{D}f)(x,t) := \frac{\partial}{\partial x} f(x,t). \tag{3.11}$$

By Corollary 3.9 in [BP17], $\partial^n f(x,t)/\partial x^n$ exists and is continuous when $n \leq \lambda$. Thus we can only guarantee existence and continuity of $\mathcal{D}f(x,t)$ for $\lambda \geq 1$, i.e. $d \geq 3$. For the interesting cases $d = 1$ and $d = 2$ we need to impose additional (but not particularly cumbersome) differentiability assumptions, namely that $\partial f(x,t)/\partial x$ exists and is continuous, to ensure the uniform convergence of its Schoenberg expansion.

Theorem 3.5. *For $d \in \mathbb{N}$ and $f \in \mathcal{P}(\mathbb{S}^{d+2} \times \mathbb{R})$, there exists a constant C such that $C + \mathcal{I}f \in \mathcal{P}(\mathbb{S}^d \times R)$.*

Proof. The proof proceeds as in [BC17], with the appropriate changes. Since $f(x,t) \in \mathcal{P}(\mathbb{S}^{d+2} \times \mathbb{R})$, it has the uniformly convergent Schoenberg expansion

$$f(x,t) = \sum_{n=0}^{\infty} a_n(t) W_n^{\lambda+1}(x),$$

with $a_n(t)$ as in the first formulation of the Berg-Porcu theorem (i.e. combining the mixing law and characteristic functions into one sequence of functions in $\mathcal{P}(\mathbb{R})$ with $\sum a_n(0) = 1$ – this eases the exposition here somewhat). Integrating term by term and using (3.10) and (3.9):

$$\begin{aligned}
\mathcal{I}f(x,t) &= \int_{-1}^{x} \sum_{n=0}^{\infty} a_n(t) W_n^{\lambda+1}(u) du \\
&= \sum_{n=0}^{\infty} a_n(t) \int_{-1}^{x} W_n^{\lambda+1}(u) du \\
&= \sum_{n=0}^{\infty} a_n(t)\, \mathcal{I} W_n^{\lambda+1} \\
&= \sum_{n=0}^{\infty} a_n(t) \frac{1}{2\mu_\lambda} \left(W_{n+1}^{\lambda}(x) - W_{n+1}^{\lambda}(-1) \right) \\
&= \sum_{n=1}^{\infty} \frac{a_{n-1}(t)}{2\mu_\lambda} W_n^{\lambda}(x) - \sum_{n=1}^{\infty} \frac{a_{n-1}(t)}{2\mu_\lambda} W_n^{\lambda}(-1).
\end{aligned}$$

So,

$$(\mathcal{I}f)(x,t) = \sum_{n=0}^{\infty} b_n(t) W_n^{\lambda}(x), \tag{3.12}$$

where

$$b_n(t) = \begin{cases} \frac{a_{n-1}(t)}{2\mu_\lambda}, & n = 1, 2, 3, \ldots \\ -\sum_{i=0}^{\infty} \frac{a_{i-1}(t)}{2\mu_\lambda} W_i^{\lambda}(-1), & n = 0. \end{cases} \tag{3.13}$$

When $n \geq 1$ the $b_n(t)$ are clearly positive definite on $\mathbb{R}$, and $2\mu_\lambda \sum_{n=1}^\infty b_n(0) = \sum_{n=0}^\infty a_n(0) < \infty$. The only term which needs more comment is $b_0(t)$. This is constant in x, so if there exists a constant C bounding b_0 we find $C + \mathcal{I}f \in \mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ as required.

When $\lambda = 0$,

$$\begin{aligned}|b_0(t)| &= \big|\frac{1}{2}\sum_{n=1}^\infty a_{n-1}(t) W_n^0(-1)\big| \\ &= \big|\frac{1}{2}\sum_{n=1}^\infty a_{n-1}(t)\big| \\ &\leq \frac{1}{2}\sum_{n=1}^\infty |a_{n-t}(t)| \\ &\leq \frac{1}{2}\sum_{n=1}^\infty a_{n-1}(0) < \infty.\end{aligned}$$

The final inequality follows from the positive-definiteness of the a_n (so $|a_n(t)| \leq a_n(0)$ for all $t \in \mathbb{R}$). So we may take $C = \frac{1}{2}\sum_{n=1}^\infty a_{n-1}(0)$.

Similarly for $\lambda > 0$, $W_n^\lambda(-x) = (-1)^n W_n^\lambda(x)$ [Sze39, §4.1.3], so

$$\begin{aligned}|b_0(t)| = \big|\sum_{m=1}^\infty \frac{(-1)^{m+1} a_{m-1}(t)}{2\mu_\lambda} W_m^\lambda(1)\big| &\leq \sum_{m=1}^\infty \frac{|a_{m-1}(t)|}{2\mu_\lambda} \\ \leq \sum_{m=1}^\infty \frac{|a_{m-1}(0)|}{2\mu_\lambda}.\end{aligned}$$

By [Sze39, Th. 9.1.3] the sum is (C,k)-summable for $k > \lambda + 1/2$ and so, since all the terms in the sum are non-negative, convergent [Har49, Th. 64].

So

$$|b_0(t)| \leq \sum_{m=1}^{\infty} \frac{a_{m-1}(0)}{2\mu_\lambda} W_m^\lambda(1) = \sum_{m=1}^{\infty} \frac{a_{m-1}(0)}{2\mu_\lambda} =: C < \infty,$$

giving the required bound for $b_0(t)$. □

Walks up dimensions, using the descente operator, are simpler still, but we now require the function f to have a continuous first partial derivative.

Theorem 3.6. *Let $d \in \mathbb{N}$ and $f : [-1,1] \times \mathbb{R} \to \mathbb{R}$ continuously differentiable in its first argument. If $f \in \mathcal{P}(\mathbb{S}^d \times \mathbb{R})$, then $\mathcal{D}f(x,t) = \partial/\partial x f(x,t) \in \mathcal{P}(\mathbb{S}^{d+2} \times \mathbb{R})$.*

Proof. We can differentiate termwise, as the series of derivatives is uniformly convergent:

$$\begin{aligned}
\frac{\partial}{\partial x} f(x,t) &= \frac{\partial}{\partial x}\left(a_0(t) + \sum_{n=1}^{\infty} a_n(t) W_n^\lambda(x)\right) \\
&= \sum_{n=1}^{\infty} a_n(t) \frac{\partial}{\partial x} W_n^\lambda(x) \\
&= \sum_{n=1}^{\infty} a_n(t) 2\mu_\lambda W_{n-1}^{\lambda+1}(x) && \text{(by (3.7))} \\
&= \sum_{n=0}^{\infty} 2\mu_\lambda a_{n+1}(t) W_n^{\lambda+1}(x).
\end{aligned}$$

Alternatively, one may extend the proof of [BC17, Lemma 2.4, Th. 2.3] along the same lines as in the proof above. □

We also have the following result for the infinite-dimensional case (the Hilbert sphere), analogous to that in [TZ17].

Theorem 3.7. *The class $\mathcal{P}(\mathbb{S}^\infty \times \mathbb{R})$ of geotemporal covariances on the Hilbert sphere cross time is closed under $\mathcal{I}$ (up to an additive constant). It is not closed under descente: if $f \in \mathcal{P}(\mathbb{S}^\infty \times \mathbb{R})$, then $\mathcal{D}f \in \mathcal{P}(\mathbb{S}^\infty \times \mathbb{R})$ only if $\sum n a_n(t) < \infty$.*

Proof. This is straightforward: functions $f \in \mathcal{P}(\mathbb{S}^\infty \times \mathbb{R})$ have the Schoenberg expansion

$$f(x,t) = \sum_{n=0}^{\infty} a_n(t) x^n.$$

Integrating with respect to x:

$$\mathcal{I}f(x,t) = \int_{-1}^{x} f(u,t)\, du = \sum_{n=0}^{\infty} \frac{a_n(t)}{n+1} \left(x^{n+1} - (-1)^{n+1}\right) = \sum_{n=0}^{\infty} b_n(t) x^n, \tag{3.14}$$

where

$$b_n(t) = \begin{cases} a_{n-1}(t)/n, & n = 1, 2, \dots \\ \sum_{i=0}^{\infty} \left(a_i(t)(-1)^i\right)/(i+1), & n = 0. \end{cases}$$

Since $b_0(t)$ is bounded by $\sum a_n(0)/(n+1) < \infty$ there is a suitable constant C such that $C + \mathcal{I}f \in \mathcal{P}(\mathbb{S}^\infty \times \mathbb{R})$.

Turning to the *descente* operator: by [BPP18, Th. 1.1] the functions $a_n(t)$ are given by

$$a_n(t) = \tilde{a}_n(x,t)\big|_{x=0}, \quad \tilde{a}_n(x,t) = \frac{1}{n!} \frac{\partial^n f(x,t)}{\partial x^n}.$$

So, as above, consider

$$\frac{\partial f(x,t)}{\partial x} = \sum_{n=0}^{\infty} b_n(t) x^n,$$

where

$$b_n(t) = \tilde{b}_n(x,t)\big|_{x=0}, \qquad \tilde{b}_n(t) = \frac{1}{n!}\frac{\partial^{n+1} f(x,t)}{\partial x^{n+1}}.$$

So,

$$\tilde{b}_n(x,t) = (n+1)\tilde{a}_{n+1}(x,t),$$

and

$$b_n(t) = (n+1)a_{n+1}(t).$$

Thus if $\sum n a_n(0)$ does not converge then $\mathcal{D}f(x,t) \notin \mathcal{P}(\mathbb{S}^\infty \times \mathbb{R})$, and the result follows taking the contrapositive. □

Beatson and zu Castell showed further than the montée and descente operators preserve *strict* positive definiteness on the sphere using the Chen, Menegatto and Sun criterion [CMS03] cf. §2.4: $f \in \mathcal{P}^+(\mathbb{S}^d)$, $d \geq 2$, if and only if infinitely many of the Schoenberg coefficients of even index, and infinitely many of odd index, are non-zero. The criterion is necessary but not sufficient for $d = 1$.

As mentioned, characterising functions in $\mathcal{P}^+(\mathbb{S}^d \times \mathbb{R})$ remains an open problem, but we expect a similar criterion to hold and thus montée and descente for $\mathbb{S}^d \times \mathbb{R}$ to preserve strict positive definiteness, also.

3.3 Walking in steps of one

The alert reader will observe that Theorems 3.3, 3.4, 3.5 and 3.7 (and the corresponding theorems in Euclidean space) hop dimension in steps of two. The reason for this is (3.8): the parameter λ of the relevant ultraspherical polynomial changes by an integer, thus alters $d = 2\lambda + 1$ by two. In the Euclidean setting, walks of one step are achieved using fractional integration. A similar approach can be employed in the spherical case, as demonstrated by Beatson and zu Castell [BC16], using weighted *Riemann-Liouville operators*

Definition 3.3 ([BC16])**.** *Define, for $f \in \mathcal{P}(\mathbb{S}^d)$ and $\lambda \geq 0$,*

$$^{\alpha}I_{+}^{\lambda}f(x) := (1+x)^{-\lambda+\alpha}\int_{-1}^{x}(x-u)^{\alpha-1}(1+u)^{\lambda}f(u)\,du, \tag{3.15}$$

$$^{\alpha}I_{-}^{\lambda}f(x) := (1-x)^{-\lambda+\alpha}\int_{x}^{1}(u-x)^{\alpha-1}(1-u)^{\lambda}f(u)\,du, \tag{3.16}$$

$$^{\alpha}\mathcal{I}_{\pm}^{\lambda} := {}^{\alpha}I_{+}^{\lambda} \pm {}^{\alpha}I_{-}^{\lambda}. \tag{3.17}$$

If f is absolutely continuous then we may also define

$$^{\alpha}D_{+}^{\lambda}f(x) := (1+x)\frac{\partial}{\partial x}\left((1+x)^{-\lambda}\int_{-1}^{x}(x-u)^{\alpha-1}(1+u)^{\lambda-\alpha}f(u)\,du\right), \tag{3.18}$$

$$^{\alpha}D_{-}^{\lambda}f(x) := (1-x)\frac{\partial}{\partial x}\left((1-x)^{-\lambda}\int_{x}^{1}(u-x)^{\alpha-1}(1-u)^{\lambda-\alpha}f(u)\,du\right), \tag{3.19}$$

$$^{\alpha}\mathcal{D}_{\pm}^{\lambda} := {}^{\alpha}D_{+}^{\lambda} \pm {}^{\alpha}D_{-}^{\lambda}. \tag{3.20}$$

Unlike the classical Riemann-Liouville operators the operators ${}^{\alpha}\mathcal{I}^{\lambda}_{\pm}$ and ${}^{\alpha}\mathcal{D}^{\lambda}_{\pm}$ do not have the semi-group property: that is, integration (or differentiation) of orders α_1 and α_2 in succession does not yield the result of integration (or differentiation) $\alpha_1 + \alpha_2$.

For one-step walks on spheres we are interested in the case where $\alpha = 1/2$. For the rest of this chapter we abuse notation somewhat, and abbreviate ${}^{\frac{1}{2}}I^{\lambda}_{\cdot}$ to $I^{\lambda}_{\cdot}$ and, similarly, ${}^{\frac{1}{2}}D^{\lambda}_{\cdot}$ to $D^{\lambda}_{\cdot}$ for ease of reading.

With the one-step operators thus defined, Beatson and zu Castell provided the following one-step analogues to montée and descente for spheres:

Theorem 3.8 ([BC16]). *If $d \in \mathbb{N}$, $f \in \mathcal{P}(\mathbb{S}^{d+1})$, then $\mathcal{I}^{\lambda}_{\pm} f \in \mathcal{P}(\mathbb{S}^{d})$.*

Theorem 3.9 ([BC16]). *If $d \in \mathbb{N}$, $f \in \mathcal{P}(\mathbb{S}^{d})$, then $\mathcal{D}^{\lambda}_{\pm} f \in \mathcal{P}(\mathbb{S}^{d+1})$*

The crux of the proof of these results, and our extension to $\mathcal{P}(\mathbb{S}^{d} \times \mathbb{R})$ given below, are the following two lemmas concerning the action of the one-step operators on ultraspherical polynomials.

Lemma 3.1 ([BC16]). *For $\lambda > 0$, $n \in \mathbb{N}$, $x \in [-1, 1]$,*

$$\begin{aligned} \mathcal{I}^{\lambda}_{+} W^{\lambda+1/2}_{n}(x) &= \frac{\sqrt{\pi}\Gamma(\lambda)}{\Gamma(\lambda+1/2)} \frac{n+2\lambda}{n+\lambda+1/2} W^{\lambda}_{n}(x), \\ \mathcal{I}^{\lambda}_{-} W^{\lambda+1/2}_{n}(x, 0) &= \frac{\sqrt{\pi}\Gamma(\lambda)}{\Gamma(\lambda+1/2)} \frac{n+2\lambda}{n+\lambda+1/2} W^{\lambda}_{n+1}(x). \end{aligned}$$

When $\lambda = 0$ simply take the limit $\lambda \to 0^+$ and obtain

$$\mathcal{I}^{0}_{+} P_n(x) = \frac{2}{n+1/2} T_n(x), \qquad \mathcal{I}^{0}_{-} P_n(x) = \frac{2}{n+1/2} T_{n+1}(x).$$

where $P_n(x)$ and T_n the Legendre and Chebychev polynomials, respectively.

Lemma 3.2 ([BC16])**.** *For $\lambda > 0$, $n \in \mathbb{N}$, $x \in [-1, 1]$,*

$$\mathcal{D}_+^\lambda W_n^\lambda(x) = \frac{\sqrt{\pi}\Gamma(\lambda + 1/2)}{\Gamma(\lambda)} \frac{n + 2\lambda}{n + \lambda} W^{\lambda+1/2} l_{n-1}(x),$$
$$\mathcal{D}_-^\lambda W_n^\lambda(x) = \frac{\sqrt{\pi}\Gamma(\lambda + 1/2)}{\Gamma(\lambda)} \frac{2n}{n + \lambda} W_n^{\lambda+1/2}(x).$$

When $\lambda = 0$ again take the limit $\lambda \to 0^+$ and obtain

$$\mathcal{D}_+^0 T_n(x) = n\pi P_{n-1}(x), \qquad \mathcal{D}_-^0 T_n(x) = n\pi P_n(x).$$

3.4 One-step walks on $\mathbb{S}^d \times \mathbb{R}$

We now turn to verifying that Beatson and zu Castell's one-step dimension walks above extend to the spatio-temporal setting. This is slightly less trivial an extension that in the two-step case, but as before most of the work has been done for us in determining the effect of the operators on the ultraspherical polynomials.

Definition 3.4. *Define, for $f(x,t) \in \mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ and $\lambda \geq 0$,*

$$I_+^\lambda f(x,t) := (1 + x)^{\lambda+1/2} \int_{-1}^{x} (x - u)^{-1/2}(1 + u)^\lambda f(u,t)du, \tag{3.21}$$

$$I_-^\lambda f(x,t) := (1 - x)^{\lambda+1/2} \int_{x}^{1} (u - x)^{-1/2}(1 - u)^\lambda f(u,t)du, \tag{3.22}$$

$$\mathcal{I}_\pm^\lambda := I_+^\lambda \pm I_-^\lambda. \tag{3.23}$$

If f is absolutely continuous, then we may also define

$$D_+^\lambda f(x,t) := (1+x)\frac{\partial}{\partial x}\left((1+x)^{-\lambda}\int_{-1}^{x}(x-u)^{-1/2}(1+u)^{\lambda-1/2}f(u,t)du\right), \tag{3.24}$$

$$D_-^\lambda f(x,t) := (1-x)\frac{\partial}{\partial x}\left((1-x)^{-\lambda}\int_{x}^{1}(u-x)^{-1/2}(1-u)^{\lambda-1/2}f(u,t)du\right), \tag{3.25}$$

$$\mathcal{D}_\pm^\lambda := D_+^\lambda \pm D_-^\lambda. \tag{3.26}$$

The results below extend those of Section 3.2 to our context.

Theorem 3.10. *If $d \in \mathbb{N}$, $f \in \mathcal{P}(\mathbb{S}^{d+1} \times \mathbb{R})$, then $\mathcal{I}_\pm^\lambda f \in \mathcal{P}(\mathbb{S}^d \times \mathbb{R})$.*

Proof. As above, the proof requires only small modifications to Beatson and zu Castell's arguments [BC16]. In both cases, the proof itself closely replicates that of Theorem 3.5: after justifying exchanging the integration with summation we simply apply the relevant result about ultraspherical polynomials, then conclude.

Let $f \in \mathcal{P}(\mathbb{S}^{d+1} \times \mathbb{R})$. Then f has the uniformly convergent expansion

$$f(x,t) = \sum_{n=0}^{\infty} a_n(t) W_n^{\lambda+1/2}(x). \tag{3.27}$$

Thus,

$$I_I^\lambda f(x,t) = (1+x)^{-\lambda}\int_{-1}^{x}(x-u)^{-1/2}(1+u)^{\lambda+1/2}f(u,t)\,du$$

$$= (1+x)^{-\lambda} \int_{-1}^{x} (x-u)^{-1/2}(1+u)^{\lambda+1/2} \sum_{n=0}^{\infty} a_n(t) W_n^{\lambda+1/2}(u)\, du.$$

To justify integrating termwise we need to establish boundedness of the operator I_+^λ from $C[-1,1]$ to $C[-1,1]$. Using the beta integral

$$\int_{-1}^{x} (x-u)^{-1/2}(1+u)^{\nu}\, du = (1+x)^{\nu+1/2} \frac{\Gamma(1/2)\Gamma(\nu+1)}{\Gamma(\nu+3/2)} \tag{3.28}$$

we obtain

$$\sup_{x\in[-1,1]} |I_+^\lambda f(x,t)| = \sup_{x\in[-1,1]} |(1+x)^{-\lambda} \int_{-1}^{x} (x-u)^{-1/2}(1+u)^{\lambda+1/2} f(u,t)\, du| \tag{3.29}$$

$$\leq \sup_{x\in[-1,1]} |1+x|^{-\lambda} \int_{-1}^{x} |(x-u)^{-1/2}(1+u)^{\lambda+1/2}||f(u,t)|\, du \tag{3.30}$$

$$\leq \frac{2\Gamma(1/2)\Gamma(\lambda+3/2)}{\Gamma(\lambda+2)} \sup_{x\in[-1,1]} |f(x,t)|. \tag{3.31}$$

So $||I_+^\lambda f|| \leq C||f||$. By the same reasoning, I_-^λ is a bounded also, and thus $\mathcal{I}_\pm^\lambda$ is. Positivity of the operators is clear from their definitions.

Now we may exchange the sum and the integral and use Lemma 3.1 to obtain:

$$\mathcal{I}_+^\lambda f(x,t) = \sum_{n=0}^{\infty} a_n(t)(1+x)^{-\lambda} \int_{-1}^{x} (x-u)^{-1/2}(1+u)^{\lambda+1/2} W_n^{\lambda+1/2}(u)\, du$$

$$= \sum_{n=0}^{\infty} a_n(t) \mathcal{I}_+^\lambda W_n^{\lambda+1/2}(x)$$

$$= \sum_{n=0}^{\infty} b_n(t) W_n^{\lambda}(x),$$

where

$$b_n(t) = \frac{\sqrt{\pi}\Gamma(\lambda)}{\Gamma(\lambda+1/2)} \frac{n+2\lambda}{n+\lambda+1/2} a_n(t),$$

and

$$\begin{aligned}
\mathcal{I}_-^{\lambda} f(x,t) &= \sum_{n=0}^{\infty} a_n(t)(1-x)^{-\lambda} \int_{-1}^{x} (u-x)^{-1/2}(1-u)^{\lambda+1/2} W_n^{\lambda+1/2}(u)\, du \\
&= \sum_{n=0}^{\infty} a_n(t) I_-^{\lambda}\, W_n^{\lambda+1/2}(x) \\
&= \sum_{n=0}^{\infty} c_n(t) W_n^{\lambda}(x),
\end{aligned}$$

where

$$c_n(t) = \frac{\sqrt{\pi}\Gamma(\lambda)}{\Gamma(\lambda+1/2)} \frac{n+1}{n+\lambda+1/2} a_{n+1}(t)$$

Clearly, both $\sum b_n(0)$ and $\sum c_n(0)$ converge, thus $\mathcal{I}_+^{\lambda} f$ and $\mathcal{I}_-^{\lambda} f$ are members of $\mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ as required. $\square$

Note that, unlike in Theorem 3.5, there is no additional constant C needed here.

To prove the corresponding statement about the operators $\mathcal{D}_{\pm}^{\lambda}$ we need the following lemma, the proof of which is identical to the corresponding result in Beatson and zu Castell [BC16, Th. 2.8].

Lemma 3.3. *Let $f \in \mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ have Schoenberg expansion $\sum a_n(t) W_n^{\lambda}(x)$.*

If both $\mathcal{D}^\lambda_\pm f \in C([-1,1] \times \mathbb{R})$, then the resulting Schoenberg expansions

$$\mathcal{D}^\lambda_+ f(x,t) = \sum_{n=0}^{\infty} b_n(t) W_n^{\lambda+1/2}(x), \qquad \mathcal{D}^\lambda_- f(x,t) = \sum_{n=0}^{\infty} c_n(t) W_n^{\lambda+1/2}(x)$$

have coefficients

$$b_n(t) = \frac{\Gamma(\lambda+1/2)\sqrt{\pi}}{\Gamma(\lambda)} \frac{2(n+2\lambda+1)}{n+\lambda+1} a_{n+1}(t), \tag{3.32}$$

$$c_n(t) = \frac{\Gamma(\lambda+1/2)\sqrt{\pi}}{\Gamma(\lambda)} \frac{2n}{n+\lambda} a_n(t). \tag{3.33}$$

Theorem 3.11. *If $d \in \mathbb{N}$, $f \in \mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ and $\mathcal{D}^\lambda_\pm f$ are continuous, then $\mathcal{D}^\lambda_\pm f \in \mathcal{P}(\mathbb{S}^{d+1} \times \mathbb{R})$.*

Proof. We adapt the proof of the corresponding result in [BC16]. On the surface matters appear slightly more complicated than in the purely spatial setting, as here we are dealing with series of positive definite functions, rather than of non-negative reals. But, we can reduce this case to the previous one.

Let $f \in \mathcal{P}(\mathbb{S}^d \times \mathbb{R})$ with Schoenberg functions (a_n). The continuity assumption guarantees Abel summability of $D^\lambda_\pm f(x,t)$ for any $(x,t) \in [-1,1] \times \mathbb{R}$, whilst Lemma 3.3 gives positive-definiteness of the resulting coefficient functions $b_n(t)$. Thus, $b_n(0) \geq 0$ and $\sum b_n(0) W_n^{\lambda+1/2}(1) = \sum b_n(0)$ is summable (since Abel summability of non-negative real numbers implies summability [Har49, Th. 91]).

To conclude, note that, for any n, $|W_n^\lambda(x)| \leq W_n^\lambda(1)$ for all $x \in [-1,1]$, and $|b_n(t)| \leq b_n(0)$ for all $t \in \mathbb{R}$. The Weierstrass M-test then demonstrates uniform convergence. So, $\mathcal{D}_{\pm}^{\lambda} f \in \mathcal{P}(\mathbb{S}^{d+1} \times \mathbb{R})$, as required. $\square$

3.5 Walking with the semi-group property

Beatson and zu Castell's operators, extended above, give walks in steps of one but lack the semi-group property – two steps of length one do not necessarily yield the same result as a single step of length two.

Here we propose an alternative method of walking through dimensions allowing, theoretically at least, for walks of arbitrary length with the semi-group property.

The idea is to use some classical integral representations of ultraspherical polynomials to find a dimension-hopping operator $\mathcal{H}_{\nu}^{\lambda}$ which maps $\mathcal{P}(\mathbb{S}^{d(\nu)})$ to $\mathcal{P}(\mathbb{S}^{d(\lambda)})$ such that the composition $\mathcal{H}_{\lambda}^{\eta} \circ \mathcal{H}_{\nu}^{\lambda} : \mathcal{P}(\mathbb{S}^{d(\nu)}) \to \mathcal{P}(\mathbb{S}^{d(\eta)})$ results in the same member of $\mathcal{P}(\mathbb{S}^{d(\eta)})$ as direct application of $\mathcal{H}_{\nu}^{\eta}$ – that is, that the operator $\mathcal{H}$ has the semi-group property.

Theorem 3.12 ([Bin72a])**.** *If $0 < \nu < \lambda$, $x \in [-1, 1]$, there exists a measure $M_{\nu}^{\lambda}(x; dy)$ on $[-1, 1]$ such that*

$$W_n^{\lambda}(x) = \int_{-1}^{1} W_n^{\nu}(y)\, M_{\nu}^{\lambda}(x; dy). \tag{3.34}$$

Moreover, when $\lambda \neq \nu$ the measure M_ν^λ is absolutely continuous with density

$$M_\nu^\lambda(x;dy) = G_\nu(dy)\sum_{m=0}^{\infty}\omega_m^\nu W_m^\lambda(x)W_m^\nu(y) \tag{3.35}$$

where, recall,

$$\begin{aligned} G_\nu(dy) &= \frac{\Gamma(\nu+1)}{\sqrt{\pi}\Gamma(\nu+1/2)}(1-y^2)^{\nu-1/2}dy, \\ \omega_m^\nu &= \frac{n+\nu}{\nu}\frac{\Gamma(n+2\nu)}{n!\Gamma(2\nu)}. \end{aligned}$$

As a result we have the following:

Theorem 3.13. *Let $1 < d_1 < d_2 < \infty$, $\lambda_1 = (d_1-1)/2$, $\lambda_2 = (d_2-1)/2$. If $f \in \mathcal{P}(\mathbb{S}^{d_1})$, then*

$$(\mathcal{H}_{\lambda_1}^{\lambda_2} f)(x) := \int_{-1}^{1} f(y)M_{\lambda_1}^{\lambda_2}(x;dy) \in \mathcal{P}(\mathbb{S}^{d_2}), \tag{3.36}$$

i.e. the operator $\mathcal{H}_{\lambda_1}^{\lambda_2}$ gives a walk on dimensions from $\mathcal{P}(d_1)$ to $\mathcal{P}(d_2)$.

Proof. If $f \in \mathcal{P}(\mathbb{S}_1^d)$, then

$$f(x) = \sum_{n=0}^{\infty} a_n W_n^{\lambda_1}(x), \quad \sum a_n = 1, a_n \geq 0.$$

So,

$$\int_{-1}^{1} f(y)M_{\lambda_1}^{\lambda_2}(x;dy) = \int_{-1}^{1}\sum_{n=0}^{\infty} a_n W_n^{\lambda_1}(y)M_{\lambda_1}^{\lambda_2}(x;dy)$$

$$= \sum_{n=0}^{\infty} \int_{-1}^{1} a_n W_n^{\lambda_1}(y) M_{\lambda_1}^{\lambda_2}(x; dy),$$

where interchanging the sum and integral is justified by the uniform convergence of the Schoenberg expansion. Then,

$$\begin{aligned} \sum_{n=0}^{\infty} \int_{-1}^{1} a_n W_n^{\lambda_1}(y) M_{\lambda_1}^{\lambda_2}(x; dy) &= \sum_{n=0}^{\infty} a_n \int_{-1}^{1} W_n^{\lambda_1}(y) M_{\lambda_1}^{\lambda_2}(x; dy) \\ &= \sum_{n=0}^{\infty} a_n W_n^{\lambda_2}(x) \in \mathcal{P}(\mathbb{S}^{d_2}), \end{aligned}$$

since the a_n remain non-negative and summable to 1. □

Theorem 3.7. is not, in-and-of-itself, terribly useful. To be able to pass directly from $f \in \mathcal{P}(\mathbb{S}^{d_1})$ to $f' \in \mathcal{P}(\mathbb{S}^{d_2})$ we need to be able to integrate against $M_{\lambda_1}^{\lambda_2}(x, dy)$ immediately – that is, we need

$$\sum_{m=0}^{\infty} \omega_m^{\nu} W_m^{\lambda}(x) W_m^{\nu}(y).$$

There is no explicit formula for this series. In an attempt to circumvent this we have found an alternative form, as a *double integral*, which is more tractable (at least numerically, as we lose the oscillatory polynomial behaviour of the sum).

Moreover, Theorem 3.14 below is interesting in its own right, completing the integral representations in [Bin72a] by showing the dependence on the higher index, λ, in a more convenient and structurally revealing way.

The Poisson kernel for the Jacobi polynomials reduces in the ultraspherical case to the generating function, cf. [Bin72a, (2.1)]

$$\sum_{n=0}^{\infty} \omega_n^{\nu} r^n W_n^{\nu}(x) = (1-r^2)/(1-2rx+x^2)^{\nu+1}, \quad r \in (-1,1). \tag{3.37}$$

Note that this is not the usual generating function for the ultraspherical polynomials [Sze39, §4.7.23].

Askey and Fitch [AF69] showed that for $x, y \in [-1,1]$, $r \in (-1,1)$, $0 \leq \nu < \lambda \leq \infty$, the series

$$\sum_{n=0}^{\infty} \omega_n^{\nu} r^n W_n^{\lambda}(x) W_n^{\nu}(y) \tag{3.38}$$

converges to a non-negative sum-function, which leads to a corresponding probability measure $M_\nu^\lambda(x)$ satisfying

$$W_n^{\lambda}(x) = \int_{-1}^{1} W_n^{\nu}(y) M_\nu^\lambda(x; dy), \quad n = 0, 1, 2, \ldots. \tag{3.39}$$

Here (see [Bin72a]) we may take $0 \leq \nu \leq \lambda \leq \infty$, $x \in [-1,1]$. Some cases give Dirac laws: if $x = \pm 1$, $M_\nu^\lambda(\pm 1) = \delta_{\pm 1}$ (as $W_n^\lambda(\pm 1) = (\pm 1)^n$). If $\lambda = \nu$, then $M_\lambda^\lambda(x) = \delta_x$ (as there is no projection to be done); so we may now restrict to $\nu < \lambda$ as before. [Bin72a, Lemma 1] gives the Abel-limit operation explicitly: for $x, y \in (-1,1)$, we may take $r = 1$ here to get

$$m_\nu^\lambda(x; y) := \sum_{n=0}^{\infty} \omega_n^{\nu} W_n^{\lambda}(x) W_n^{\nu}(y) \geq 0 \tag{3.40}$$

a non-negative function in $L_1(G_\nu)$, finite-valued unless $x = y$ and $\nu < \lambda \leq$

$\nu + 1$. It is in fact the *Radon-Nikodym derivative* $dM_\nu^\lambda(x; dy)/dG_\nu(dy)$:

$$M_\nu^\lambda(x; dy) = G_\nu(dy) \cdot m_\nu^\lambda(x; y) = G_\nu(dx) \cdot \sum_{n=0}^{\infty} \omega_n^\nu W_n^\lambda(x) W_n^\nu(y). \tag{3.41}$$

Following [Bin72a], for $\lambda > \nu$ write H_ν^λ for the probability measure of Beta type on $[0, 1]$ given by the *Sonine law*

$$H_\nu^\lambda(dx) := \frac{2\Gamma(\lambda + \frac{1}{2})}{\Gamma(\nu + \frac{1}{2})\Gamma(\lambda - \nu)} \cdot x^{2\nu}(1 - x^2)^{\lambda - \nu - \frac{1}{2}} dx. \tag{3.42}$$

This occurs in Sonine's first finite integral for the Bessel function [Wat62, p. 373]: for

$$\Lambda_\mu(t) := \Gamma(\nu + 1) J_\nu(t)(t/2)^{-\mu}, \tag{3.43}$$

$$\Lambda_{\lambda - \frac{1}{2}}(t) = \int_0^1 \Lambda_{\nu - \frac{1}{2}}(ut) H_\nu^\lambda(du) \tag{3.44}$$

(the drop by a half-integer in parameter here reflects the drop in dimension in $\mathbb{S}^d \subset \mathbb{R}^{d+1}$).

For the product of W_n terms in (3.40), we need Gegenbauer's multiplication theorem for the ultraspherical polynomials [Wat62, p. 369],

$$W_n^\nu(x) W_n^\nu(y) = \int_{-1}^{1} W_n^\nu(xy + \sigma\sqrt{1 - x^2}\sqrt{1 - y^2}) G_{\nu - \frac{1}{2}}(d\sigma). \tag{3.45}$$

To cope with the drop in index (dimension) in (3.40), we need the *Feldheim-*

Vilenkin integral [Bin72a, (2.11)], [AF69],

$$W_n^\lambda(x) = \left[\frac{2\Gamma(\lambda+\frac{1}{2})}{\Gamma(\nu+\frac{1}{2})\Gamma(\lambda-\nu)}\right]\int_0^1 u^{2\nu}(1-u^2)^{\lambda-\nu-1}$$
$$\cdot\,[x^2-x^2u^2+u^2]^{\frac{1}{2}n}W_n^\nu\left(\frac{x}{\sqrt{x^2-x^2u^2+u^2}}\right)du. \tag{3.46}$$

We can then formulate our result.

Theorem 3.14. *For $r \in (-1,1)$, the sum of the Askey-Fitch series (3.40) above is given by the integral (3.47) below:*

$$\int_0^1 H_\nu^\lambda(du)\int_{-1}^1 G_{\nu-\frac{1}{2}}(dv)\,\frac{[1-r^2(x^2-x^2u^2+u^2)]}{I^{\nu+1}}, \tag{3.47}$$

where I is given by

$$I := 1-2r\cdot\frac{xy+uv\sqrt{1-x^2}\sqrt{1-y^2}}{\sqrt{x^2-x^2u^2+u^2}}+\frac{(xy+uv\sqrt{1-x^2}\sqrt{1-y^2})^2}{(x^2-x^2u^2+u^2)}. \tag{3.48}$$

Moreover, this holds also for $r=1$ unless $\nu<\lambda\leq\nu+1$.

Proof. We sum the series by reducing it to the generating function (3.37). There are two steps: reduction of λ to ν by the Feldheim-Vilenkin integral (3.46) and reduction of two W_n terms to one by Gegenbauer's multiplication theorem (3.45).

We follow [Bin72a]. As there, we may substitute for $W_n^\lambda(x)$ from (3.46) into the series (3.40) and integrate term-wise, rewriting (3.40) as

$$\int_0^1 H_\nu^\lambda(du) \sum_{n=0}^\infty \omega_n^\nu (r[x^2 - x^2u^2 + u^2]^{\frac{1}{2}})^n \cdot W_n^\nu(y) W_n^\nu \left(\frac{x}{\sqrt{x^2 - x^2u^2 + u^2}} \right). \tag{3.49}$$

We use Gegenbauer's multiplication theorem (3.45) with

$$r \mapsto r\sqrt{x^2 - x^2u^2 + u^2},$$

and replace the product of W_n^ν factors in the above, at the cost of another integration over $G_{\nu-\frac{1}{2}}(dv)$, by a single W_n^ν term, with argument

$$\frac{xy}{\sqrt{x^2 - x^2u^2 + y^2}} + v\sqrt{1-y^2}.\sqrt{1 - \frac{x^2}{x^2 - x^2u^2 + u^2}} = \frac{xy + uv\sqrt{1-x^2}\sqrt{1-y^2}}{\sqrt{x^2 - x^2u^2 + u^2}}. \tag{3.50}$$

The integrand is now of the form $\sum \omega_n^\nu r^n W_n^\nu(\cdot)$, and the result now follows from (3.37). $\square$

This result allows us to rephrase Theorem 3.13 in a more (numerically) tractable form, by replacing the double sum in the measure $M_{\lambda_1}^{\lambda_2}(\cdot; dy)$ with (3.47). That is, if $f \in \mathcal{P}(\mathbb{S}^{d_1})$ a walk on dimension to d_2 is obtained via

$$\int_{-1}^1 f(y) M_{\lambda_1}^{\lambda_2}(x; dy) \in \mathcal{P}(\mathbb{S}^{d_2}), \tag{3.51}$$

where

$$M_{\lambda_1}^{\lambda_2}(x; dy) = G_{\lambda_1}(dy) \int_0^1 H_{\lambda_1}^{\lambda_2}(du) \int_{-1}^1 G_{\lambda_1 - \frac{1}{2}}(dv)\, \frac{[1 - r^2(x^2 - x^2u^2 + u^2)]}{I^{\lambda_1+1}}, \tag{3.52}$$

where I is given by

$$I := 1 - 2r \cdot \frac{xy + uv\sqrt{1-x^2}\sqrt{1-y^2}}{\sqrt{x^2 - x^2u^2 + u^2}} + \frac{(xy + uv\sqrt{1-x^2}\sqrt{1-y^2})^2}{(x^2 - x^2u^2 + u^2)}. \tag{3.53}$$

Of course, applying this to produce new parametric families via projection is still complicated by the fact we were unable to obtain a closed form for the integral above. Nonetheless, the result completes and complements the work in [AF69; Bin72a] by displaying the dependence on the higher index λ in a structurally revealing way: for simplicity, let $r = 1$ so that

$$I = \left(1 - \frac{xy + uv\sqrt{1-x^2}\sqrt{1-y^2}}{\sqrt{x^2 - x^2u^2 + u^2}}\right)^2, \tag{3.54}$$

and (3.47) is given by

$$\int_0^1 H_\nu^\lambda(du) \int_{-1}^1 G_{\nu - \frac{1}{2}}(dv) \frac{1 - (x^2 - x^2u^2 + u^2)}{I^{\nu+1}}. \tag{3.55}$$

Using the definition of H_ν^λ and the probability measure $G_{\nu+\frac{1}{2}}$ and simplifying, (3.47) becomes

$$\frac{2}{\sqrt{\pi}} \int_0^1 \frac{\Gamma(\lambda + \frac{1}{2})}{\Gamma(\lambda - \nu)} u^{2\nu}(1-u^2)^{\lambda - \nu - \frac{1}{2}} \int_{-1}^1 (1 - v^2)^{\nu - \frac{1}{2}}$$

$$\cdot \left[\frac{1 - (x^2 - x^2u^2 + u^2)}{I^{2(\nu+1)}} \right] dvdu. \tag{3.56}$$

Note that the higher index λ occurs only in the outer integral. Moreover, the interactions between the indexes occur also only in the outer integral: in the Gamma function $\Gamma(\lambda - \nu)$ and the power $\lambda - \nu - 1/2$ of $(1 - u^2)$.

Chapter 4

Path properties

As we have previously noted, centred Gaussian processes are determined entirely by their covariance functions. When the process is parametrised by the sphere $\mathbb{S}^d$ with continuous covariance the covariance function is given by $f \circ \cos$ where f has, by Theorem 2.4, a representation of the form

$$f(x) = \sum_{n=0}^{\infty} a_n W_n^{\lambda}(x), \quad a_n \geq 0, \sum_{n=0}^{\infty} a_n = 1, \lambda = (d-1)/2, \tag{4.1}$$

and thus the process's behaviour is encoded entirely within the mixing law (a_n). In this chapter we exploit that fact to quantify several properties of the paths of a Gaussian random field on the sphere.

Starting with Belyaev's dichotomy we survey this very active area of research, looking at both summability and integrability conditions for path continuity. We extend a result of Malyrenko's to the Hilbert sphere $\mathbb{S}^{\infty}$, and close by exploring some extensions to the geo-temporal setting.

4.1 Belyaev's dichotomy and the Dudley integral

As usual, let $X = \{X_t : t \in M\}$ be a real-valued zero-mean Gaussian process, *on* (defined on, indexed by) the sphere $\mathbb{S}^d$ (or any other compact metric space M). The law of X is determined by either of the covariance or the incremental variance:

$$c(s,t) := cov(X_s, X_t) = E[X_s X_t], \qquad i(s,t) := E[(X_t - X_s)^2] \tag{4.2}$$

(respectively positive and negative definite, or of positive and negative type); recall from Chapter 1 that we can pass between them by

$$i(s,t) = c(s,s) + c(t,t) - 2c(s,t), \quad c(s,t) = \frac{1}{2}(i(s,o) + i(t,o) - i(s,t)), \tag{4.3}$$

with o some base point (a 'North Pole'). As usual, we restrict attention to *isotropic* processes, where these are functions only of the geodesic distance $d(s,t)$, or of $x := \cos d(s,t) \in [-1,1]$ $(s, t \in \mathbb{S}^d)$:

$$c(s,t) = C(x), \qquad i(s,t) = I(x). \tag{4.4}$$

If the covariance function is continuous, then it is immediately clear that the process X is *mean-square continuous*, i.e. that for all $s, t \in \mathbb{S}^d$

$$s \to t \Rightarrow \mathbb{E}[\,|X(s) - X(t)|^2] \to 0. \tag{4.5}$$

But, continuity in mean-square does not imply (nor is implied by) continuity of the sample paths of the process X. To make statements about that, we need some new tools. For reference here, we use [MR06, Ch. 5] which exposits the theory for processes on more general separable metric spaces than the sphere. Indeed, all of the results in this section can be stated quite generally, at the very least for processes on compact metric spaces whose Karhunen-Loève expansion (see Theorem 4.1 below) converges a.s. at fixed points in their domain.

First, we need the notion of a reproducing kernel Hilbert space (RKHS).

Definition 4.1 (Reproducing kernel Hilbert space (RKHS) [MR06, Th. 5.3.1])**.** *Let Γ be a continuous covariance function of $\mathbb{S}^d \times \mathbb{S}^d$ (not necessarily isotropic). Then the reproducing kernel Hilbert space of Γ is the separable Hilbert space $H(\Gamma)$ of continuous real-valued functions on $\mathbb{S}^d$ such that*

$$\Gamma(t, \cdot) \in H(\Gamma), \quad t \in \mathbb{S}^d; \tag{4.6}$$

$$\langle f(\cdot), \Gamma(t, \cdot)\rangle = f(t), \quad f \in H(\Gamma), t \in \mathbb{S}^d. \tag{4.7}$$

The property (4.7) is the reproducing property which gives RKHSs their name – $\Gamma = \Gamma(\cdot, \cdot)$ is called the reproducing kernel of $H(\Gamma)$.

An alternative definition of RKHS's is as Hilbert spaces whose point evaluation maps[1] are continuous. Definition 4.1 allows us to prove the following representation theorem for the process X.

Theorem 4.1 (Karhunen-Loève Expansion [MR06, Th. 5.3.2, Cor. 5.3.4]). *Let X be a mean zero Gaussian process on a separable metric space M with continuous covariance. Then X has a version given by*

$$X'(t) = \sum_{i=0}^{\infty} \gamma_i(t)\xi_i, \tag{4.8}$$

where the γ_i are continuous functions on M, ξ_i are independent Gaussian random variables and the convergence is in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ and almost sure for fixed t.

The Karhunen-Loève expansion is well-known, but we include the proof here due to the importance of the result.

Proof. We follow [MR06, Th. 5.3.2, Cor. 5.3.4]. Suppose X has continuous covariance Γ with reproducing kernel $H(\Gamma)$. Let

$$\mathcal{L}_2(X) := \text{Closure of } \left\{ \sum_{i=1}^{n} a_i X(t_i) : a_i \in \mathbb{R}, t_i \in M, i = 1, 2, \ldots n \right\} \tag{4.9}$$

in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, and further

$$S := \left\{ \sum_{i=1}^{n} a_i \Gamma(t_i, \cdot) : a_i \in \mathbb{R}, t_i \in M, i = 1, 2, \ldots n \right\}. \tag{4.10}$$

[1]For fixed x the point evaluation map is the map $\delta_x : \Gamma \mapsto \mathbb{R}$, $\delta_x : f \mapsto f(x)$.

Then we can define a linear, bijective, norm-preserving map $\Theta_{\mathbb{P}} : S \to \mathcal{L}_2(X)$ by

$$\Theta_{\mathbb{P}}\left(\sum_{i=1}^{n} a_i \Gamma(t_i, \cdot)\right) = \sum_{i=1}^{n} a_i X(t_i). \tag{4.11}$$

This is the *Loève isometry*. Note that the right hand size is a Gaussian random variable (it is the sum of Gaussian random variables).

Let γ_i be a complete orthonormal set in $H(\Gamma)$ and set $\xi_i = \Theta_{\mathbb{P}}(\gamma_i)$. Then for each $i \in \mathbb{N}$ ξ_i is a Gaussian random variable. Moreover, since the γ_i is an orthonormal set, the ξ_i form a complete orthonormal set in $\mathcal{L}_2(X)$ (and are thus independent Gaussian random variables). Since X is clearly in $\mathcal{L}_2(X)$, we then have that

$$X(t) = \sum_{i=1}^{\infty} \mathbb{E}(X(t)\xi_j)\xi_i. \tag{4.12}$$

(Note: the left-hand side is a random variable, whilst the right-hand side is an equivalence class of $L^2(\Omega, \mathcal{F}, \mathbb{P})$ and the equality must be interpreted accordingly, as a statement about equivalence classes.)

Now, since $\Omega_{\mathbb{P}}$ is an isometry,

$$\mathbb{E}(\Theta_{\mathbb{P}}(f)\Theta_{\mathbb{P}}(g)) = \langle f, g \rangle, \tag{4.13}$$

and so, using also the reproducing kernel property of Γ,

$$\mathbb{E}(X(t)\xi_i) = \langle \Gamma(t, \cdot), \gamma_i(\cdot) \rangle = \gamma_i(t), \tag{4.14}$$

giving the Karhunen-Loève expansion

$$X(t) = \sum_{i=1}^{\infty} \gamma_i(t)\xi_i, \tag{4.15}$$

with convergence in $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Fix $t \in M$. Then $(\gamma_i(t)\xi_i)_{i\in\mathbb{N}}$ form a sequence of independent Gaussian random variables whose sum $\sum \gamma_i(t)\xi_i$ converges almost surely, by Kolmogorov's three-series criterion [Kal97, Th. 3.18], if and only if

(a) $\sum \gamma_i(t)\xi_i$ converges in distribution – this is given by the L^2 convergence found above;

(b) $\sum \mathbb{P}\{|\gamma_i(t)\xi_i| > 1\} < \infty$;

(c) $\sum \mathbb{E}\left[\gamma_i(t)\xi_i : |\gamma_i(t)\xi_i| \leq 1\right]$ converges – this is immediate, since the ξ_i have mean zero;

(d) $\sum \mathrm{Var}\left[\gamma_i(t)\xi_i : |\gamma_i(t)\xi_i| \leq 1\right]$ converges.

To show (d) consider the sum's upper bound $\sum \mathrm{Var}\left[\gamma_i(t)\xi_i\right] = \sum \gamma_i^2(t)$. Since the $\{\xi_i\}$ are independent standard Gaussians, for fixed $t \in M$:

$$\infty > \mathbb{E}\left[X(t)^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^{n} \gamma_i(t)\xi_i\right)^2\right] \tag{4.16}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_i(t)\gamma_j(t)\mathbb{E}\left[\xi_i\xi_j\right] \tag{4.17}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_i(t)\gamma_j(t)\delta_{ij} = \sum_{i=1}^{n} \gamma_i^2(t). \tag{4.18}$$

So (d) holds. Applying Chebyshev's inequality to (b) and using the above:

$$\sum_{i=1}^{n} \mathbb{P}\left\{|\gamma_i(t)\xi_i| > 1\right\} \leq \mathrm{Var}(|\gamma_i(t)\xi_i|) = \sum \gamma_i^2(t) < \infty. \tag{4.19}$$

Thus, for fixed $t \in M$, (4.8) converges almost surely. □

It can be shown further [MR06, Cor. 14.6.4] that if X has continuous sample paths then (4.8) converges *uniformly* on compact spaces M with probability one.

But when does X have continuous sample paths? Continuity of the covariance function is uninformative, here, but the Karhunen-Loève expansion opens the door to a library of zero-one laws.

We recall one of the best known of these: Belyaev's dichotomy for Gaussian processes ([Bel61]; [MR06, Th 5.3.10]) – colloquially, Belyaev's dichotomy says that Gaussian paths are either 'very nice', or 'very nasty'.

To prove this, we need to introduce a new tool: the *oscillation function*. For general (non-random) functions f this is defined as

$$W_f(t) := \lim_{\epsilon \to 0} \sup_{u,v \in B_\epsilon(t)} |f(u) - f(v)|, \tag{4.20}$$

where $B_\epsilon(t)$ is the closed ball of radius ϵ centred at $t \in M$. The oscillation function is 0 if and only if f is continuous. We have

$$W_f(t) = M_f(t) - m_f(t), \tag{4.21}$$

where

$$M_f(t) := \lim_{\epsilon \to 0} \sup_{u \in B_\epsilon(t)} f(u), \tag{4.22}$$

$$m_f(t) := \lim_{\epsilon \to 0} \inf_{u \in B_\epsilon(t)} f(u), \tag{4.23}$$

using the convention that $\infty - \infty = 0$. Then,

Definition 4.2 ([MR06, Th. 5.3.7])**.** *The* oscillation function *of a Gaussian process X on a separable metric space M is the upper-semi-continuous function $\alpha : M \to \mathbb{R}$ with the properties*

$$\mathbb{P}(W_{X(\cdot,\omega)}(t)) = \alpha(t) : t \in M) = 1; \tag{4.24}$$

$$\mathbb{P}\left(M_{X(\cdot,\omega)}(t) = X(t,\omega) + \frac{\alpha(t)}{2}, m_{X(\cdot,\omega)}(t) = X(t,\omega) - \frac{\alpha(t)}{2} \right) = 1 \quad \forall t \in \mathbb{S}^d. \tag{4.25}$$

We can now state and prove Belyaev's Dichotomy for spheres $\mathbb{S}^d \subset \mathbb{R}^{d+1}$.

Theorem 4.2 (Belyaev's Dichotomy [MR06, Th. 5.3.10], [Bel61])**.** *Let $\{X(t) : t \in \mathbb{S}^d\}$ be an isotropic Gaussian process on $(\mathbb{S}^d, d(\cdot,\cdot))$. Then either (a) X has continuous paths almost surely on all open $S \subset \mathbb{S}^d$; or (b) X is unbounded almost surely on all open $S \subset \mathbb{S}^d$.*

Proof. Let α be the oscillation function of X. Since X is isotropic its incremental variance $i(s,t) = \mathbb{E}[|X(s) - X(t)|^2]$ depends only on $d(s,t)$, and (since the increments have mean zero) the law of $(X(t)-X(s))$ is therefore a function solely of the distance $d(s,t)$ between s and $t \in \mathbb{S}^d$. Then, by (4.24), α is constant on all open subsets of $\mathbb{S}^d$. So either:

a) $\alpha = 0$, meaning X is continuous almost surely on all open $S \subset \mathbb{S}^d$.

b) $\alpha = a > 0$. This follows by [MR06, Th. 5.3.9]. Let E be a dense subset of $S \subset \mathbb{S}^d$ with separability set D, so $D \cap S = E$. Then for fixed $t \in I'$ define

$$F_n(t) := \{u \in \mathbb{S}^d : d(t,u) < 1/n\}. \tag{4.26}$$

Then by definition of the oscillation function, and noting that if $u \in D \cap F_j(t)$ for $j \geq n$, $u \in D \cap F_{n/2}(t)$:

$$\begin{aligned}
M_{X(\cdot,\omega)}(t) &= \lim_{s \to t} \sup_{s \in \mathbb{S}^d} X(s,\omega) && (4.27)\\
&= \lim_{n\to\infty} \sup_{s \in D \cap F_n(t)} X(s,\omega) && (4.28)\\
&\geq \lim_{n\to\infty} \sup_{s \in D \cap F_n(t)} \left(\lim_{j\to\infty} \sup_{u \in D \cap F_j(s)} X(u,\omega) \right) && (4.29)\\
&= \lim_{n\to\infty} \sup_{s \in D \cap F_n(t)} (X(s,\omega) + \alpha(s)/2) && (4.30)\\
&= \lim_{n\to\infty} \sup_{s \in D \cap F_n(t)} (X(s,\omega) + a/2) && (4.31)\\
&= M_{X(\cdot,\omega)}(t) + a/2. && (4.32)
\end{aligned}$$

Since $a > 0$ it follows that $\mathbb{P}(M_{X(\cdot,\omega)}(t) = \infty; t \in D \cap S) = \mathbb{P}(M_{X(\cdot,\omega)}(t) = \infty; t \in E) = 1$ and hence, since E is dense in $\mathbb{S}^d$, $M_{X(\cdot,\omega)}(t) = \infty$ for all $t \in I$ with probability one. The corresponding argument for $m_{X(\cdot,\omega)}(t)$ is identical. □

The natural question is, then: on what side of the dichotomy does a given process X fall? That is: what are the necessary and sufficient conditions for a process X to have continuous sample paths?

This is one of the oldest problems in stochastic-process theory, and much is known by way of necessary conditions for continuity [MR06, §6.2], and sufficient conditions [MR06, §6.1]; see also [MS70; MS72; Gar72]. To formulate these results in the context of the sphere, and give a necessary and sufficient condition for path continuity in terms of the Schoenberg coefficients, we need, as in the standard theory, to introduce the concept of *metric entropy*.

Definition 4.3. *For a Gaussian process $\{X_t : t \in \mathbb{S}^d\}$, the Dudley metric is defined as:*

$$d_X(s,t) := \sqrt{E[(X_s - X_t)^2]} = \sqrt{i(s,t)}, \qquad s,t \in \mathbb{S}^d, \tag{4.33}$$

where $i(s,t)$ is the incremental variance familiar from previous chapters.

Note that the Dudley metric is, more precisely speaking, a pseudo-metric: $d_X(s,t) = 0 \not\Rightarrow s = t$.

For $u > 0$, write $N(u)$ for the minimum number of d_X-balls of radius u required to cover $\mathbb{S}^d$; then if $H(u) := \log N(u)$, $H := \{H(u) : u > 0\}$ is called the *metric entropy with respect to the pseudo-metric d_X* (the term is due to Lorentz, the concept to Kolmogorov using the term 'ϵ-entropy'). The *Dudley* *integral* is

$$\int_0^{\epsilon} \sqrt{H(u)}du, \qquad \epsilon > 0. \tag{4.34}$$

When the process X is isotropic, we can obtain a clean necessary and sufficient condition for continuity, namely, finiteness of (4.34) [MR06, Th. 6.1.2]. In the Euclidean case this result is due to Dudley [Dud67], with the general case discussed in [Dud73, §5]. The case of isotropic fields on compact subspaces of two-point homogeneous spaces is examined in Lifshits [Lif95, Th. 3, §15].

Estimating the metric entropy is a non-trivial task. But, it is possible to find a condition for the finiteness of (4.34) without needing to estimate $H(u)$, by instead making use of the incremental variance.

To do this, we need to briefly introduce the *non-decreasing rearrangement* of the incremental variance $i(s,t) = (d_X(s,t))^2$. Note that i is a function from $\mathbb{S}^d \times \mathbb{S}^d \to \mathbb{R}^+$ (although, since X is isotropic it depends only on $d(s,t)$). Define

$$m_i(\epsilon) := \lambda((s,t) \in \mathbb{S}^d \times \mathbb{S}^d : i(s,t) < \epsilon), \tag{4.35}$$

where $\lambda\{\cdot\}$ denotes, as usual, Haar measure. Then the non-decreasing rearrangement of i is

$$\overline{i(u)} := \sup \{y : m_i(y) < u\}. \tag{4.36}$$

Then we may define our alternative to the Dudley integral:

$$J(d_X) = J(i) := \int_0^{\epsilon} \frac{\overline{i(u)}}{\sqrt{-\log u}} \frac{du}{u}. \tag{4.37}$$

This is equivalent to (4.34) (when (4.34 is finite) by a technical lemma from Marcus & Pisier [MP81, Lemma 3.6]:

Lemma 4.1 ([MP81])**.** *If the Dudley integral $I(i) := \int_0^{\epsilon} \sqrt{H_i(u)}\, du < \infty$ (where H_i denotes the metric entropy with respect to the Dudley metric associated with incremental variance i) then*

$$-C_1\bar{i} + \frac{1}{2}J(i) \leq I(\epsilon) \leq C_2\bar{i} + 2J(i), \tag{4.38}$$

where $0 < C_1, C_2 < \infty$ and $\bar{i} := \sup\{i(x) : x \in \mathbb{S}^d\}$.

Moreover, by [JM74, Cor. 2.5],

$$J(i) = \int_0^{\epsilon} \frac{\overline{i(u)}}{\sqrt{-\log u}} \frac{du}{u} \leq \int_0^{\epsilon} \frac{i(u)}{\sqrt{-\log u}} \frac{du}{u}. \tag{4.39}$$

Thus, combining, the Dudley integral (4.34) is finite if

$$\int_0^{\epsilon} \frac{i(u)}{\sqrt{-\log u}} \frac{du}{u} < \infty. \tag{4.40}$$

We can now, for isotropic processes on spheres, formulate this explicitly.

Recall that, when X is isotropic $i(s,t) = I(x)$, where $x = \cos d(s,t)$. Take

$$\phi(u) := \sup\{\sqrt{I(\cos d(s,t))} : d(s,t) \le u\}, \qquad s,t \in \mathbb{S}^d; \tag{4.41}$$

then clearly

$$\int_0^\epsilon \frac{i(u)}{\sqrt{-\log u}} \frac{du}{u} < \infty \Rightarrow \int_0^\epsilon \frac{\phi(u)}{\sqrt{-\log u}} \frac{du}{u} < \infty. \tag{4.42}$$

The Dudley condition then becomes, taking $v := d(s,t)$ and using Schoenberg's theorem,

$$\int_0^1 \sqrt{\frac{\sup_{v \le u} I(\cos v)}{-\log u}} \frac{du}{u} < \infty; \tag{4.43}$$

$$\int_0^1 \sqrt{\frac{\sup_{v \le u}(1 - \sum_0^\infty a_n W_n^\lambda(\cos v))}{-\log u}} \frac{du}{u} < \infty. \tag{4.44}$$

Finiteness of this integral, then, guarantees continuous sample paths (on all open subsets). When the integral fails to converge the paths are unbounded on all open subsets of $\mathbb{S}^d$. But, there is a sense in which they are 'nearly continuous': a 'localisation of pathology'. If (4.44) does not hold for a process X, then its Karhunen-Loève expansion,

$$X(t,\omega) = \sum_{i=0}^\infty \gamma_i(t)\xi_i(\omega), \tag{4.45}$$

is not uniformly convergent a.s., but it is convergent a.s. and in L^2. In particular, this shows that a.s. $X(t) = X(t,\omega)$ is a *measurable* function of t. So, we can invoke Lusin's continuity theorem (or Lusin's restriction theorem, of 1912: [Dud90, Th. 7.5.2], [Rud87, §2.24]).

Theorem 4.3 (Lusin's Continuity Theorem). *Suppose f is a measurable function on a compact space M. Let $\epsilon > 0$. Then there exists a $g \in C(M)$ such that*

$$\lambda(\{x : f(x) \neq g(x)\}) < \epsilon \tag{4.46}$$

and

$$\sup_{x \in M} |g(x)| \leq \sup_{x \in M} |f(x)|. \tag{4.47}$$

Applying this to $X(t)$ we see that there exists a *continuous* process X', taking values identical to X (for fixed ω) apart from on a set of arbitrary small measure.

Thus, when slightly restricted, discontinuous paths become continuous. A similar idea is pursued in [Adl90], using the idea of Lebesgue density. Adler restricts his attention to the sample paths of $\{X(t) : t \in [0,1]\}$. By a result of Geman [Adl90, p.114], the set A has zero Lebesgue density at $t = 0$ iff

$$\int_A \Psi \left(\frac{\lambda(A \cap (0,t))}{t} \right) \frac{dt}{t} < \infty.$$

for some continuous strictly increasing Ψ such that $\Psi(0) = 0$. Adler quotes Don Geman proposing the following conjecture [Adl90, p.114]: "If, after we have drawn a discontinuous, unbounded, Gaussian sample path on the blackboard, we were to step far enough backwards so that we could no longer see sets of zero Lebesgue density, the sample path would become continous and bounded." The above Lusin argument gives this if we can no longer see sets of arbitrary small measure, but not necessarily sets of Lebesgue density zero.

4.2 Malyarenko's theorem

Examination of (4.44) above reveals that the paths of a Gaussian process on the sphere are continuous if and only if the coefficients a_n in its covariance's Schoenberg expansion (cf. Theorem 2.4)

$$\operatorname{cov}(X_s, X_t) = \sum_{n=0}^{\infty} a_n W_n^{\lambda}(\cos d(s,t)), \quad s,t \in \mathbb{S}^d \tag{4.48}$$

decay quickly enough[2] – slow decay means wild behaviour of the paths, but if the decay is fast enough, the paths become very smooth, e.g. if $a_n = \mathcal{O}(1/n^{\alpha})$ then the paths will be a.s. in $C^{\lfloor \alpha/2 \rfloor}$.

While the Dudley conditions in the section above resolve the matter completely in principle, in practice one cannot test them, for the obvious reasons: even if one can determine the mixing law (a_n) in closed form, passage between the mixing law (a_n) and the ultraspherical series $\sum a_n W_n^{\lambda}$, the supremum, and the integration together form an intimidating mathematical barrier.

While there is no definite solution to this (any more than there is in the

[2]We note here that in some of the literature, particularly that aimed at cosmological applications, (a rescaling of) the mixing law (a_n) is referred to as the *angular power spectrum*.

classical case of Fourier series [Zyg68]), there is an answer in a principal case of practical interest, that when the mixing law (a_n) is *regularly varying* (see e.g. [BGT87]). Here the results are due to Malyarenko [Mal05; Mal13], based on early works of Bingham [Bin79], Askey and Wainger [AW65]).

First we need to introduce the notion of a slowly varying function, to be able to formulate Malyrenko's theorem.

Definition 4.4. *A measurable function $\ell : (0, \infty) \to \mathbb{R}$ is* slowly varying *if*

$$\frac{\ell(\lambda x)}{\ell(x)} \to 1, \qquad x \to \infty, \tag{4.49}$$

for all $\lambda > 0$.

Theorem 4.4 (Malyrenko [Mal13, Th. 4.8])**.** *Let X be a mean-zero Gaussian isotropic random field on the sphere $\mathbb{S}^d$ with incremental variance $I : [-1, 1] \to \mathbb{R}$. Then, for ℓ slowly varying,*

$$A_n := \sum_{k=n}^{\infty} a_k \sim \ell(n)/n^{\gamma}, \quad n \to \infty, \quad \gamma \in (0, 2) \tag{4.50}$$

if and only if

$$I(v) = 1 - \sum_{n=0}^{\infty} a_n W_n^{\lambda}(\cos v) \sim \Gamma(\lambda + \frac{1}{2}).\frac{\Gamma(1 - \frac{1}{2}\gamma)}{2^{\gamma}\Gamma(\lambda + \frac{1}{2} - \frac{1}{2}\gamma)}.v^{\gamma}\ell(1/v), \quad v \downarrow 0. \tag{4.51}$$

Proof. We sketch Malyrenko's proof here. The implication from A_n to $I(v)$

is Abelian; the converse is Tauberian. We have

$$I(v) = \sum_{n=0}^{\infty} a_n (1 - W_n^{\lambda}(v)) = \sum_{n=0}^{\infty} (A_n - A_{n+1})(1 - W_n^{\lambda}(v)). \tag{4.52}$$

Writing this by partial summation:

$$I(v) = \sum_{n=0}^{\infty} A_{n+1}(W_n^{\lambda}(v) - W_{n+1}^{\lambda}(v)). \tag{4.53}$$

The difference of ultraspherical polynomials here may be expressed as a single Jacobi polynomial (Erdélyi et al. [Erd+81, Vol. II, 10.8]). Recall that the Jacobi polynomials are a two-index family $P_n^{(\alpha,\beta)}$ ($\alpha, \beta \geq -\frac{1}{2}$; we take $\alpha \geq \beta$). When $\alpha = \beta$, the Jacobi polynomials reduce to the ultraspherical polynomials, with

$$\alpha = \beta = \lambda - \frac{1}{2} = \frac{1}{2}(d-2). \tag{4.54}$$

We use the normalisation [Mal2, 4.3.1]

$$R_n^{(\alpha,\beta)}(x) := P_n^{(\alpha,\beta)}(x) / P_n^{(\alpha,\beta)}(1). \tag{4.55}$$

Then ([Mal13, p. 127])

$$R_n^{\alpha,\beta}(\cos\theta) - R_{n+1}^{\alpha,\beta}(\cos\theta) = \frac{(2n+\alpha+\beta+2)}{(\alpha+1)} \sin^2 \frac{1}{2}\theta \; R_n^{\alpha+1,\beta}(\cos\theta). \tag{4.56}$$

So

$$I(\cos\theta) = \frac{2\sin^2 \frac{1}{2}\theta}{(\alpha+1)} \sum_{n=0}^{\infty} (n+\alpha+1) A_n R_n^{\alpha+1,\beta}(\cos\theta). \tag{4.57}$$

The $\sin^2\frac{1}{2}\theta$ (equivalently, $\theta^2/4$) term on the right accounts for the upper limit 2 on γ in the result; that the incremental variance is non-negative accounts for the lower limit of 0. The results of [Bin78; Bin79] now apply to the sequence $(n+\alpha+1)A_n = (n+\alpha+1)\sum_n^\infty a_k$ with the σ there as $1-\gamma$.

The Tauberian conditions needed follow from $a_n \geq 0$ (so A_n is non-negative and non-decreasing). $\square$

In fact, this theorem applies on any compact two-point homogeneous space, where in general isotropic processes have covariance functions given by

$$\operatorname{cov}(X_s, X_t) = \sum_{n=0}^{\infty} a_n R_n^{(\alpha,\beta)}(\cos d(s,t)) \tag{4.58}$$

where $R_n^{(\alpha,\beta)}$ are normalised Jacobi polynomials (as in the above proof).

Malyarenko's theorem is very similar to that of [Bin72c] on Hankel transforms, the link being provided by Szegő's Hilb-type asymptotic formula for the Jacobi polynomials [Sze39, Th. 8.21.12].

The Belyaev integral is convergent here, and so Malyarenko's theorem provides us with an ample range of examples of the continuous case in the Belyaev dichotomy. The supremum operation in (4.44) is not an obstacle here, since any regularly varying function of *non-zero* index is asymptotically monotone [BGT87, §1.5.2]).

It is possible to extend to the boundary case $\gamma = 0$ here, when the tail A_n of the mixing law is slowly varying, but convergence of the Dudley integral now hinges on the behaviour of ℓ at infinity. This is shown by familiar examples – $\sum 1/(n(\log n)^k)$ is convergent if $k > 1$, and divergent if $k \le 1$. One can also extend to the case $\gamma = 2$ [Bin72c].

A complementary approach to Malyrenko's is taken by Kerkyacharian et al. [Ker+18, §7.22], who show that if the mixing law satisfies $a_n = O(1/n^{1+\gamma})$ for $\gamma > 0$ (so $A_n := \sum_n^\infty a_k = O(1/n^\gamma)$), then the sample paths of the process X are a.s. in the Besov space $B^\alpha_{\infty,1}$ for all $\alpha < \gamma$ (there is much more on Besov spaces to come below, in §4.4 – in the context of [Ker+18] see Giné and Nickl [GN16] for the theory of Besov spaces and Fukushima et al. [FOT11] for the necessary Dirichlet structure on the index set, $\mathbb{S}^d$ here). Thus, *the faster the decay of the mixing law, the smoother the paths of the process.*

4.3 The Hilbert sphere

We now turn to a new result: an extension of Theorem 4.4 to the Hilbert sphere. The Hilbert sphere $\mathbb{S}^\infty$ is not locally compact, and because of this one may expect very different behaviour for it from that on Euclidean spheres.

Gaussian processes on $\mathbb{S}^\infty$ are *deterministic* (see [Ber74] for the definition): the behaviour of the process locally determines it everywhere [Ber80, Th. 4.1]. In fact, Berman shows that, by replacing independence assumptions with orthogonality ones, *any* centred stochastic process on $\mathbb{S}^\infty$ with covariance of the form (2.25) is deterministic.

Recall that ultraspherical polynomials may be defined for $\lambda = \infty$ by $W_n^\infty(x) = x^n$ and the general covariance for a process X on the Hilbert sphere is of the form (2.25):

$$\sum_{n=0}^{\infty} a_n x^n, \quad a_n \geq 0, \sum_{n=0}^{\infty} a_n = 1. \tag{4.59}$$

This is, as is remarked in [Ber80], a much simpler form than the finite-dimensional case, a fact which opens up a lot of interesting theory, especially in the Gaussian case – see [Ber80] for a definitive overview. For more recent work on covariance functions and Gaussian processes on the Hilbert sphere, see [BPP18; Jäg19].

Since

$$\frac{\Gamma(\nu + \frac{1}{2} + \frac{1}{2}\gamma)}{\Gamma(\nu + \frac{1}{2})} \sim \nu^{\frac{1}{2}\gamma} \to \infty \quad (\nu \to \infty), \tag{4.60}$$

the infinite-dimensional case of Theorem 4.4 does not follow formally by letting $\nu \to \infty$. Instead, we have the following:

Theorem 4.5. *In the notation of Malyarenko's theorem, with* $\gamma \in (0, 2)$

$$A_n := \sum_{k=n}^{\infty} a_k \sim \ell(n)/n^{\frac{1}{2}\gamma}, \qquad n \to \infty \tag{4.61}$$

if and only if

$$I(v) = 1 - \sum_{n=0}^{\infty} a_n (\cos v)^n \sim \frac{\Gamma(1 - \frac{1}{2}\gamma)}{2^{\frac{1}{2}\gamma}} \cdot v^\gamma \ell(1/v^2), \qquad v \downarrow 0. \tag{4.62}$$

Proof. The functions in $\mathcal{P}(\mathbb{S}^\infty) = \mathcal{P}^\infty$ are probability generating functions (in t, say), or (putting $t = e^{-s}$) Laplace-Stieltjes transforms.

Let $s = -\log \cos v$, so that $\cos v = e^{-s}$. The Schoenberg series $\sum a_n e^{sn}$ thus takes the form of the Laplace-Stieltjes transform of the mixing law.

[BGT87, Cor. 8.1.7] links the tail behaviour of a distribution F with its Laplace-Stieltjes transform $\hat{F}$. It states that, for $\alpha \in [0, 1)$,

$$1 - F(x) \sim \ell(x)/x^{\alpha}, \quad x \to \infty, \tag{4.63}$$

if and only if

$$1 - \hat{F}(s) \sim \Gamma(1-\alpha)s^{\alpha}\ell(1/s), \qquad s \to 0. \tag{4.64}$$

So, taking $\alpha = \gamma/2 \in [0, 1)$, (4.61) holds if and only if

$$1 - \sum_{n=0}^{\infty} a_n e^{-sn} \sim \Gamma(1 - \frac{1}{2}\gamma)s^{\frac{1}{2}\gamma}\ell(1/s). \tag{4.65}$$

Comparing expansions around zero of $\cos v$ and e^{-s}, we have $s \sim \frac{1}{2}v^2$ and so, taking $\alpha = \gamma/2 \in [0, 1)$ (4.65) is equivalent to

$$1 - \sum_{n=0}^{\infty} a_n (\cos v)^n \sim \frac{\Gamma(1 - \frac{1}{2}\gamma)}{2^{\frac{1}{2}\gamma}} \cdot v^{\gamma}\ell(1/v^2), \qquad v \downarrow 0, \tag{4.66}$$

as required. □

It is not surprising that the tails here are heavier than in the finite dimensional case of Theorem 4.4 – there there are 'more ways of going off to infinity'. Thus the relevant probability laws (a_n) here have regularly varying tails in $(0, 1)$, rather than in $(0, 2)$ as in the Euclidean case of Malyarenko's theorem.

The constants introduced (in going between the 'Abelian' and 'Tauberian' sides) in results of this type are the values, for $s = \gamma$, of the *Mellin transform*

$$\hat{k}(s) := \int_0^\infty u^s k(u) du/u \quad (s \in \mathbb{C})$$

of the kernel k in the relevant Mellin-Stieltjes convolution (see e.g. [BGT87, Ch. 4, 5]).

4.4 Integrability conditions

The question of path-continuity of process on spheres is also addressed in the work of Lang and Schwab [LS15] (cf. [AL14]) and Lan, Marinucci and Xiao [LMX18]. The picture is much as above: the faster the decay of the Schoenberg coefficients/mixing law, the better: the more regular the paths of the process (and, as above, the faster the decay of the incremental variance at the origin).

In [LS15, §4, Assumption 4.1], Lang and Schwab assume a decay condition on the mixing law (a_n) measured by a *summability condition* (rather than by rate of decay as in §4.3 above): in our notation, they assume

$$\sum_{n=0}^{\infty} a_n n^\gamma < \infty \qquad (\gamma > 0). \tag{4.67}$$

In view of the work above, we re-write this by partial summation as

$$\sum_{n=0}^{\infty} A_n.n^{\gamma-1} < \infty : \qquad \sum_{n=0}^{\infty}(n+\alpha+1)A_n \cdot n^{\gamma-2} < \infty. \tag{4.68}$$

As in §4.3, and in [LS15, §4], the case $\gamma \in (0,2)$ is especially important, so we begin with that. Then the summability condition (4.67) may [Bin79, Th. 1] be translated into a corresponding *integrability condition* on the incremental variance at the origin: (4.67) implies

$$\int_{0+}^{\pi/2} I(\cos\theta) \cdot \theta^{-\gamma} \frac{d\theta}{\theta} < \infty. \tag{4.69}$$

As $\int_{0+} d\theta/\theta$ diverges, this gives in particular that

$$I(\cos\theta) = o(\theta^{\gamma}) \quad (\theta \downarrow 0). \tag{4.70}$$

This strengthens the result of [LS15, Lemma 4.2] from $O(.)$ to $o(.)$ (though in view of the 'ϵ-gap' in [LS15, Th. 4.7], where it is used, this does not matter).

This leads quickly to the path-regularity result ([LS15]; cf. [LMX18]):

Theorem 4.6 ([LS15, Th. 4.5]). *If X is an isotropic Gaussian random field on the sphere whose Schoenberg coefficients (a_n) satisfy the summability condition (4.67), then for any $\delta < \gamma/2$ X has a continuous version: for $k := \lfloor \gamma/2 \rfloor$, the modification is k times continuously differentiable, with kth derivative Hölder continuous with exponent $\delta - k$.*

Proof. (Sketch). The proof consists of three steps:

(i) For $n \in N$, $x, y \in \mathbb{S}^d$

$$E[|X(x) - X(y)|^{2n}] \leq C_{\gamma,n} d(x,y)^{\gamma n},$$

with $d(.,.)$ geodesic distance as before [LS15, Lemma 4.3].
(ii) The Kolmogorov-Chentsov theorem on manifolds [AL14] gives the result for $\gamma \in (0, 2]$.
(iii) For $\gamma > 2$, fractional differentiation can be used to the range above.

We refer for full detail to [LS15; AL14]. □

The crucial step is the application of the Kolmogorov-Chentsov theorem for manifolds. In the Euclidean setting, this classical result is as follows:

Theorem 4.7 (The Kolmogorov-Chentsov Theorem)**.** *If X is a stochastic process on R^d and*

$$\mathbb{E}[\,|X_t - X_s|^p] \leq c|t-s|^{d+s+kp}, \qquad s,t \in \mathbb{R}^d \tag{4.71}$$

for some constants $p > 0$, $s \in (0,p)$, $k \in \mathbb{N} \cup 0$, then for any $\sigma < s/p$,

$$|X_t - X_s| \leq C(\omega)|t-s|^{k+\sigma}, \qquad s,t \in \mathbb{R}^d \tag{4.72}$$

for some random variable $C(\omega) < \infty$ with probability one. In other words: X has a.s. a $(k+\sigma)$–Hölder continuous version.

The notion of a Besov space and Besov Embedding Theorem, generalisations of the earlier Sobolev spaces and embedding theorem, together allow for an extremely short proof of this powerful theorem, due to Schilling [Sch00].

Define, using the notation (as in Schilling) $||u|X||$ to denote the norm of u with respect to the space X, the *Besov space* on $\mathbb{R}^d$ to be the space

$$B^s_{pp}(\mathbb{R}^n) := \left\{ u \in L^p(\mathbb{R}^d) : ||u|B^s_{pp}(\mathbb{R}^n)|| < \infty \right\} \tag{4.73}$$

with $0 < p \leq \infty$, $n \max(1/p - 1, 0) < s < M$, $M \in \mathbb{N}$ with quasi-norm

$$||u|B^s_{pp}(\mathbb{R}^d)|| := ||u|L^p(R^d)|| + \left(\int_{|h| \leq \nu} |h|^{-sp} ||\Delta^M_h u|L^p(\mathbb{R}^d)||^p \frac{dh}{|h|^n} \right)^{1/p}, \tag{4.74}$$

and

$$\Delta^1_h u := u(\cdot + h) - u(\cdot) \tag{4.75}$$

$$\Delta^M_h u := \Delta^{M-1}_h (\Delta^1_h u). \tag{4.76}$$

Theorem 4.8 (Besov Embedding Theorem [Sch00; Tri83])**.** $B^{s+n/p}_{pp}(\mathbb{R}^d) \subset C^s(\mathbb{R}^d)$ *for* $0 < p \leq \infty$, $0 < s \notin \mathbb{N}$, *up to modification on a set of measure zero.*

Note that Theorem 4.8 is a generalisation of Sobolev's Embedding Theorem, which (as the name suggests) refers only to Sobolev spaces $W^s_p(\mathbb{R}^d)$. The Sobolev spaces $W^s_p(\mathbb{R}^d)$ coincide entirely with the Besov spaces $B^s_{pp}(\mathbb{R}^d)$ when $p \geq 1$.

Proof of Theorem 4.7 [Sch00]. The expectation $\mathbb{E}[|X_t - X_s|^p]$ is finite, so the random variable $X_t(\omega)$ is finite a.s. and thus its paths $t \to X_t(\omega)$ are in the Besov space $B_{pp}^{\sigma+k+n/p}(\mathbb{R}^d)$ a.s. So, by Sobolev's (or Besov's) Embedding Theorem, the process X has a version X' in the Hölder space $C^{\sigma+k}(\mathbb{R}^d)$ for each $\sigma < s/p$. □

The power of Schilling's approach is that it easily extends to the non-Euclidean case, giving an alternative proof of the Lang-Schwab Theorem. For this, we need a) Besov spaces on spheres; and b) a Sobolev/Besov Embedding Theorem.

For (a), see [FFP16]. The definition here is a little more complex. On the sphere $\mathbb{S}^d = SO(d+1)/SO(d)$ define the vector fields

$$X_{i,j} = x_j \partial_{x_i} - x_i \partial_{x_j}, \quad i < j. \tag{4.77}$$

These generator a family of one-parameter groups of rotations in $SO(d+1)$, defined by

$$\exp \tau X_{i,j} \cdot (x_1, \dots, x_{d+1}) = (x_1, \dots, x_i \cos \tau - x_j \sin \tau, \dots x_i \sin \tau + x_j \cos \tau, \dots x_{d+1}) \tag{4.78}$$

(see [FFP16, §4.4]). Then, c.f. [FFP16, §10.2] define

$$T_j(\tau)u(x) := u(\exp \tau X_{i,j} \cdot x),\ x \in \mathbb{S}^d \tag{4.79}$$

and the new difference operator

$$\Omega_{h,p}^{M}(u) := \sum_{1\leq j_1,\dots,j_M\leq d} \sup_{0\leq \tau_{j_1}\leq h} \dots \sup_{0\leq \tau_{j_M}\leq h} ||(T_{j_1}(\tau_{j_1})-I)\dots(T_{j_M}(\tau_{j_M})-I)u|L^p(\mathbb{S}^d)||, \tag{4.80}$$

where I is the identity operator in $L^p(\mathbb{S}^d)$ and, recall, $||u|X||$ denotes the norm of u with respect to the space X. Then the Besov space $B_{pp}^s(\mathbb{S}^d)$ consists of the functions in $u \in L^p(\mathbb{S}^d)$ such that

$$||u|B_{pp}^s(\mathbb{S}^d)|| := ||u|L^p(\mathbb{S}^d)|| + \left(\int_0^\infty h^{-sp}\left(\Omega_{h,p}^M(u)\right)^p \frac{dh}{h}\right)^{1/p} < \infty. \tag{4.81}$$

The final ingredient needed for a Schilling-style proof of Lang-Schwab theorem (Th. 4.6) is a Besov Embedding Theorem applicable to spheres – this is provided by Han [Han95] for spaces of homogeneous type, with the statement directly analogous to that of Theorem 4.8. Spheres are of homogeneous type, and so Schilling's short proof of the Kolmogorov-Chentsov theorem generalises to spheres, simplifying the proof of Theorem 4.6.

4.5 Spherical Harmonics

We've seen that processes defined on any separable metric space M have a convergent (in L^2 and almost surely) Karhunen-Loève expansion

$$X(t) = \sum_{i=0}^{\infty} \gamma_i(t)\xi_i, \qquad t \in M. \tag{4.82}$$

Given a process's Karhunen-Loève expansion, we can read off its covariance:

$$\begin{aligned}
\operatorname{cov}(X_s, X_t) &= \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma_i(s)\xi_i \sum_{j=0}^{\infty} \gamma_j(t)\xi_j\right] && (4.83)\\
&= \sum_{i,j=0}^{\infty} \gamma_i(s)\gamma_j(t)\mathbb{E}\left[\xi_i\xi_j\right] && (4.84)\\
&= \sum_{i,j=0}^{\infty} \gamma_i(s)\gamma_j(t)\delta_{ij} && (4.85)\\
&= \sum_{i=0}^{\infty} \gamma_i(s)\gamma_i(t), && (4.86)
\end{aligned}$$

where (4.85) follows from the fact that $\xi_i \sim N(0,1)$ are independent and thus $\mathbb{E}\left[\xi_i\xi_j\right] = \delta_{ij}$. Clearly we can also go in the other direction, and derive a given process's Karhunen-Loève expansion from its covariance.

We can thus exploit Schoenberg's expansion of the covariance function for isotropic random fields on the sphere to get a more explicit Karhunen-Loève expansion. Although this is not important for the theoretical results on path properties above, it is the key for simulation of such random fields [LS15, §5].

The vital ingredients here are the *spherical harmonics* – recall from Section 2.3 that these are the restrictions to the sphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ of homogeneous polynomials which are harmonic: solutions to the Laplace equation in $\mathbb{R}^{d+1}$

[AAR99, Ch. 9]. For each degree $\lambda = 0, 1, 2, \ldots$ there are

$$c(l,d) := \frac{2l+d-1}{d-1}\binom{l+d-2}{l} \tag{4.87}$$

linearly independent spherical harmonics of degree l. Denote the m^{th} spherical harmonic of degree l by Y_{lm}.

Spherical harmonics are closely related to the ultraspherical polynomials, via the *addition theorem for spherical harmonics* [AAR99, Th. 9.6.3]:

$$\sum_{m=1}^{c(l,d)} Y_{lm}(s)Y_{lm}(t) = W_l^{\lambda}(\langle s,t\rangle) \cdot \frac{c(l,d)}{\omega_d}, \tag{4.88}$$

where $\lambda = (d-1)/2$, ω_d is (as in (2.16)) the surface area of the $d-$sphere and $\langle \cdot, \cdot \rangle$ is the inner product in $\mathbb{R}^{d+1}$ (so in particular, $\cos\langle \cdot, \cdot \rangle$ is geodesic distance on $\mathbb{S}^d$).

Now, let X be an isotropic process on $\mathbb{S}^d$ with mixing law (a_n). Write

$$v_n := a_n \cdot \frac{\omega_d}{c(n,d)}, \quad n = 0, 1, \ldots. \tag{4.89}$$

For readers familiar with the cosmological applications of random fields on the sphere (c.f. [MP11] and its references) this will be familiar, when $d = 2$, as the angular power spectrum we referred to previously.

So, comparing (4.86) with the general Schoenberg expansion of a covariance on the sphere:

$$\sum_{l=0}^{\infty} \gamma_i(t)\gamma_j(s) = \sum_{l=0}^{\infty} a_l W_l^\lambda(\langle s, t\rangle) \tag{4.90}$$

$$= \sum_{l=0}^{\infty} v_l \frac{c(l,d)}{\omega_d} W_l^\lambda(\langle s, t\rangle) \tag{4.91}$$

$$= \sum_{l=0}^{\infty} v_l \sum_{m=1}^{c(l,d)} Y_{lm}(s) Y_{lm}(t) \tag{4.92}$$

$$= \sum_{lm} \sum_{l'm'} \delta_{ll'}\delta_{mm'} v_l Y_{lm}(s) Y_{l'm'}(t) \tag{4.93}$$

$$= \sum_{lm} \sum_{l'm'} \mathbb{E}[\xi_{lm}\xi_{l'm'}] Y_{lm}(s) Y_{l'm'}(t) \tag{4.94}$$

$$= \mathbb{E}\left[\sum_{lm} Y_{lm}(s)\xi_{lm} \sum_{l'm'} Y_{l'm'}(t)\xi_{l'm'}\right] \tag{4.95}$$

where $\xi_{lm} \sim N(0, v_l)$ and writing $\sum_{lm}$ for the double sum $\sum_{l=0}^{\infty}\sum_{m=1}^{c(l,d)}$. That is, the Karhunen-Loève expansion of an istropic process on $\mathbb{S}^d$ takes the form

$$X(t) = \sum_{lm} \frac{Y_{lm}(t)}{v_l} \xi_{lm} \tag{4.96}$$

where the Y_{lm} are the relevant spherical harmonics, and $\xi_{lm} \sim N(0, 1)$ are independent.

The spectral expansion (4.96) is well-known, and well-used – Marinucci and Peccati's [MP11] use the case $d = 2$ to model cosmic microwave background (CMB) radiation; and it provides the easiest-to-implement method of generating sample isotropic fields on $\mathbb{S}^d$ numerically. Below we shall see that this approach can be extended to the geo-temporal case.

4.6 The geo-temporal case

In light of (4.96) it seems appropriate to find a corresponding spectral expansion of random fields on $\mathbb{S}^d \times \mathbb{R}$.

Theorem 4.9. *Let X be a geo-temporal random field with Karhunen-Loève expansion*

$$X(x,t) = \sum_{lm} \sqrt{v_l}\, a_{lm} c_l \phi_l(t) Y_{lm}(x), \quad a_{lm} \sim N(0,1) \tag{4.97}$$

where Y_{lm} are spherical harmonics, $\phi_l(t)$ a characteristic function, c_l constants such that $\sum v_l c(l,d) c_l^2 < \infty$, and writing $\sum_{lm}$ for the double sum $\sum_{l=0}^{\infty} \sum_{m=1}^{c(l,d)}$. Then X has covariance

$$C(X(x,s), X(y,t)) = c \sum_{n=0}^{\infty} a_n \phi_n(t-s) W_n^{\lambda}(\langle x,y \rangle), \tag{4.98}$$

where the coefficients (a_n) are as in Theorem 2.7b. That is, (4.97) *generates the most general Gaussian random field on $\mathbb{S}^d \times \mathbb{R}$ isotropic in space and stationary in time.*

Proof. The covariance calculation is now

$$\operatorname{cov}(X(x,s), X(y,t)) = \mathbb{E}\left[\sum_{lm} \sqrt{v_l}\, a_{lm} c_l \phi_l(s) Y_{lm}(x) \sum_{l'm'} \sqrt{v_{l'}}\, a_{l'm'} c_{l'} \phi_{l'}(t) Y_{l'm'}(y) \right] \tag{4.99}$$

$$= \sum_{lm} \sum_{l'm'} \mathbb{E}[a_{lm} a_{l'm'}] \sqrt{v_l v_{l'}} c_l c_{l'} \phi_l(s) \phi_{l'}(t) Y_{lm}(x) Y_{l'm'}(y) \tag{4.100}$$

$$= \sum_{lm} \sum_{l'm'} \delta_{ll'} \delta_{mm'} \sqrt{v_l v_{l'}} c_l c_{l'} \phi_l(s) \phi_{l'}(t) Y_{lm}(x) Y_{l'm'}(y) \tag{4.101}$$

$$= \sum_{l=0}^{\infty} v_l c_l^2 \phi_l(s) \phi_l(t) \sum_{m=0}^{\infty} Y_{lm}(x) Y_{lm}(y) \tag{4.102}$$

$$= \frac{1}{\omega_d} \sum_{l=0}^{\infty} v_l c(l,d) c_l^2 \phi_l(s) \phi_l(t) W_l^{\lambda}(\langle x, y \rangle) \tag{4.103}$$

$$= \sum_{l=0}^{\infty} a_l c_l^2 \phi_l(s) \phi_l(t) W_l^{\lambda}(\langle x, y \rangle) \tag{4.104}$$

using the addition formula for spherical harmonics. Since the field is stationary in time the temporal factor in the covariance depends only on the difference $s - t$, and thus we may take $s = 0$ and deal with t only. After rescaling so $\sum_{n=0}^{\infty} a_n = 1$, this gives (4.98), as required. □

Interest in path properties of geo-temporal Gaussian processes is still developing and the current literature is very sparse – we know of only [Jon63; CAP18] addressing the subject. [CAP18] describe regularity properties of spatio-temporal processes using Sobolev and interpolation spaces, aided by two new spectral expansions for spatio-temporal processes: a double Karhunen-Loève expansion and a decomposition of the $\phi_n(t)$ above into their Hermite expansions $\sum_k b_{n,k} H_k(t)$, with H_k the normalised Hermite polynomial of degree k. Their approach, which is based on Theorem 2.7a, handles space and time asymmetrically, whereas (4.97) above gives a symmetric expansion.

4.7 Complements

4.7.1 Other approaches to continuity for Gaussian processes

A very recent approach to studying path properties of Gaussian processes on the sphere is due to work by Lan, Marinucci and Xiao [LMX18; LX18] obtained in 2018, using their concept of *strong local non-determinism* to determine a.s. Hölder conditions for spherical Gaussian random fields ($d = 2$; 'exact modulus of non-differentiability') and their local times, assuming the Schoenberg coefficients are bounded above and below by powers.

Azmoodeh et al. [Azm+14] give necessary and sufficient conditions for Hölder continuity of Gaussian processes (but without an exact modulus of continuity, as in [LMX18]), for general (rather than spherical) Gaussian processes. Marcus and Rosen [MR06, Ch. 7] give a fairly exhaustive account of the theory of *exact* moduli of continuity for such Gaussian processes.

4.7.2 Measuring smoothness

'How to measure smoothness' is the title of Chapter 1 in the standard work on function spaces, Triebel [Tri83]. The three-parameter families most commonly used for measuring smoothness are the Besov spaces B^s_{pq} described above. For more background here, we refer to [Tri83] (there, the spaces F^s_{pq}, usually called Triebel-Lizorkin spaces, are also treated, but we do not consider these), and for $B^s_{pq}(M)$ to [FFP16]. There are also weighted versions $B^s_{pq}(\rho)$, with weight-function ρ; see [Sch00, §3]

4.7.3 Sobolev spaces

The approach of [LS15] (and extended in [CAP18]) quantifies Gaussian paths in terms of Sobolev spaces $W^m_p(\Omega)$ ($m \in \mathbb{N}, p \in [1, \infty)$, Ω a domain), rather than the (more general) Besov spaces we have used. Sobolev spaces are defined [AF03, Ch. III] as vector subspaces of $L^p(\Omega)$ whose weak (or distributional) partial derivatives of order $\alpha \leq m$ lie in L^p. As in L^p, the elements of W^m_p are equivalence classes of functions (under equality a.e.) rather than individual functions and, again like L^p, the W^m_p are Banach spaces.

Chapter 5

The Ciesielski Isomorphism on Spheres

Let $\alpha \in (0,1)$, and define $\mathrm{Lip}^{\alpha}([0,1]^d)$ to be the space of Hölder continuous functions of order α, with the additional requirement that they vanish at zero. Using the notation $||u|X||$ to denote the norm of u with respect to the space X, let

$$||f \mid \mathrm{Lip}^{\alpha}(\mathbb{S}^2)|| := \sup_{s,t\in\mathbb{S}^2} \frac{|f(s)-f(t)|}{|t-s|^{\alpha}}. \tag{5.1}$$

We then have the following classical result due to Ciesielski [Cie60a; Cie60b], relating a function space to a sequence space.

Theorem 5.1 (Ciesielski's Isomorphism [Sem82, Th. 3.5.10])**.** *Let* $\alpha \in (0,1)$*. Then the space* $\mathrm{Lip}^{\alpha}([0,1]^d)$ *is isomorphic to* ℓ^{∞}*.*

Ciesielski's isomorphism has found myriad applications in probability theory, ranging from stochastic integration [GIP16] to large deviation theory [AIP13]. Our motivation here comes from the use of Ciesielski's result to describe the path properties of a Gaussian process. This question was first addressed in the one-dimensional setting by Ciesielski [Cie61].

[RS98] address the question of regularity of random fields on the d-dimensional interval $[0,1]^d$, using Ciesielski's approach to find a condition for Hölder continuity of a version of a random field X in terms of the coefficients of its Schauder expansion

$$\sum_{j=0}^{\infty}\sum_{v\in V_j}\lambda_{j,v}\Lambda_{j,v} \tag{5.2}$$

where the $\Lambda_{j,v}$ form a Schauder basis of $[0,1]^d$ with respect to a suitable choice of triangulation $V_1 \subset V_2 \subset \ldots$[1]. When the parameter space of the random field is $[0,1]^d$ the coefficients $\lambda_{j,v}$ are second differences of the form $\Delta_h^2 X(t) = X(t+h) + X(t-h) - 2X(t)$, the importance of which for path regularity was first noted by Zygmund [Zyg45]. [RS98] go on to obtain a central limit theorem for sequences of random elements in a Hölder space. A follow-up paper [RS04] answers the same question for Banach-space *valued* random fields.

In this chapter we take the preliminary steps towards expanding this framework to spheres: specifically, we obtain a Ciesielski isomorphism between the space of Hölder-continuous functions of order α on $\mathbb{S}^2$ and the sequence

[1]Recall that a Schauder basis of a Banach space F is a sequence (Λ_n) of functions in F where for each $f \in F$ there exists a unique sequence of scalars (λ_n) such that $||f - \sum_1^N \lambda_n\Lambda_n|| \to 0$ as $N \to \infty$.

space ℓ^∞. The proof of Theorem 5.1 follows from an equivalence between the continuity properties of a function f and the rate of convergence of the coefficients of f's Faber-Schauder decomposition [Sem82, Theorem 3.5.2], and it is this result which is the technical hurdle here. The construction of the Faber-Schauder decomposition hinges on the use of bases of splines of degree one – i.e. functions which are “pyramidal”: continuous and affine on a triangulation of $[0,1]^d$. Such functions cannot exist on the sphere – it is non-Euclidean, so the parallel postulate does not hold. But, the idea can be adapted for our purposes, borrowing from the theory of spherical splines. The result is Theorem 5.2, below.

The modulus of smoothness we can use in Theorem 5.2 is rather general – take any function $\rho : [0, 2\pi] \to \mathbb{R}$ with the following properties:

1. $\rho(0) = 0$, $\rho(\delta) > 0$ for $0 < \delta < 2\pi$;
2. ρ is non-decreasing on $[0, 2\pi]$;
3. $\rho(2\delta) < c_1\rho(\delta)$;
4. $\int_0^\delta \rho(u)/u\, du < c_2\rho(\delta)$;
5. $\delta \int_\delta^{2\pi} \rho(u)/u^2\, du < c_3\rho(\delta)$

for suitable positive constants c_1, c_2, c_3. One example of a function ρ satisfying (1)–(5) is

$$\rho(\delta) := \delta^\alpha \log^\beta(c/\delta), \tag{5.3}$$

with $c \geq 2\pi \exp(\beta/\alpha)$ if $\beta > 0$ and $c > 2\pi \exp(-\beta/(1-\alpha))$ if $\beta < 0$. Throughout this chapter any function denoted ρ is assumed to satisfy (1)–(5).

5.1 Spherical Triangulations

The first step is to define a triangulation on $\mathbb{S}^2$, which we need to find our Schauder basis of $C(\mathbb{S}^2)$. We start with an octahedron projected (gnomonically) onto the sphere – i.e. eight spherical triangles of equal size with six vertices – $(v_0, v_1, v_2, v_3, v_4, v_5)$ given by

$$\begin{aligned} v_0 &= (0,0,1) \\ v_1 &= (1,0,0) \\ v_2 &= (0,1,0) \\ v_3 &= (-1,0,0) \\ v_4 &= (0,-1,0) \\ v_5 &= (0,0,-1). \end{aligned}$$

At each recursion step, subdivide each triangle into four triangles by drawing edges between the mid-points of each existing edge i.e. (wlog the vertices of the existing triangle are (v_0, v_1, v_2)) define new vertices w_0, w_1, w_2 by

$$w_0 = \frac{v_1 + v_2}{|v_1 + v_2|};$$

$$w_1 = \frac{v_0 + v_2}{|v_0 + v_2|};$$
$$w_2 = \frac{v_0 + v_1}{|v_0 + v_1|}.$$

It is not possible to have a completely uniform triangulation of the sphere containing more than twenty spherical triangles (the icosahedron is the Platonic solid with the most faces). Nonetheless, the octahedral triangulation boasts nice enough uniformity properties for our purpose: denoting by T_n the octahedral triangulation a level n, the maximum diameter of $\Delta \in T_n$, d_n, is bounded by

$$d_n \leq A \cdot 2^{-n}, \tag{5.4}$$

where A is a suitable constant [BDS08]. Moreover, it follows from the proof of [BDS08, Theorem 1] that, for all $n \in \mathbb{N}$, for all $\Delta \in T_n$, the area $\alpha(\Delta)$ is bounded below by

$$\alpha(\Delta) \geq B \cdot 2^{-2n}, \tag{5.5}$$

where B is a constant.

Now we need an analogue of a *pyramidal basis* for $C(\mathbb{S}^2)$. In the planar case this basis is formed of splines of degree one (see [Sem82, Ch. 3]), so a natural first pick is to try spherical splines of degree one, slightly modified.

Given a spherical triangle $\Delta = (v_1, v_2, v_3)$ define the *spherical barycentric coordinates* of a point $x \in \mathbb{S}^2$ to be the (unique) triplet $(b_{v_1}(x), b_{v_2}(x), b_{v_3}(x))$

such that

$$x = b_{v_1}(x)v_1 + b_{v_2}(x)v_2 + b_{v_3}(x)v_3. \tag{5.6}$$

Clearly, for $\alpha \in \mathbb{R}$ $b_v(\alpha w) = \alpha b_v(w)$. Moreover $b_{v_i}(v_j) = \delta_{ij}$, $b_v(x) > 0$ for all $x \in \mathrm{Int}(\Delta)$ and if x lies on the edge of Δ joining v_1 and v_2, say, then $b_{v_3}(x) = 0$.

The values of the barycentric co-ordinates are given by ratios of volumes:

$$b_1 = \frac{\mathrm{vol}(t_1)}{\mathrm{vol}(t)}, \tag{5.7}$$

where t is the spherical pyramid with vertices $(0, v_1, v_2, v_3)$ and t_1 the spherical pyramid with vertices $(0, x, v_2, v_3)$. b_2 and b_3 are defined similarly.

The volume of a spherical pyramid $t = (0, v_1, v_2, v_3)$ with face triangle $\Delta = (v_1, v_2, v_3)$ is given by

$$\mathrm{vol}(t) = \mathrm{vol}(\mathbb{S}^2)\frac{\alpha(\Delta)}{\alpha(\mathbb{S}^2)} = \frac{4\pi}{3}\frac{\alpha(\Delta)}{4\pi} = \frac{\alpha(\Delta)}{3}, \tag{5.8}$$

where $\alpha(\cdot)$ denotes surface area. So the b_1 are equivalently ratios of surface areas of spherical triangles.

Now define, for each triangulation T_j and each $v \in T_j$ the *pyramidal* function

$\Lambda_{j,v} : \mathbb{S}^2 \to \mathbb{R}$ by

$$\Lambda_{j,v}(x) := \begin{cases} b_v(x) & x \in \{\Delta : v \in \mathrm{Vert}(\Delta)\}, \\ 0 & \text{else.} \end{cases} \tag{5.9}$$

Then $\Lambda_{j,v}$ has the following properties:

1. $\Lambda_{j,v}(v) = 1$
2. $\Lambda_{j,v}(w) = 0$ for all $w \in \mathrm{Vert}(T_j) \backslash \{v\}$

Let $T_1, T_2, \ldots$ be the sequence of refinements of the octahedral triangulation of the sphere. Let $W_n = \mathrm{vert}(T_n)$, $V_1 = W_1$, $V_{n+1} = W_{n+1} \backslash V_n$, $V = \cup V_n$. Then we can formulate the following definition.

Definition 5.1. *A basis* $\{\Lambda_j : j \in \mathbb{N}\}$ *of* $C(\mathbb{S}^2)$ *is* pyramidal *if and only if*

1. *for every* $j \in \mathbb{N}$ *there exists a* $\nu \in \mathbb{N}$ *such that* Λ_j *is pyramidal with respect to* T_ν *– denote the smallest such* ν *by* $\nu(j)$*;*
2. *if* $j < j'$ *then* $\nu(j) \leq \nu(j')$*;*
3. *each* $v \in V$ *is the peak vertex of precisely one function* Λ_j*;*
4. *if* $v \in W_m$ *and* v *is a peak vertex of* Λ_j *then* $\nu(j) \leq m$*.*

Lemma 5.1. *For the sequential octahedral triangulation of* $\mathbb{S}^2$ *there is a sequence in* $C(\mathbb{S}^2)$ *satisfying (1) – (4).*

Proof. Pick the sequence of functions $\mathcal{L} = \{\Lambda_{j,v} : j \in \mathbb{N}, v \in V_j\}$, where $\Lambda_{j,v}$ is the (unique) function pyramidal wrt T_j with peak vertex v. For ease, order the vertices lexicographically by their co-ordinates projected back onto the octahedron. For each $j \in \mathbb{N}$ (1) is satisfied, with $\nu = j$. (2) and (3) follow by the construction of V_j. For (4), let $v \in W_m$. Then either $v \in V_m$, and $\nu(j) = m$, or $v \in W_{m-1}$ and $\nu(j) < m$. □

We will see that the $\mathcal{L}$ forms a basis for $C(\mathbb{S}^2)$, i.e. any $f \in C(\mathbb{S}^2)$ has a unique, uniformly convergent, expansion

$$f(t) = \sum_{j \in \mathbb{N}} \sum_{v \in V_j} \lambda_{j,v}(f) \Lambda_{j,v}(t). \tag{5.10}$$

In fact, $\mathcal{L}$ is precisely the class $\mathfrak{s}_1$ of spherical splines of degree one, based on Bernstein-Bézier polynomials – see e.g. [ANS96].

Denote by $S^{(n)}$ the partial sum over all vertices in the first n triangulations:

$$S^{(n)} = \sum_{j=1}^{n} \sum_{v \in V_j} \lambda_{j,v} \Lambda_{j,v}. \tag{5.11}$$

Note that (1) and (4) imply that if $v < w$ then $\Lambda_{j,w}(v) = 0$.

Let $w \in W_n$. Then

$$f(w) - S^{(n)}(w) = \sum_{j=n+1}^{\infty} \sum_{v \in V_j} \lambda_{j,v} \Lambda_{j,v}(w) = 0, \tag{5.12}$$

as for all $j > n$ we have $w \in \mathrm{vert}(T_j)$, so $\Lambda_{j,v}(w) = 0$ (as if $w \in W_n$ then

$w \notin V_j$ for $j > n$). Thus, for $v \in W_1$

$$f(v) = S^{(1)}(v) = \sum_{w \in W_1} \lambda_{1,w} \Lambda_{1,w}(v) = \lambda_{1,v}, \tag{5.13}$$

and for $v \in V_n$, $n \geq 2$,

$$f(v) = S^n(v) = S^{(n-1)}(v) + \sum_{w \in V_n} \lambda_{n,w} \Lambda_{n,w}(v) \tag{5.14}$$

$$= S^{(n-1)}(v) + \lambda_{n,v}, \tag{5.15}$$

i.e. the coefficients are given by

$$\lambda_{1,v} = f(v), \qquad v \in W_1, \tag{5.16}$$

$$\lambda_{n,v} = f(v) - S^{(n-1)}(v), \qquad v \in V_n,\ n \geq 2. \tag{5.17}$$

Now we can finally show that $\mathcal{L}$ is indeed a basis for $C(\mathbb{S}^2)$.

Proposition 5.1. *$\mathcal{L} = \{\Lambda_{j,v} : j \in \mathbb{N}, v \in V_j\}$ is a basis of $C(\mathbb{S}^2)$.*

Proof. By [Sem, 1.3.2] $\mathcal{L}$ is a basis of $C(\mathbb{S}^2)$ if (a) $\Lambda_{j,v}(v) \neq 0$, $\Lambda_{j,v}(w) = 0$, $w \in \mathrm{Vert}(T_j) \backslash \{v\}$, which is true by construction, and (b) $\sum \lambda\Lambda$ converges uniformly.

To show this, note that $d_n \to 0$, where $d_n = \sup\{\mathrm{diam}\,\Delta : \Delta \in T_n\}$. Take n sufficiently large. For any $x \in \mathbb{S}^2$ there exists a triangle Δ with vertex v

such that $d(x, v) \leq \operatorname{diam}(\Delta) \leq d_n$, and x. Then,

$$|S^{(n)}(x) - f(x)| = |S^{(n)}(x) - f(v) + f(v) - f(x)| \tag{5.18}$$

$$\leq |S^{(n)}(x) - f(v)| + |f(v) - f(x)|, \tag{5.19}$$

by the triangle inequality. Define

$$\omega_f(\delta) := \sup_{t,s,\in\mathbb{S}^2, d(t,s)\leq\delta} |f(t) - f(s)| \tag{5.20}$$

so $|f(v) - f(x)| \leq \omega_f(d_n)$. Meanwhile, since $S^{(n)}$ is continuous and $f(v) = S^{(n)}(v)$,

$$|S^{(n)}(x) - f(v)| = |S^{(n)}(x) - S^{(n)}(v)| \to 0, \tag{5.21}$$

i.e. $||f - S^{(n)}|| \to 0$. Moreover.

$$|\lambda_j, v| \leq \omega_f(d_{n-1}) + \omega_f(d_n) \leq 2\omega_f(d_{n-1}), \qquad v \in V^n, \tag{5.22}$$

completing the proof. □

5.2 The Ciesielski Isomorphism

Now that we have defined the relevant basis of $C(\mathbb{S}^2)$, we can formulate our main result, from which the Ciesielski isomorphism follows.

Theorem 5.2. *For every $f \in C(\mathbb{S}^2)$ the following are equivalent:*

$$(i) \qquad \omega_f(\delta) = \mathcal{O}\left(\rho(\delta)\right); \tag{5.23}$$

$$(ii) \qquad \sup_{v \in V_n} |\lambda_{n,v}| = \mathcal{O}\left(\rho(2^{-n})\right). \tag{5.24}$$

To prove this, we need a series of lemmas.

Lemma 5.2. *For all $j \in \mathbb{N}$, all $x \in \mathbb{S}^2$, $\sum_{v \in V_j} |\Lambda_{j,v}(x)| \leq 3$.*

Proof. By continuity it suffices to check the result for $x \in \Delta$, some spherical triangle in T_j. If $v \in V_j$ is not a vertex of Δ then $\Lambda_{j,v}(t) = 0$. Otherwise, either 2 or 3 vertices of Δ are in V_j, so

$$\sum_{v \in V_j} |\Lambda_{j,v}(x)| \leq \sum_{v_i \text{ vertices of } \Delta} |\Lambda_{j,v_i}(x)| \leq 3, \tag{5.25}$$

since Λ_{j,v_i} takes value 1 at v_i then decreases to 0 at the other vertices. □

Lemma 5.3. *If $f \in C(\mathbb{S}^2)$ and there exists $k \in \mathbb{N}$ such that*

$$\sup\left\{|\lambda_{n,v}| : v \in V_n\right\} \leq (\rho(2^{-n})) \tag{5.26}$$

for $n \geq k$, then

$$||f - S^{(n-1)}(f)|| \leq 6c_1c_2\rho(2^{-n}), \qquad n \geq k. \tag{5.27}$$

Proof. Let $n \geq k$. Then, using Lemma 5.2, the proof follows [Sem82] closely:

$$
\begin{aligned}
|f(x) - S^{(n-1)}(x)| &= \left| \sum_{j=n}^{\infty} \sum_{v \in V_j} \lambda_{j,v} \Lambda_{j,v}(x) \right| && (5.28)\\
&\leq \sum_{j=n}^{\infty} \sum_{v \in V_j} |\lambda_{j,v}||\Lambda_{j,v}(x)| && (5.29)\\
&\leq \sum_{j=n}^{\infty} 3 \sup_{v \in V_j} |\lambda_{j,v}| \leq \sum_{j=n}^{\infty} 3\rho(2^{-j}) = 6 \sum_{j=n}^{\infty} 2^{-j} \frac{\rho(2^{-j})}{2^{-j+1}} && (5.30)\\
&\leq 6 \sum_{j=n}^{\infty} \int_{2^{-j}}^{2^{-j+1}} \frac{\rho(t)}{t} \,\mathrm{d}t && (5.31)\\
&= 6 \int_0^{2^{-n+1}} \frac{\rho(t)}{t} \,\mathrm{d}t \leq 6c_2 \rho(2^{-n+1}) \leq 6c_1c_2\rho(2^{-n}). && (5.32)
\end{aligned}
$$

□

Lemma 5.4. *There exists a constant c_0 such that for each $n \in \mathbb{N}$ and each pair $x_0, y_0 \in \mathbb{S}^2$ such that $d(x_0, y_0) \leq \pi 2^{-n}$ there exists a sequence $x_0 = z_0, z_1, \ldots, z_k = y_0$ and spherical triangles Δ_j in T, $j = 1, \ldots, k$ such that $k \leq c_0$ and for each j successive points z_{j-1}, z_j lie in the same simplex Δ_j.*

Proof. The proof again follows [Sem82], with the appropriate adjustments for the spherical geometry. Let x_0, y_0 be such that $d(x_0, y_0) \leq \pi 2^{-n}$. Then there is a spherical cap K such that $\operatorname{diam}(K) = \pi 2^{-n}$ and $x_0, y_0 \in K$. Denote by $T_n^K := \{\Delta \in T_n : \Delta \cap K \neq \emptyset\}$ the set of spherical triangles in the triangulation T_n intersecting K. $|T_n^K|$ is then the maximum number of simplexes that the sequence $z_0, \ldots z_k$ can walk through.

Let $\hat{K} := \{z \in \mathbb{S}^2 : d(z, K) \leq A \cdot 2^{-n}\}$. By (5.4) it follows that $\cup\{\Delta : \Delta \in$

$T_n^K\} \subset \hat{K}$. Now, the area of the spherical cap $\hat{K}$ is

$$2\pi\left[1 - \cos\left((A + 2\pi)2^{-n}\right)\right] \leq \pi\left[(A + 2\pi)2^{-n}\right]^2, \tag{5.33}$$

using the bound $1 - \cos x \leq x^2/2$. By (5.5) the area of Δ in T_n is greater than $B \cdot 2^{-2n}$, and thus the number of triangles $|T_n^K|$ in T_n^K is less than

$$c_0 = \frac{\pi(A + 2\pi)^2}{B}, \tag{5.34}$$

which does not depend on n.

The rest follows precisely as in [Sem82, p. 69]: let Γ denote the set of all finite sequences of spherical triangles $\Delta_1, \ldots, \Delta_k$ with $x_0 \in \Delta_1$ and $\Delta_{j-1} \cap \Delta_j \neq \emptyset$ (further, $\Delta_i \neq \Delta_j$ for $i < j$ – otherwise the intermediary triangles could be dropped from the sequence).

For each $\Delta \in T_n^K$ there is a sequence (Δ_j) in Γ with $\Delta = \Delta_j$ for some j. Thus, there is a sequence $\Delta_1, \ldots \Delta_k$ in Γ where $y_0 \in \Delta_k$. So $k \leq c_0$, and we can choose any $z_j \in \Delta_j \cap \Delta_{j+1}$ to fill the sequence. □

We now proceed to the proof of the main theorem

Proof of Theorem 5.2. ($\Rightarrow$): Let $\omega_f(\delta) = \mathcal{O}(\rho(\delta))$. Then since the coefficients

$$\lambda_{n,v} = f(v) - S^{(n-1)}(v) \tag{5.35}$$

are bounded:

$$|\lambda_{n,v}| \leq \omega_f(d_{n-1}) \leq \omega_f(2^{-(n+1)}) \leq \omega_f(2^{-n}), \tag{5.36}$$

we have that $\sup_{v\in V_n} |\lambda_{n,v}| = \mathcal{O}(\rho(2^{-n}))$.

($\Leftarrow$): Now suppose

$$\sup_{v\in V_n} |\lambda_{n,v}| = \mathcal{O}(\rho(2^{-n}). \tag{5.37}$$

Then by Lemma 5.3,

$$||f - S^{(n-1)}(f)|| \leq 6c_1c_2\rho(2^{-n}). \tag{5.38}$$

First, following Semadeni [Sem82, p. 71], estimate $|f(x) - f(y)|$ when x, y belong to the same triangle Δ in T_n for some sufficiently large n. Then, by (5.4) $d(x, y) \leq A\cdot 2^{-n} < \pi/4$. For each $k \leq n$ there exists a (unique) $\Delta_k \in T_k$ such that $\Delta \subset \Delta_k$.

Let v be a vertex of Δ_k, and denote the remaining two vertices by v_2, v_3. Let $\Delta_x = (x, v_2, v_3)$ and $\Delta_y = (y, v_2, v_3)$ be the triangles with vertices x, v_2, v_3 and y, v_2, v_3 respectively. Since by (5.7) the barycentric co-ordinate $b_v(\cdot)$ are ratios of the volumes of $\Delta_{(\cdot)}$ and Δ_k, using (5.8) we have

$$\begin{aligned} |\Lambda_{k,v}(x) - \Lambda_{k,v}(y)| &= | b_v(x) - b_v(y) | && (5.39)\\ &= \left| \frac{\alpha(\Delta_x)}{\alpha(\Delta_k)} - \frac{\alpha(\Delta_y)}{\alpha(\Delta_k)} \right| = \frac{1}{|\alpha(\Delta_k)|}|\alpha(\Delta_x) - \alpha(\Delta_y)|. && (5.40)\end{aligned}$$

By (5.5) $\alpha(\Delta_k) \geq B \cdot 2^{-2k}$ so

$$|\Lambda_{k,v}(x) - \Lambda_{k,v}(y)| \leq \frac{2^{2k}}{B} |\alpha(\Delta_x) - \alpha(\Delta_y)|. \tag{5.41}$$

The area $|\alpha(\Delta_x) - \alpha(\Delta_y)|$ is the area between the two triangles Δ_x and Δ_y and is enclosed in a spherical rectangle of area $2 \cdot d(x, y) \cdot \operatorname{diam}(\Delta_k)$. The area of this spherical rectangle $\square$ is given by

$$\alpha(\square) = 4 \arcsin\left(\tan\left(\frac{2d(x,y)}{2}\right)\tan\left(\frac{\operatorname{diam}(\Delta_k)}{2}\right)\right). \tag{5.42}$$

Noting that $k \geq 1$ $\operatorname{diam}(\Delta_k) \leq \pi/2$, so $\tan(\operatorname{diam}(\Delta_k)/2) \leq 1$, and since arcsin is increasing and $\arcsin(\tan(x)) \leq 2x$ for $x \in [0, \pi/4]$, we have

$$\alpha(\square) \leq 4 \arcsin\left(\tan\left(d(x,y)\right)\right) \leq 8d(x,y) \leq 8A \cdot 2^{-n}. \tag{5.43}$$

Putting this together,

$$\left|\Lambda_{k,v}(x) - \Lambda_{k,v}(y)\right| \leq \frac{2A2^{2k-n}}{B}. \tag{5.44}$$

Now,

$$|f(x) - f(y)| \leq \sum_{k=0}^{n-1} \sum_{v \in V_k} |\lambda_{k,v}||\Lambda_{k,v}(x) - \Lambda_{k,v}(y)| + 2||f - S^{(n-1)}(f)|| \tag{5.45}$$

$$\leq \sum_{k=1}^{n-1} \rho(2^{-k}) \sum_{v \in \mathrm{Vert}(\Delta_k)} |\Lambda_{k,v}(x) - \Lambda_{k,v}(y)| + 12c_1c_2\rho(2^{-n}) \tag{5.46}$$

$$\leq \sum_{k=1}^{n-1} \rho(2^{-k}) \frac{2A2^{2k-n}}{B} + 12c_1c_2\rho(2^{-n}) \tag{5.47}$$

$$\leq C2^{-n} \sum_{k=1}^{n} 2^{k-1}\rho(2^{-k}). \tag{5.48}$$

$$\tag{5.49}$$

Now consider the general case of $x, y \in \mathbb{S}^2$, $d(x,y) < \delta < \pi/4$. Then there exists a $n \geq 2$ such that $2^{-(n+1)}\pi \leq d(x,y) \leq 2^{-n}\pi$. So, by Lemma 5.4. and the above,

$$|f(x) - f(y)| \leq \sum_{l=1}^{k} |f(z_l) - f(z_{l-1})| \tag{5.50}$$

$$\leq C'2^{-n} \sum_{k=1}^{n} 2^{k-1}\rho(2^{-k}) \tag{5.51}$$

$$\leq C'2^{-n} \left[\sum_{k=1}^{2^n-1} \int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{\rho(t)}{t^2}\,\mathrm{d}t \right] \tag{5.52}$$

$$\leq C'2^{-n} \left[\int_{2^{-n}}^{1} \frac{\rho(t)}{t^2}\,\mathrm{d}t \right] \tag{5.53}$$

$$\leq C' \frac{d(x,y)}{\pi} \left[\int_{2^{-n}}^{1} \frac{\rho(t)}{t^2}\,\mathrm{d}t \right] \tag{5.54}$$

$$\leq C' \frac{d(x,y)}{\pi} \left[\int_{\frac{d(x,y)}{\pi}}^{1} \frac{\rho(t)}{t^2}\,\mathrm{d}t \right]. \tag{5.55}$$

Recall properties (2) and (5) of ρ:

$$2.\ \rho \text{ is non-decreasing;} \tag{5.56}$$

$$5.\ \delta \int_{\delta}^{2\pi} \frac{\rho(u)}{u^2}\,\mathrm{d}u < c_3\rho(\delta). \tag{5.57}$$

Since ρ is positive, $\int_\delta^1 \rho(u)/u^2 \leq \int_\delta^{2\pi} \rho(u)/u^2$, so (allowing the constant C' to vary as necessary),

$$|f(x) - f(y)| \le C'\rho\left(\frac{d(x,y)}{\pi}\right) \tag{5.58}$$

$$\le C'\rho(d(x,y)) \le C'\rho(\delta). \tag{5.59}$$

□

Now, letting $\mathrm{Lip}^\alpha(\mathbb{S}^2)$ be the space of Hölder continuous functions f of order α with norm

$$||f \,|\, \mathrm{Lip}^\alpha(\mathbb{S}^2)|| := \sup_{s,t\in\mathbb{S}^2} \frac{|f(s) - f(t)|}{d(s,t)^\alpha} \tag{5.60}$$

we have the following spherical version of Theorem 5.1.

Corollary 5.1. *Let $\alpha \in (0,1)$. Then $\mathrm{Lip}^\alpha(\mathbb{S}^2)$ is isomorphic to ℓ^∞.*

Proof. We need to find a linear, bounded, operator $u_\alpha : \ell^\infty \to \mathrm{Lip}^\alpha(\mathbb{S}^2)$. Let $a = (a_n)$ be a sequence in ℓ^∞, fix $t \in \mathbb{S}^2$, and consider the series

$$\sum_{n=1}^{\infty} a_n b_n \Lambda_n(t) = \sum_{j=1}^{\infty} \sum_{v\in V_j} a_{j,v} b_j \Lambda_{j,v}(t) \tag{5.61}$$

where $b_j = 2^{-j\alpha}$. Then, for fixed j,

$$\sum_{v\in V_j} |a_{j,v} b_j \Lambda_{j,v}(t)| \le ||a||_\infty 2^{-j\alpha} \sum_{v\in V_j} \Lambda_{j,v}(t) \tag{5.62}$$

$$\le ||a||_\infty 2^{-j\alpha}, \tag{5.63}$$

and since $\sum 2^{-j\alpha} < \infty$ the series (5.60) converges to, say, $f \in C(\mathbb{S}^2)$. Moreover,

$$||f||_\infty \leq ||a||_\infty \sum_j 2^{-j\alpha} \tag{5.64}$$

$$= ||a||_\infty \frac{1}{1 - 2^{-\alpha}}, \tag{5.65}$$

so f is bounded, and $f \in \mathrm{Lip}^\alpha(\mathbb{S}^2)$.

Let $\rho(\delta) := \delta^\alpha$. Applying Theorem 5.2, we see that $\omega_f(\delta) \leq C\delta^\alpha$ for a constant C depending only on $||a||_\infty$ (crucially, not on the sequence itself). Thus, $u_\alpha(a) := f$ is a bounded, linear operator from ℓ^∞ to $\mathrm{Lip}^\alpha(\mathbb{S}^2)$.

The result then follows using the Open Mapping Theorem [Rud91, Th. 2.11] – if T is a bounded linear operator from X to Y then $X \approx Y$ if and only if T is bijective. To see that the mapping $u_\alpha : \ell^\infty \to \mathrm{Lip}^\alpha(\mathbb{S}^2)$ is bijective, we construct its inverse.

Let $f \in \mathrm{Lip}^\alpha(\mathbb{S}^2)$ be given by $f(t) = \sum_{n=0}^\infty \sum_{v \in V_n} \lambda_{n,v} \Lambda_{n,v}(t)$ and note that, for fixed n, by Theorem 5.2 and the definition of $\mathrm{Lip}^\alpha(\mathbb{S}^2)$, $\sup_v |\lambda_{n,v}| = \mathcal{O}(2^{-n\alpha})$. Then the sequence $(a_{n,v})$ defined by $a_{n,v} := 2^{n\alpha}\lambda_{n,v}$ is such that

$$|a_{n,v}| \leq 2^{n\alpha} \sup_{v \in V_n} |\lambda_{n,v}| \leq C 2^{n\alpha} 2^{-n\alpha} \leq C, \tag{5.66}$$

i.e. $(a_n) = (a_{n,v})$ is a bounded sequence and the mapping $f \mapsto (a_n)$ is the required inverse $u_\alpha^{-1} : Lip^\alpha(\mathbb{S}^2) \to \ell^\infty$. $\square$

Chapter 6

Summary and future directions

6.1 Summary

In this thesis we have explored several areas of the theory of Gaussian random fields on the sphere, ranging from the prototypical example of Brownian motion in Chapter 1 to their path properties in Chapter 4.

In the geo-temporal setting we have proved a new formulation of the characterisation theorem (Theorem 2.7b) for functions positive definite on $\mathbb{S}^d \times \mathbb{R}$, before in Chapter 3 verifying that Beatson and zu Castell's [BC17; BC16] operators for functions on $\mathbb{S}^d$ extend to the geo-temporal case. The paths of processes of $\mathbb{S}^d \times \mathbb{R}$ are not yet well-studied: we showed in Chapter 4.6 that our formulation of the Berg-Porcu theorem allows for a very symmetric Karhunen-Loève expansion compared to those proposed in [CAP18].

We have also made some contributions to the purely spatial theory. In Chapter 3 we proposed a new method for dimension hopping, based on spherical fractional integration with the semi-group property. After surveying the current state of the theory of the path properties of Gaussian processes on $\mathbb{S}^d$ in Chapter 4, we showed that Schilling's neat approach to the Kolmogorov-Chentsov theorem [Sch00] may be applied on the sphere, streamlining work by Lang & Schwab [LS15]. We also extended a result of Malyrenko's to the case of fields on the Hilbert sphere $\mathbb{S}^\infty$. Finally, in Chapter 5 we obtained a Ciesielski isomorphism for functions on the 2-sphere, which we hope will have interesting applications.

6.2 Current and Future Directions

The work in this thesis is ongoing, and we have much more to say in the future. Below we summarise some of our main strands of thinking.

6.2.1 Applications of the Ciesielski Isomorphism

Theorem 5.2 is the equivalent result on the 2-sphere to a theorem of Ciesielski [Cie60a] (generalised slightly by Rhyll, we used the exposition in Semadeni [Sem82, Ch. 3]). In [RS98], Račkauskas and Suquet use this to find a sufficient condition for Hölder continuity of random processes on $[0, 1]^d$. In particular, they find the following.

Theorem 6.1 (Račkauskas and Suquet [RS98, Theorem 10]). *Let $p \geq 1$ and σ is an increasing function on $\mathbb{R}$. If a continuous process $\{X(t) : t \in [0,1]^d\}$ satisfies, for each $t \in [0,1]^d$ and each $h = (h_1, \ldots, h_d)$ where $0 \leq h_i \leq \min(t_i, 1-t_i)$,*

$$\mathbb{P}\left\{|\Delta_h^2 X(t)| > \lambda\right\} \leq \frac{C}{\lambda^p}\sigma^p(|h|), \tag{6.1}$$

where

$$\int_0^1 \frac{\sigma(u)}{u^{1+d/p}\rho(u)} du < \infty, \tag{6.2}$$

then X has a version with almost all paths Hölder continuous with modulus of continuity ρ.

Using the results of Chapter 5 to find analogous statements to the above for processes parametrised by the sphere is the focus of my current work. It would also be of interest to find a d-dimensional version of Theorem 5.2.

6.2.2 Path properties of geo-temporal processes

As remarked in Chapter 4, although the literature on path properties of random fields on spheres is vast, geo-temporal processes are much less studied (unsurprisingly, as they were first classified in 2017). The questions here must be delicately posed: are we interested in smoothness in the spatial component, the temporal one, or both? Recent work by Kühn and Schilling in the context of spatio-temporal Lévy processes [KS19], discussed by Schilling at a London Probability Seminar, indicates that in that scenario there is a trade-off between smoothness in space or smoothness in time. I intend to explore whether similar behaviour exists for geo-temporal processes.

6.2.3 The SPDE approach

Throughout this thesis we have taken a very probabilistic view of our subject, focussing on Gaussian processes through the behaviour of their covariance functions. This, as we have seen, takes us a long way theoretically. In practice, it takes us less far – fully specifying the covariance of a geo-temporal dataset is prohibitively expensive [PAF18]. Moreover, one of the most popular parametric families of covariances, the Matérn covariance

$$M_\nu(x) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} x^\nu \mathcal{K}_\nu(x), \qquad x \geq 0, \tag{6.3}$$

with $\mathcal{K}_\nu$ a modified Bessel function, is not positive definite when $x = d(s,t)$, geodesic distance on $\mathbb{S}^2$ [Gne13]. An ingenious approach by Lindgren et al. [LRL11] works around this limitation. Their idea is to model the field as the solution to the SPDE

$$(\kappa^2 - \Delta)^\alpha X(s) = \mathcal{W}(s), \qquad s \in \mathbb{S}^2, \tag{6.4}$$

where $\kappa > 0$, $\alpha > 0$, Δ is the Laplace-Beltrami operator and $\mathcal{W}$ is a white-noise process. The solution X is a Gaussian process on $\mathbb{S}^2$ with Matérn covariance (the parameter ν in the Matérn covariance function is linked to the parameter α in the SPDE), found without specifying the covariance explicitly using a Gaussian Markov field approximation of the full process. This SPDE model can be paired with a hierarchial model to create a spatio-temporal random field [Cam+13].

This approach is used by a variety of practitioners, including by the team at the Malaria Atlas Project to model the global endemicity of malaria-causing species of the *Plasmodium* parasite [Bat+19]. I am joining this group as a postdoctoral scientist in September 2019.

Bibliography

[AF03] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*. Academic Press, 2003. (1st ed., R. A. Adams, 1975).

[Adl90] R. J. Adler. *An introduction to continuity, extrema and related topics for general Gaussian processes*. Vol. 12. Inst. Math. Statist. Lecture Notes–Monographs. IMS, 1990.

[AT07] R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer, 2007.

[AP17] A. Alegria and E. Porcu. The dimple problem related to space-time modelling under a Lagrangian framework. *J. Multivariate Analysis* 162 (2017), 110–121.

[Ale+19] A. Alegria et al. Covariance functions for multivariate Gaussian fields evolving temporally over Planet Earth. *Stochastic Environmental Research and Risk Assessment* (2019), pp. 1–16.

[ANS96] P. Alfeld, M. Neamtu, and L. L. Schumaker. Bernstein-Bézier polynomials on spheres and sphere-like surfaces. *Computer Aided Geometric Design* 13.(4) (1996), 333 –349.

[AL14] R. Andreev and A. Lang. Kolmogorov-Chentsov theorem and differentiability of random fields on manifolds. *Potential Analysis* 41 (2014), 761 –769.

[AIP13] A. Andresen, P. Imkeller, and N. Perkowski. Large deviations for Hilbert-space-valued Wiener processes: a sequence space approach. *Malliavin calculus and stochastic analysis.* Vol. 34. Springer Proc. Math. Stat. Springer, 2013, p. 115–138.

[AAR99] G. E. Andrews, R. Askey, and R. Roy. *Special Functions.* Cambridge University Press, 1999.

[AB76] R. Askey and N. H. Bingham. Gaussian processes on compact symmetric spaces. *Z. Wahrschein.* 37 (1976), 127 –143.

[AF69] R. Askey and J. Fitch. Integral representations for Jacobi polynomials, and some applications. *J. Math. Anal. Appl.* 25 (1969), 411–437.

[AW65] R. Askey and S. Wainger. On the behaviour of special classes of ultraspherical polynomials I & II. *J. Anal. Math.* 15 (1965), 193 –244 and 245 –262.

[Azm+14] E. Azmoodeh et al. Necessary and sufficient conditions for Hölder continuity of Gaussian processes. *Statist. Probab. Letters* 94 (2014), 230–235.

[Bat+19] K. E. Battle et al. Mapping the global endemicity and clinical burden of Plasmodium vivax, 2000–17: a spatial and temporal modelling study. *The Lancet* 394.(10195) (2019), 332 –343.

[BCX14] R. K. Beatson, W. zu Castell, and Y. Xa. A Pólya criterion for (strict) positive-definiteness on the sphere. *IMA J. Numer. Anal.* 34 (2014), 550–568.

[BC16] R. K. Beatson and W. Zu Castell. One-step recurrences for stationary random fields on the sphere. *SIGMA* 12 (2016), 19p.

[BC17] R. K. Beatson and W. Zu Castell. Dimension hopping and families of strictly positive definite zonal basis functions on spheres. *J. Approximation Theory* 221 (2017), 22 –37.

[BHV08] B. Bekka, P. de la Harpe, and A. Valette. *Kazhdan's property (T).* Vol. 11. New Mathematical Monographs. Cambridge University Press, 2008.

[Bel61] Y. K. Belyaev. Continuity and Hölder's conditions for sample functions of stationary Gaussian processes. *Proc. Fourth Berkeley Symp. Math. Stat. Probab. Vol II: Contributions to Probability Theory.* Ed. by J. Neyman. U. California Press, 1961, p. 23–33.

[BPP18] C. Berg, A. P. Peron, and E. Porcu. Schoenberg's theorem for real and complex Hilbert spheres revisited. *J. Approximation Theory* 228 (2018), 58 –78.

[BP17] C. Berg and E. Porcu. From Schoenberg coefficients to Schoenberg functions. *Constructive Approximation* 45 (2017), 217 – 241.

[Ber74] S. M. Berman. A Gaussian paradox: determinism and discontinuity of sample paths. *Ann. Prob.* 2 (1974), 950–953.

[Ber80] S. M. Berman. Isotropic Gaussian processes on the Hilbert sphere. *Ann. Prob.* 6 (1980), 1093–1106.

[Bin72a] N. H. Bingham. Integral representations for ultraspherical polynomials. *J. London Math. Soc.* 6 (1972), 1 –11.

[Bin72b] N. H. Bingham. Random walks on spheres. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 22 (1972), 169 –192.

[Bin72c] N. H. Bingham. Tauberian theorems for integral transforms of Hankel type. *J. London Math. Soc* 5 (1972), 493 –503.

[Bin73] N. H. Bingham. Positive definite functions on spheres. *Proc. Cambridge Phil. Soc.* 73 (1973), 145–156.

[Bin78] N. H. Bingham. Tauberian theorems for Jacobi series. *Proc. London Math. Soc.* 36 (1978), 285 –309.

[Bin79] N. H. Bingham. Integrability theorems for Jacobi series. *Publ. Inst. Math. Belgrade* 26 (1979), 45–56.

[BGT87] N. H. Bingham, C. M. Goldie, and J. Teugels. *Regular variation.* Cambridge University Press, 1987.

[BMS16] N. H. Bingham, Aleksandar Mijatović, and Tasmin L. Symons. Brownian manifolds, negative type and geo-temporal covariances. *Communications on Stochastic Analysis* 10 (2016), 421 –432.

[BS19] N. H. Bingham and Tasmin L. Symons. Dimension walks on the sphere cross line. *Statist. Probab. Lett.* 147 (2019), 12 –17.

[BH94] W. R. Bloom and H. Heyer. *Harmonic analysis of probability measures on hypergroups.* Vol. 20. De Gruyter Studies in Math. Walter de Gruyter, 1994.

[Blu70] L. M. Blumenthal. *Theory and applications of distance geometry.* 2nd. Chelsea, New York, 1970.

[BDS08] N. Boal, V. Domingue, and F.J. Sayas. Asymptotic Properties of some triangulations of the sphere. *J. Comp. Appl. Math.* 211 (2008), 11–22.

[Boc33] S. Bochner. Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse. *Math. Annalen* 108 (1933), 378 –410.

[BC49] S. Bochner and K. Chandrasekharan. *Fourier transforms.* Annals of Mathematics Studies 19. Princeton University Press, 1949.

[Cam+13] M. Cameletti et al. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Advances in Statistical Analysis* 97 (2013), 109–131.

[CAP18] J. Clarke de la Cerda, A. Alegria, and E. Porcu. Regularity properties and simulations of Gaussian random fields on the sphere cross time. *Electronic J. Stat.* 12 (2018), 399–426.

[CMS03] D. Chen, V. A. Menegatto, and X. Sun. A necessary and sufficient coefficient for strictly positive definite functions on spheres. *Proc. Amer. Math. Soc.* 131.(9) (2003), 2733–2740.

[Che57] N. N. Chentsov. Lévy Brownian motion for several parameters and generalised white noise. *Th. Prob. Appl.* 2 (1957), 265–266.

[Cie60a] Z. Ciesielski. On the isomorphisms of the spaces H_α and m. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* 8 (1960), 217–222.

[Cie60b] Z. Ciesielski. Some properties of Schauder basis of the space $C_{\langle 0,1\rangle}$. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* 8 (1960), 141–144.

[Cie61] Z. Ciesielski. Hölder conditions for realizations of Gaussian processes. *Trans. Amer. Math. Soc.* 99 (1961), 403–413.

[Cre93] N. Cressie. *Statistics for Spatial Data.* Wiley, 1993.

[DL97] M. M. Deza and M. Laurent. *Geometry of cuts and metrics.* Springer, 1997.

[Dud67] R. M. Dudley. The sizes of compact subsets in Hilbert space and continuity of Gaussian processes. *J. Func. Anal.* 1 (1967), 290–330.

[Dud73] R. M. Dudley. Sample functions of the Gaussian process. *Ann. Prob.* 1 (1973), 66–103.

[Dud90] R. M. Dudley. *Real analysis and probability.* Wadsworth & Brooks / Cole, 1990. (Revised, CUP, 2002.)

[Erd+81] A. Erdélyi et al. *Higher transcendental functions, Vol, I-III.* Krieger, Melbourne, 1981. 1st ed., McGraw-Hill, 1953 (Vol. I, II), 1955 (Vol. III).

[Far73] J. Faraut. *Fonction brownienne sur une variété Riemannienne.* Vol. 321. *Sém. Prob. VII, Lecture Notes in Math.* Springer, 1973, pp. 61 –76.

[Far08] J. Faraut. *Analysis on Lie groups.* Vol. 110. Cambridge Studies in Advanced Math. Cambridge University Press, 2008.

[Far12] J. Faraut. Asymptotics of spherical functions for large rank: An introduction. *Representation theory, complex analysis and integral geometry.* Ed. by B. Krötz, O. Offen, and E. Sayag. Birkhäuser/Springer, 2012, p. 251–275.

[FH74] J. Faraut and K. Harzallah. Distances hilbertiennes invariantes sur un espace homogène. *Ann. Inst. Fourier* 24 (1974), 171–283.

[FFP16] H. G. Feichtinger, H. Führ, and I. Pesenson. Geometric space-frequency analysis on manifolds. *J. Fourier Anal. Appl.* 22 (2016), 1294–1355.

[FOT11] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes.* W. de Gruyter, 2011.

[Gan67] R. Gangolli. Positive definite kernels on homogeneous spaces and certain stochastic processes related to Lévy's Brownian motion of several parameters. *Ann. Inst. H. Poincaré B (NS)* 3 (1967), 121–226.

[Gar72] A. M. Garsia. Continuity properties of Gaussian processes with multidimensional time parameter. *Proc. Sixth Berkeley Symp. Math. Stat. Probab. Volume II: Probability Theory.* Ed. by J.

Neyman L. LeCam and E. M. Scott. U. California Press, 1972, p. 369–374.

[GC99] G. Gaspari and S. E. Cohn. Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Met. Soc.* 135 (1999), 723 –757.

[Gel14] Z. A. Gelbaum. Fractional Brownian fields over manifolds. *Trans. Amer. Math. Soc.* 366 (2014), 4781–4814.

[GS16] A. E. Gelfand and E. M. Schliep. Spatial statistics and Gaussian processes: A beautiful marriage. *Spatial Statistics* 18 (2016). Spatial Statistics Avignon: Emerging Patterns, 86 –104.

[GN43] I. M. Gelfand and M. A. Naimark. On the embedding of normed rings into the ring of operators in Hilbert space (Russian). *Mat. Sbornik NS* 12 (1943), 197–213.

[GN16] E. Giné and R. Nickl. *Mathematical foundations of infinite-dimensional statistical models.* Cambridge University Press, 2016.

[Gne99] T. Gneiting. Correlation functions for atmospheric data analysis. *Quart. J. Roy. Meteorol. Soc.* 125 (1999), 2449–2464.

[Gne01] T. Gneiting. Criteria of Pólya type for radial positive definite functions. *Proc. Amer. Math. Soc.* 129 (2001), 2309 –2318.

[Gne02a] T. Gneiting. Compacted supported correlation functions. *J. Multi. Anal.* 83 (2002), 493 –508.

[Gne02b] T. Gneiting. Non-separable, stationary covariance functions for space-time data. *J. Amer. Stat. Assoc.* (2002), 590–600.

[Gne13] T. Gneiting. Strictly and non-strictly positive definite functions on spheres. *Bernoulli* 19 (2013), 1327 –1349.

[GGG07] T. Gneiting, M. G. Genton, and P. Guttorp. Geostatistical space-time models: stationarity, separability and full symmetry. *Statistics of spatio-temporal systems*. Ed. by B. Finkenstädt, L. Held, and V. Isham. Chapman and Hall/CRC, 2007, p. 151 –175.

[GIP16] M. Gubinelli, P. Imkeller, and N. Perkowski. A Fourier analytic approach to pathwise stochastic integration. *Electron. J. Probab.* 21 (2016), Paper No. 2, 37 pages.

[GMP16] J. C. Guella, V. A. Menegatto, and A. P. Peron. An extension of a theorem of Schoenberg to products of spheres. *Banach J. Math. Analysis* 10 (2016), 671–685.

[Han95] Y. Han. The embedding theorem for the Besov and Triebel-Lizorkin spaces on spaces on homogeneous type. *Proc. Amer. Math. Soc.* 123 (1995), 2181–2189.

[Har49] G. H. Hardy. *Divergent Series*. Oxford University Press, 1949.

[Hel62] S. Helgason. *Differential geometry and symmetric spaces*. Academic Press, 1962.

[HKM02] P. G. Hjorth, S. L. Kokkendorff, and S. Markvorsen. Hyperbolic spaces are of strictly negative type. *Proc. Amer. Math. Soc.* 130 (2002), 175–181.

[Jäg19] J. Jäger. A note on the derivatives of isotropic positive definite functions on the Hilbert sphere. `arXiv:1905.08655`. 2019.

[JM74] N. C. Jain and M. B. Marcus. Sufficient conditions for the continuity of stationary Gaussian processes and applications to random series of functions. *Annales de l'Institut Fourier* 24.(2) (1974), 117–141.

[Jon63] R. H. Jones. Stochastic processes on a sphere. *Ann. Math. Stat.* 34 (1963), 213–218.

[Kal97] O. Kallenberg. *Foundations of Modern Probability.* Springer, 1997.

[Ker+18] G. Kerkyacharian et al. Regularity of Gaussian processes on Dirichlet spaces. *Constr. Approx.* 47 (2018), 277–320.

[Kin63] J. F. C. Kingman. Random walks with spherical symmetry. *Acta. Math.* 109 (1963), 11 –53.

[Kos69] B. Kostant. On the existence and irreducibility of certain series of representations. *Bull. Amer. Math. Soc.* 75 (1969), 627–642.

[KS19] F. Kühn and R. Schilling. Strong convergence of the Euler–Maruyama approximation for a class of Lévy-driven SDEs. *Stochastic Process. Appl.* 129 (2019), 2654–2680.

[LX18] X. Lan and Y. Xiao. Hölder Conditions of Local Times and Exact Moduli of non-differentiability for Spherical Gaussian fields. `arXiv:1807.03706`. 2018.

[LMX18] Xiaohong Lan, Domenico Marinucci, and Yimin Xiao. Strong local nondeterminism and exact modulus of continuity for spherical Gaussian fields. *Stoch. Proc. Appl.* 128 (2018), 1294–1315.

[LS15] A. Lang and C. Schwab. Isotropic Gaussian random fields on the sphere: regularity, fast simulation and stochastic partial differential equations. *Ann. Appl. Prob.* 25 (2015), 3047–3094.

[Lee18] J. M. Lee. *Introduction to Riemannian manifolds*. Second edition. Springer, 2018.

[Lév48] P. Lévy. *Processus stochastiques et mouvement brownien*. Gauthier-Villars, Paris, 1948. 2nd ed. 1965.

[Lév59] P. Lévy. Le mouvement brownien fonction d'un point sur la sphère de Riemann. *Rend. Circ. Mat. Palermo* 8 (1959), 297–310.

[Lév66] P. Lévy. Fonctions browniennes dans l'espace Euclidien et dans l'espace de Hilbert. *Research Papers in Statistics* (J. Neyman Festschrift). Ed. by F. N. David. Wiley, 1966, p. 189–223.

[Lif95] M. A. Liftshits. *Gaussian random functions*. Vol. 332. Math. Appl. Kluwer Academic, 1995.

[LRL11] F. Lindgren, H. Rue, and J. Lindstroem. An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc.: Series B* 73 (2011), 423–498.

[Luk70] E. Lukacs. *Characteristic functions*. Second edition, revised and enlarged. Hafner Publishing Co., 1970.

[Lyo13] R. Lyons. Distance covariance in metric spaces. *Annals of Probability* 41 (2013), 3284–3305.

[Lyo14] R. Lyons. Hyperbolic space has strong negative type. *Illinois J. Math.* 58 (2014), 1009–1013.

[Ma17] Chunsheng Ma. Time-varying isotropic vector random fields on spheres. *J. Theoretical Probability* 30 (2017), 1763–1785.

[Mal05] A. A. Malyarenko. Abelian and Tauberian theorems for random fields on two-point homogeneous spaces. *Th. Prob. Math. Stat.* 69 (2005), 115–127.

[Mal13] A. A. Malyarenko. *Invariant random fields on spaces with a group action*. Springer, 2013.

[MP81] M. B. Marcus and G. Pisier. *Random Fourier Series with Applications to Harmonic Analysis*. Princeton University Press, 1981.

[MR06] M. B. Marcus and J. Rosen. *Markov Processes, Gaussian processes and local times*. Cambridge University Press, 2006.

[MS70] M. B. Marcus and L. A. Shepp. Continuity of Gaussian processes. *Trans. Amer. Math. Soc.* 151 (1970), 377–392.

[MS72] M. B. Marcus and L. A. Shepp. Sample behaviour of Gaussian processes. *Proc. Sixth Berkeley Symp. Math. Stat. Probab. Volume II: Probability Theory*. U. California Press, 1972, p. 423–441.

[MP11] D. Marinucci and G. Peccati. *Random fields on the sphere. Representation, limit theorems and cosmological applications*. London Math. Soc. Lecture Note Series 389, Cambridge University Press, 2011.

[MPP17] E. Massa, A. Peron, and E. Porcu. Positive definite functions on complex spheres and their walks through dimensions. *SIGMA* 13 (2017), 088.

[Mat70] G. Matheron. La théorie des variables régionalisées et ses applications. *Les Cahiers du Centre du Morphologie Mathématique de Fontainebleu* (5) (1970).

[McK63] H. P. McKean. Brownian motion with several-dimensional time. *Th. Prob. Appl.* 8 (1963), 357–378.

[Pól18] G. Pólya. Über die Nullstellen gewisser ganzer Funktionen. *Math. Zeitschr.* 2 (1918), 352–383.

[Pól49] G. Pólya. *Remarks on characteristic functions.* Proc. Berkeley Symp. Math. Stat. Probab. U. California Press, 1949, pp. 115–123.

[PAF18] E. Porcu, A. Alegria, and R. Furrer. Modeling temporally evolving and spatially globally dependent data. *Int. Stat. Review* 86 (2018), 344–377.

[Pyt+19] A. Python et al. A Bayesian approach to modelling subnational spatial dynamics of worldwide non-state terrorism, 2010–2016. *J. Roy. Stat. Soc. A* 182.(1) (2019), 323–344.

[RS98] A. Račauskas and C. Suquet. Random fields and central limit theorems in some generalized Hölder spaces. *Prob. Theory and Math. Statist. Proc. 7th Vilnius Conference.* Ed. by B. Grigelionis et al. TEV Vilnius and VSP Utrecht, 1998, p. 599–616.

[RS04] A. Račkauskas and C. Suquet. Central limit theorems in Hölder topologies for Banach space valued random fields. *Th. Probab. Appl.* 49.(1) (2004), 109–125.

[Rud87] W. Rudin. *Real and complex analysis.* 3rd edition. Mc-Graw-Hill, 1987. (1st ed. 1966, 2nd ed. 1974).

[Rud91] W. Rudin. *Functional analysis.* 2nd edition. Mc-Graw-Hill, 1991. (1st ed. 1973).

[Sch00] R. Schilling. Sobolev embedding for stochastic processes. *Expositiones Mathematicae* 18 (2000), 239–242.

[Sch37] I. J. Schoenberg. On certain metric spaces arising from Euclidean spaces by a change of metric, and their embedding in Hilbert space. *Ann. Math.* 38 (1937), 787–793.

[Sch38] I. J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.* 44 (1938), 522–536.

[Sch42] I. J. Schoenberg. Positive definite functions on spheres. *Duke Math. J.* 9 (1942), 96 –108. (Selected Papers I, 172 – 185).

[Seg47] I. E. Segal. The group algebra of a locally compact group. *Trans. Amer. Math. Soc.* 61 (1947), 69–105.

[Sem82] Z. Semadeni. *Schauder bases in Banach spaces of continuous functions.* Vol. 918. Lecture Notes in Math. Springer, 1982.

[Sze39] G. Szegő. *Orthogonal polynomials.* American Mathematical Society, Colloquium Publiations Volume XXIII, 1939.

[SR09] G. J. Székely and M. L. Rizzo. Brownian distance covariance. *Annals of Applied Statistics* 3.(4) (2009), 1236–65.

[SRB07] G. J. Székely, M. L. Rizzo, and N. K. Bakirov. Measuring and testing dependence by correlation of distances. *Ann. Statist.* 35 (2007), 2769–2794.

[Tri83] H. Triebel. *Theory of function spaces*. Vol. 78. Monographs in Math. Birkhäuser, 1983.

[TZ17] M. Trübner and J. F. Ziegel. Dervivates of isotropic positive definite functions on spheres. *Proc. Amer. Math. Soc.* 145 (2017), 3017 –3031.

[Tu07] L. W. Tu. *An Introduction to Manifolds*. Springer, 2007.

[Wat62] G. N. Watson. *A treatise on the theory of Bessel functions*. 2nd. Cambridge University Press, 1962.

[Wen95] H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv. Comp. Math.* 4 (1995), 389 –396.

[Wil91] D. Williams. *Probability with martingales*. Cambridge University Press, 1991.

[Wil01] D. Williams. *Weighing the odds: A course in probability and statistics*. Cambridge University Press, 2001.

[Wol07] J. A. Wolf. *Harmonic analysis on commutative spaces*. Amer. Math. Soc., 2007.

[Xu18] Y. Xu. Positive definite functions on the unit sphere and integrals of Jacobi polynomials. *Proc. Amer. Math. Soc.* 146 (2018), 2039–2048.

[XC92] Y. Xu and E. W. Cheney. Strictly Positive Definite Functions on Spheres. *Proc. Amer. Math. Soc.* 116 (1992), 977 –981.

[Yad83] M. Ǐ. Yadrenko. *Spectral theory of random fields.* Translation Series in Mathematics and Engineering. New York: Optimisation Software, 1983.

[Zie14] J. Ziegel. Convolution roots and differentiability of isotropic positive definite functions on spheres. *Proc. Amer. Math. Soc.* 142 (2014), 2053 –2077.

[Zyg45] A. Zygmund. Smooth Functions. *Duke Math. J.* 12 (1945), 47 –76.

[Zyg68] A. Zygmund. *Trigonometric series, Vol. I, II.* Cambridge University Press, 1968.