REGULARITY OF GAUSSIAN PROCESSES ON DIRICHLET SPACES

GERARD KERKYACHARIAN, SHIGEYOSHI OGAWA, PENCHO PETRUSHEV, AND DOMINIQUE PICARD

Abstract. We study the regularity of centered Gaussian processes $(Z_x(\omega))_{x\in M}$, indexed by compact metric spaces (M,ρ). It is shown that the almost everywhere Besov space regularity of such a process is (almost) equivalent to the Besov regularity of the covariance $K(x,y)=\mathbb{E}(Z_xZ_y)$ under the assumption that (i) there is an underlying Dirichlet structure on M which determines the Besov space regularity, and (ii) the operator K with kernel $K(x,y)$ and the underlying operator A of the Dirichlet structure commute. As an application of this result we establish the Besov space regularity of Gaussian processes indexed by compact homogeneous spaces and, in particular, by the sphere.

1. Introduction

Gaussian processes have been at the heart of probability theory for very long time. There is a huge literature about it (see among many others [32], [29], [30] [2], [1] [34]). They also have been playing a key role in applications for many years and seem to experience an active revival in the recent domains of machine learning (see among others [36], [39]) as well as in Bayesian nonparametric statistics (see for example [48], [26]).

In many areas it is important to develop regularization procedures or sparse representations. Finding adequate regularizations as well as the quantification of the sparsity play an essential role in the accuracy of the algorithms in statistical theory as well as in Approximation theory. A way to regularize or to improve sparsity which is at the same time genuine and easily explainable is to impose regularity conditions.

The regularity of Gaussian processes has also been for a long time in the essentials of probability theory. It goes back to Kolmogorov in the 1930s (see among many others [18], [45], [47] [28], [31]).

In applications, an important effort has been put on the construction of Gaussian processes on manifolds or more general domains, with the two especially challenging examples of spaces of matrices and spaces of graphs to contribute to the emerging field of signal processing on graphs and extending high-dimensional data analysis to networks and other irregular domains.

Date: August 29, 2016.

2010 *Mathematics Subject Classification.* 58J35, 46E35, 42C15, 43A85.

Key words and phrases. Heat kernel, Gaussian processes, Besov spaces.

G. Kerkyacharian and D. Picard's work was supported by ANR Forewer.

Corresponding author: Dominique Picard, E-mail: picard@math.univ-paris-diderot.fr.

Motivated by these aspects we explore in this paper the regularity of Gaussian processes indexed by compact metric domains verifying some conditions in such a way that regularity conditions can be identified.

In effect, to prove regularity properties, we need a theory of regularity, compatible with the classical examples: Lipschitz properties and differentiability. At the same time we want to be able to handle more complicated geometries. For this aspect we borrow the geometrical framework developed in [14], [25].

Many of the constructions for regularity theorems are based on moments bounds for the increments of the process. Our approach here is quite different, it utilizes the spectral properties of the covariance operator. In particular, we use the Littlewood-Paley theory (this point of view was implicitly in [12]) to show that the Besov space regularity of the process is (almost) equivalent to the Besov regularity of the covariance operator. Especially, it is shown that the almost everywhere Besov space regularity of such a process is (almost) equivalent to the Besov regularity of the covariance $K(x,y) = \mathbb{E}(Z_x Z_y)$.

It is also important to notice that unlike many results in the literature, the regularity is expressed using the genuine distance of the domain, not the distance induced by the covariance.

We illustrate our approach by revisiting the Brownian motion as well as the fractional Brownian motion on the interval. We show the standard Besov regularity of these processes but also prove that they can be associated to a genuine geometry which finally appears in a nontrivial way.

We also illustrate our main result on the more refined case of two points homogeneous spaces and the special case of the unit sphere $\mathbb{S}^d$ in $\mathbb{R}^{d+1}$.

In the two subsequent sections, we recall the needed background information about Gaussian processes, the geometrical framework introduced in [14], [25], and how it allows to develop a smooth functional calculus as well as a description of regularity. In Section 4, we state the main results of the paper: the regularity theorem, The Ito-Nisio representation and the link with the RKHS. Subsection 4.3 details the seminal case of the Neumann operator and the standard Brownian motion. In this case, the salient fact is not the regularity result (which is known) but the original geometry corresponding to this process. The proofs of the main results are carried out in Section 5. Section 6 recalls some basic facts (and less basic) about positive and negative definite functions on two point homogeneous spaces. Section 7 establishes the Besov regularity of Gaussian processes indexed by the sphere. Section 8 is an appendix where we detail some facts on positive definite and negative definite functions as well as Gaussian probability on separable Banach spaces.

2. Gaussian processes: Background

In this section we recall some basic facts about Gaussian processes and establish useful notation.

2.1. General setting for Gaussian processes. Let $(\Omega, \mathcal{A}, P)$ be a probability space. Consider a centered Gaussian process on a set M, i.e. a family of real random variables $Z_x(\omega)$ with $x \in M$ and $\omega \in \Omega$ such that for all $n \in \mathbb{N}$, $x_1, \dots, x_n \in M$, and $\alpha_1, \dots, \alpha_n \in \mathbb{R}$

$$\sum_{i=1}^{n} \alpha_i Z_{x_i} \quad \text{is a centered Gaussian random variable.}$$

The covariance function $K(x,y)$ associated to such a process $(Z_x)_{x\in M}$ is defined by

$$K(x,y) := \mathbb{E}(Z_x Z_y) \quad \text{for} \;\; (x,y) \in M \times M.$$

It is readily seen that $K(x,y)$ is real-valued, symmetric, and positive definite (P.D.), i.e.

$$K(x,y) = K(y,x) \in \mathbb{R}, \quad \text{and}$$

$$\forall n \in \mathbb{N},\ \forall\, x_1,\dots,x_n \in M,\ \forall \alpha_1,\dots,\alpha_n \in \mathbb{R}, \quad \sum_{i,j\le n} \alpha_i\alpha_j K(x_i,x_j) \ge 0.$$

Remark 2.1. *In this paper, we only consider real Gaussian variables and real Hilbert spaces.*

2.2. **Gaussian processes with a zest of topology.** We now consider the following more specific setting. Let M be a compact space and let μ be a Radon measure on $(M,\mathcal{B})$ with support M and $\mathcal{B}$ being the Borel sigma algebra on M. Assuming that $(\Omega,\mathcal{A},P)$ is a probability space we let

$$Z : (M,\mathcal{B}) \otimes (\Omega,\mathcal{A}) \mapsto Z_x(\omega) \in \mathbb{R} \text{ be a measurable map}$$

such that $(Z_x)_{x\in M}$ is a Gaussian process. In addition, we assume that $K(x,y)$ is a symmetric, continuous, and positive definite function on $M \times M$. Then obviously the operator K defined by

$$Kf(x) := \int_M K(x,y) f(y) d\mu(y), \quad f \in L^2(M,\mu),$$

is a self-adjoint compact positive operator (even trace-class) on $L^2(M,\mu)$. Moreover, $K(L^2) \subset C(M)$, the Banach space of continuous functions on M. Let $\nu_1 \ge \nu_2 \ge \cdots > 0$ be the sequence of eigenvalues of K repeated according to their multiplicities and let $(u_k)_{k\ge1}$ be the sequence of respective normalized eigenfunctions:

$$\int_M K(x,y) u_k(y) d\mu(y) = \nu_k u_k(x).$$

The functions u_k are continuous real-valued functions and the sequence $(u_k)_{k\ge1}$ is an orthonormal basis for $L^2(M,\mu)$. By Mercer's theorem we have the following representation:

$$K(x,y) = \sum_k \nu_k u_k(x) u_k(y),$$

where the convergence is uniform.

Let $\mathcal{H} \subset L^2(\Omega,P)$ be the closed Gaussian space spanned by finite linear combinations of $(Z_x)_{x\in M}$. Clearly, interpreting the following integral as Bochner integral with value in the Hilbert space $\mathcal{H}$, we can define

$$B_k(\omega) = \frac{1}{\sqrt{\nu_k}} \int_M Z_x(\omega) u_k(x) d\mu(x) \in \mathcal{H}.$$

It is not difficult to prove that B_k is a sequence of independent $N(0,1)$ variables and that the process

$$\tilde{Z}_x(\omega) := \sum_k \sqrt{\nu_k} u_k(x) B_k(\omega)$$

is a modification of $Z_x(\omega)$, i.e. $P(Z_x = \tilde{Z}_x) = 1$, $\forall x \in M$.

We are interested in the regularity of the "trajectory" $x \in M \mapsto Z_x(\omega)$ for almost all $\omega \in \Omega$ and for a suitable modification of $Z_x(\omega)$ and for this reason, we will focus on the version $\tilde{Z}_x(\omega)$.

3. Regularity spaces on metric spaces with Dirichlet structure

On a compact metric space (M, ρ) one has the scale of s-Lipschitz spaces defined by the norm

$$\|f\|_{\mathrm{Lip}_s} := \|f\|_\infty + \sup_{x \neq y} \frac{|f(x) - f(y)|}{\rho(x,y)^s}, \quad 0 < s \leq 1. \tag{3.1}$$

In Euclidian spaces a function can be much more regular than Lipschitz, for instance differentiable at different order, or belong to some Sobolev space, or even in a more refine way to a Besov space. For this purpose, we consider metric measured spaces with Dirichlet structure. This setting is rich enough to develop a Littlewood-Paley theory in almost complete analogy with the classical case on $\mathbb{R}^d$, see [14, 25]. In particular, it allows to develop Besov spaces B^s_{pq} with all set of indices. At the same time this framework is sufficiently general to cover a number of interesting cases as will be shown in what follows. We next describe the underlying setting in detail.

3.1. **Metric spaces with Dirichlet structure.**

We assume that (M, μ) is a compact connected measure space, where μ is a Radon measure with support M. Also, assume that A is a self-adjoint non-negative operator mapping real-valued to real-valued functions with dense domain $D(A) \subset L^2(M, \mu)$. Let $P_t = e^{-tA}$, $t > 0$, be the associate self-adjoint semi-group. Furthermore, we assume that A determines a local and regular Dirichlet structure, see [14] and for details [19], [43], [41], [42], [44], [11], [22]. In fact, we assume that P_t is a Markov semi-group (A verifies the Beurling-Deny condition):

$$0 \leq f \leq 1 \text{ and } f \in L^2 \text{ imply } 0 \leq P_t f \leq 1,$$

and also $P_t \mathbb{1}_M = \mathbb{1}_M$ (equivalently $A\mathbb{1}_M = 0$). From this it follows that P_t can be extended as a contraction operator on $L^p(M, \mu)$ for $1 \leq p \leq \infty$, i.e. $\|P_t f\|_p \leq \|f\|_p$, and $P_t P_s f = P_{t+s} f$, $t, s > 0$.

The next assumption is that there exists a sufficiently rich subspace $\tilde{D} \subset D(A)$ such that $f \in \tilde{D} \Longrightarrow f^2 \in D(A)$ (see [11]). Then we define a bilinear operator "square gradiant" $\Gamma : \tilde{D} \times \tilde{D} \mapsto L^1$ by

$$\Gamma(f, g) := -\frac{1}{2}[A(fg) - fA(g) - gA(f)].$$

Consequently, $\Gamma(f, f) \geq 0$ and $\int_M A(f) g d\mu = \int_M \Gamma(f, g) d\mu$ (formula for integration by parts).

Main assumptions:

(1) Let

$$\rho(x, y) := \sup_{\Gamma(f,f) \leq 1} f(x) - f(y) \quad \text{for } x, y \in M. \tag{3.2}$$

We assume that ρ is a metric on M that generates the original topology on M.

(2) *The doubling property:* Denote $B(x,r) = \{y \in M : \rho(x,y) < r\}$. The assumption is that there exists a constant $c_0 > 0$ such that

$$\mu(B(x,2r)) \le c_0\mu(B(x,r)), \quad \forall x \in M,\ \forall r > 0. \tag{3.3}$$

This means that (M,ρ,μ) is a homogeneous space in the sense of Coifman and Weiss [13]. Observe that from (3.3) it follows that

$$\mu(B(x,\lambda r)) \le c_0\lambda^d\mu(B(x,r)) \quad \text{for } x \in M,\ r > 0, \text{ and } \lambda > 1, \tag{3.4}$$

where $d := \log_2 c_0$; the constant d plays the role of a dimension.

(3) *Poincaré inequality:* There exists a constant $c > 0$ such that for all $f \in \tilde{D}$, $x \in M$, and $r > 0$

$$\inf_{\lambda\in\mathbb{R}} \int_{B(x,r)} (f-\lambda)^2 d\mu \le cr^2 \int_{B(x,r)} \Gamma(f,f)d\mu.$$

As a consequence the associated semi-group $P_t = e^{-tA}$, $t > 0$, consists of integral operators of continuous (heat) kernel $p_t(x,y) \ge 0$, with the following properties:

(a) Gaussian localization: For all $x, y \in M$ and $t > 0$

$$\frac{c_1 \exp\{-\frac{\rho^2(x,y)}{c_2 t}\}}{\sqrt{\mu(B(x,\sqrt{t}))\mu(B(y,\sqrt{t}))}} \le p_t(x,y) \le \frac{c_3 \exp\{-\frac{\rho^2(x,y)}{c_4 t}\}}{\sqrt{\mu(B(x,\sqrt{t}))\mu(B(y,\sqrt{t}))}}. \tag{3.5}$$

(b) Hölder continuity: There exists a constant $\kappa > 0$ such that

$$\big|p_t(x,y) - p_t(x,y')\big| \le c_1\Big(\frac{\rho(y,y')}{\sqrt{t}}\Big)^{\kappa} \frac{\exp\{-\frac{\rho^2(x,y)}{c_2 t}\}}{\sqrt{\mu(B(x,\sqrt{t}))\mu(B(y,\sqrt{t}))}} \tag{3.6}$$

for $x, y, y' \in M$ and $t > 0$, whenever $\rho(y,y') \le \sqrt{t}$.

(c) Markov property:

$$\int_M p_t(x,y)d\mu(y) = 1 \quad \text{for } x \in M \text{ and } t > 0. \tag{3.7}$$

Above $c_1, c_2, c_3, c_4 > 0$ are structural constants.

Remark 3.1. *The setting described above is quite general. This setting covers, in particular, the case of compact Riemannian manifolds. It naturally contains the cases of the sphere, interval, ball, and simplex with weights. For more details, see* [14].

Notation. Throughout we will use the notation $|E| := \mu(E)$ and $\mathbb{1}_E$ will stand for the characteristic function of $E \subset M$. Also $\|\cdot\|_p = \|\cdot\|_{L^p} := \|\cdot\|_{L^p(M,\mu)}$. Positive constants will be denoted by $c, c', c_1, C, C', \dots$ and they may vary at every occurrence. The notation $a \sim b$ will stand for $c_1 \le a/b \le c_2$. As usual we will denote by $\mathbb{N}$ the set of all natural numbers and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

From the compactness of M and the fact that A is an essentially self-adjoint non-negative operator it follows that the spectrum of A is discrete and of the form: $0 \le \lambda_1 < \lambda_2 < \cdots$. Furthermore, the respective eigenspaces $\mathcal{H}_{\lambda_k} := \mathrm{Ker}(A - \lambda_k \,\mathrm{Id})$ are finite dimensional and

$$L^2(M,\mu) = \bigoplus_{k\ge1} \mathcal{H}_{\lambda_k}.$$

Denoting by $P_{\mathcal{H}_{\lambda_k}}$ the orthogonal projector onto $\mathcal{H}_{\lambda_k}$ the above means that all $f \in L^2(M,\mu)$ can be expressed in the following form $f = \sum_{k\ge1} P_{\mathcal{H}_{\lambda_k}} f$. In addition,

$$Af = \sum_{k\ge1} \lambda_k P_{\mathcal{H}_{\lambda_k}} f, \quad \forall f \in D(A), \quad \text{and} \quad P_t f = \sum_{k\ge1} e^{-t\lambda_k} P_{\mathcal{H}_{\lambda_k}} f, \quad \forall f \in L^2. \tag{3.8}$$

More generally, for a function $g \in L^\infty(\mathbb{R}_+)$ the operator $g(\sqrt{A})$ is defined by

$$g(\sqrt{A})f := \sum_{k\ge1} g(\sqrt{\lambda_k}) P_{\mathcal{H}_{\lambda_k}} f, \quad \forall f \in L^2. \tag{3.9}$$

The spectral spaces Σ_λ, $\lambda > 0$, associated with $\sqrt{A}$ are defined by

$$\Sigma_\lambda := \bigoplus_{\sqrt{\lambda_k}\le\lambda} \mathcal{H}_{\lambda_k}.$$

Observe that $\Sigma_\lambda \subset L_\infty$ and hence $\Sigma_\lambda \subset L^p$ for $1 \le p \le \infty$.

From now on we will assume that the eigenvalues $(\lambda_k)_{k\ge1}$ are enumerated with algebraic multiplicities taken into account, i.e. if the algebraic multiplicity of λ is m then λ is repeated m times in the sequence $0 \le \lambda_1 \le \lambda_2 \le \cdots$. We let $(u_k)_{k\ge1}$ be respective real orthogonal and normalized in L^2 eigenfunctions of A, that is, $Au_k = \lambda_k u_k$.

Let $\Pi_\delta(x,y) := \sum_{\sqrt{\lambda_k}\le\delta^{-1}} u_k(x)u_k(y)$, $\delta > 0$, be the kernel of the orthogonal projector onto $\Sigma_{1/\delta}$. Then as is shown in [14, Lemma 3.19]

$$\Pi_\delta(x,x) \sim |B(x,\delta)|^{-1}. \tag{3.10}$$

A key trait of our setting is that it allows to develop a smooth functional calculus. In particular, if $g \in C^\infty(\mathbb{R})$ is even, then the operator $g(t\sqrt{A})$ defined in (3.9) is an integral operator with kernel $g(t\sqrt{A})(x,y)$ having this localization: For any $\sigma > 0$ there exists a constant $c_\sigma > 0$ such that

$$\big|g(t\sqrt{A})(x,y)\big| \le c_\sigma |B(x,t)|^{-1}\big(1 + t^{-1}\rho(x,y)\big)^{-\sigma}, \quad \forall x,y \in M. \tag{3.11}$$

Furthermore, $g(t\sqrt{A})(x,y)$ is Hölder continuous. An immediate consequence of (3.11) is that the operator $g(t\sqrt{A})$ is bounded on $L^p(M)$:

$$\|g(t\sqrt{A})f\|_p \le c\|f\|_p, \quad \forall f \in L^p(M), \quad 1 \le p \le \infty. \tag{3.12}$$

For more details and proofs, see [14, 25].

For discretization (sampling) purposes, we will use *maximal δ-nets*. Recall that a set $\mathcal{X} \subset M$ is a maximal δ-net on M ($\delta > 0$) if $\rho(x,y) \ge \delta$ for all $x,y \in \mathcal{X}$, $x \ne y$, and $\mathcal{X}$ is maximal with this property. It is easily seen that a maximal δ-net on M always exists. Of course, if $\delta > \mathrm{Diam}(M)$, then $\mathcal{X}$ will consists of a single point. The following useful assertion is part of Theorem 4.2 in [14].

Proposition 3.2. *There exists a constant $\gamma > 0$, depending only on the structural constant of our setting, such that for any $\lambda > 0$ and $\delta := \gamma/\lambda$ there exists a δ-net $\mathcal{X}$ obeying*

$$2^{-1}\|g\|_\infty \le \max_{\xi\in\mathcal{X}} |g(\xi)| \le \|g\|_\infty, \quad \forall g \in \Sigma_\lambda. \tag{3.13}$$

Finally, if $N(\delta, M)$ is the covering number of M (or the cardinality of a maximal $\delta-$net), then

(3.14)
$$\dim(\Sigma_{\frac{1}{\sqrt{t}}}) \sim \int_M |B(x,\sqrt{t})|^{-1} d\mu(x) \sim N(\sqrt{t}, M) \sim \|e^{-tA}\|_{HS}^2 \le ct^{-d/2}, \quad t>0.$$

Here $\|e^{-tA}\|_{HS}^2 := \int_M \int_M |p_t(x,y)|^2 d\mu(x) d\mu(y)$ is the Hilbert-Schmidt norm.

3.2. **Regularity spaces.** In the general setting described above, the full scales of Besov and Tribel-Lizorkin spaces are available [14, 25].

The Sobolev spaces $W_p^k = W_p^k(A)$, $k \ge 1$, $1 \le p \le \infty$, are standardly defined by

$$W_p^k := \big\{ f \in D(A^{\frac{k}{2}}) : \|f\|_{W_p^k} := \|f\|_p + \|A^{\frac{k}{2}} f\|_p < \infty \big\}. \tag{3.15}$$

The Besov space $B_{pq}^s = B_{pq}^s$, $s>0$, $1 \le p, q \le \infty$, is defined by interpolation as in [35]:

$$B_{pq}^s := \big(L^p, W_p^k\big)_{\theta, q}, \quad \theta := s/k, \tag{3.16}$$

where $\big(L^p, W_p^k\big)_{\theta,q}$ is the real interpolation space between L^p and W_p^k, see [14].

The following Littlewood-Paley decomposition of functions will play an important role in the sequel. Suppose $\Phi \in C^\infty(\mathbb{R})$ is real-valued, even, and such that $\operatorname{supp} \Phi \subset [-2,2]$, $0 \le \Phi \le 1$, and $\Phi(\lambda) = 1$ for $\lambda \in [0,1]$. Let $\Psi(\lambda) := \Phi(\lambda) - \Phi(2\lambda)$. Evidently $\operatorname{supp} \Psi \cap \mathbb{R}_+ \subset [1/2, 2]$. Set

$$\Psi_0 := \Phi \quad \text{and} \quad \Psi_j(\lambda) := \Psi(2^{-j}\lambda) \quad \text{for } j \ge 1. \tag{3.17}$$

It is readily seen that $\Psi_0, \Psi \in C^\infty(\mathbb{R})$, Ψ_0, Ψ are even, $\operatorname{supp} \Psi_0 \subset [-2,2]$, $\operatorname{supp} \Psi_j \cap \mathbb{R}_+ \subset [2^{j-1}, 2^{j+1}]$, $j \ge 1$, and $\sum_{j \ge 0} \Psi_j(\lambda) = 1$ for $\lambda \in \mathbb{R}_+$. Consequently, for any $f \in L^p(M,\mu)$, $1 \le p \le \infty$, one has

$$f = \sum_{j \ge 0} \Psi_j(\sqrt{A}) f \quad \text{in} \quad L^p. \tag{3.18}$$

The Littlewood-Paley characterization of Besov spaces uses the functions Ψ_j from above: If $s>0$ and $1 \le p, q \le \infty$, then for a function $f \in L^p(M,\mu)$ we have

$$f \in B_{p,q}^s \Longleftrightarrow \|\Psi_j(\sqrt{A}) f\|_p = \varepsilon_j 2^{-js}, \;\; j \ge 0, \;\; \text{with} \;\; \{\varepsilon_j\} \in \ell^q. \tag{3.19}$$

Furthermore, if $f \in B_{p,q}^s$, then $\|f\|_{B_{p,q}^s} \sim \|\{\varepsilon_j\}\|_{\ell^q}$. We refer the reader to [14, 25] for proofs and more details on Besov spaces in the setting from §3.1. In particular, the following proposition clarifying the relationship between $B_{\infty,\infty}^s$ and $\operatorname{Lip} s$ (see [14, Proposition 6.4]).

Proposition 3.3. (*a*) *For any* $0 < s \le 1$ *we have* $\operatorname{Lip} s \subset B_{\infty,\infty}^s$.

(*b*) *Assuming that* $\kappa > 0$ *is the constant from* (3.6)*, then* $B_{\infty,\infty}^s \subset \operatorname{Lip} s$ *for* $0 < s < \kappa$*.*

Remark 3.4. *In the most interesting cases* $\kappa = 1$*, Proposition 3.3 implies that* $\operatorname{Lip} s = B_{\infty,\infty}^s$ *for* $0 < s < 1$*.*

4. Main results

In this section we state and discuss our main results. The proofs are carried out in the next section.

We consider a centered Gaussian process $(Z_x)_{x\in M}$ with covariance function $K(x,y)$ as described in § 2.2, indexed by a metric space M with Dirichlet structure just as described in § 3.1. We will adhere to the assumptions and notation from § 3.1.

4.1. **Commutation property.** We now make the *fundamental assumption* that the operator K with kernel $K(x,y)$ and the operator A from § 3.1 commute in the following sense:

Definition 4.1. *If K is a bounded operator on a Banach space $\mathbb{B}$ ($K\in\mathcal{L}(\mathbb{B})$) and A is an unbounded operator with domain $D(A)\subset\mathbb{B}$, we say that K and A commute if $K(D(A))\subset D(A)$ and*

$$KAf = AKf, \quad \forall f\in D(A).$$

Remark 4.2. *Let A be the infinitesimal generator of a contraction semi-group P_t. Then K and A commute in the sense of Definition 4.1 if and only if*

$$KP_t = P_tK, \quad \forall t>0.$$

We refer the reader to [16]*, Theorem 6.1.27.*

We now return to the covariance operator K and the underlying self-adjoint non-negative operator A from our setting. In light of Proposition 4.2 our assumption that K and A commute implies that they have the same eigenspaces.

Recall that the eigenvalues of A are ordered in a sequence $0=\lambda_1\le\lambda_2\le\cdots$, where the eigenvalues are repeated according to their multiplicities, and the respective eigenfunctions $(u_k)_{k\ge1}$ are real-valued, orthogonal, and normalized in L^2. Let $(\nu_k)_{k\ge1}$ be the eigenvalues of the covariance operator K. Then

$$Au_k=\lambda_k u_k \quad \text{and} \quad Ku_k=\nu_k u_k, \quad k\ge1. \tag{4.1}$$

Remark 4.3. *As a consequence of the commutation property of K and A, the operator AK is defined everywhere on $L^2(M,\mu)$ and is closed. Therefore, AK is a continuous operator from $L^2(M,\mu)$ to $L^2(M,\mu)$. Clearly,*

$$KAf=\sum_{k\ge1}\langle f,u_k\rangle\lambda_k\nu_k u_k, \quad \forall f\in L^2 \textit{ and hence } \quad \|KA\|_{\mathcal{L}(L^2)}=\sup_{k\ge1}\lambda_k\nu_k<\infty.$$

Remark 4.4. *Assume that we are in the geometric setting described in §3.1, associated to an operator A. As in §4.1, suppose $K(x,y)$ is a P.D. kernel such that the associate operator K commutes with A. It is easy to see that*

$$A\mathbb{1}_M=0 \quad \textit{and} \quad \dim\operatorname{Ker}(A)=1. \tag{4.2}$$

Indeed, the Markov property (3.7) *yields $A\mathbb{1}_M=0$. To show that $\dim\operatorname{Ker}(A)=1$, assume that $Af=0$, $f\in D(A)$. Then $\Gamma(f,f)=0$. Assume that $f\neq$ constant. Then $f(x)\neq f(y)$ for some $x,y\in M$, $x\neq y$. Since $\Gamma(f,f)=0$ we have $\Gamma(af,af)=0$ for each $a>0$. Then by* (3.2) *$\rho(x,y)\ge a|f(x)-f(f)|$, $\forall a>0$, implying $\rho(x,y)=\infty$, which is a contradiction because M is connected (see* [14]*). Therefore, $Af=0$ implies $f=$ constant and hence $\dim\operatorname{Ker}(A)=1$.*

As a consequence of (4.2) *we have*

$$AK\mathbb{1}_M = KA\mathbb{1}_M = 0.$$

However, as $\dim \mathrm{Ker}(A) = 1$*, necessarily* $K\mathbb{1}_M = C\mathbb{1}_M$*.*

4.2. **Regularity theorem, Ito-Nisio representation and RKHS.** We now come to the main results of this paper.

Theorem 4.5. *Let* $(Z_x)_{x\in M}$ *be a centered Gaussian process with covariance function* $K(x,y) := \mathbb{E}(Z_x Z_y)$ *indexed by a metric space* M *with Dirichlet structure induced by a self-adjoint operator* A *such that* K *and* A *commute in the sense of Definition 4.1. Then the following assertions hold:*

(*a*) *If the covariance kernel* $K(x,y)$ *has the regularity described by*

$$\sup_{x\in M} \|K(x,\bullet)\|_{B^s_{\infty,\infty}} < \infty \quad \textit{for some } \; s > 0, \tag{4.3}$$

then the Gaussian process $Z_x(\omega)$ *has the following regularity: For any* $0 < \alpha < \frac{s}{2}$

$$Z_x(\omega) \in B^\alpha_{\infty,1} \quad \textit{for almost all}\, \omega \quad (B^\alpha_{\infty,1} \subset B^\alpha_{\infty,\infty}).$$

(*b*) *Conversely, suppose there exists* $\alpha > 0$ *such that* $Z_x(\omega) \in B^\alpha_{\infty,\infty}$ *for almost all* ω*. Then*

$$\sup_{x\in M} \|K(x,\bullet)\|_{B^{2\alpha}_{\infty,\infty}} < \infty.$$

Remark 4.6. *It is interesting to observe that because of the second part of the theorem condition* (4.3) *is necessary.*

Another key point is that in the above theorem the Besov space smoothness parameter $s > 0$ *can be arbitrarily large, while* $0 < s \le 1$ *in the case when the regularity is characterized in terms of Lipschitz spaces.*

4.2.1. *Ito-Nisio representation.* The following theorem gives an Ito-Nisio representation of the process.

Theorem 4.7 (Wiener measure)**.** *In the setting from above, if* $K(x,y)$ *is a continuous positive definite function on* M *such that*

$$\sup_{x\in M} \|K(x,\bullet)\|_{B^s_{\infty,\infty}} < \infty$$

and the associated kernel operator K *commutes with* A*, then there is a unique probability measure* Q *on the Borelian sets of* $B^\alpha_{\infty,1}$*,* $\alpha < \frac{s}{2}$*, such that the family of random variables:*

$$\forall x \in M,\ \omega \in B^\alpha_{\infty,1} \xrightarrow{\delta_x} \omega(x) \in \mathbb{R}$$

is a centered Gaussian process of covariance $K(x,y)$*.*

4.2.2. *Reproducing Kernel Hilbert Spaces (RKHS).* We finally connect condition (4.3) with the RKHS associated to the process Z_x (see the appendix).

As is well known the covariance kernel K determines a real Hilbert space $\mathbb{H}_K$ of functions for which the evaluation:

$$\forall x \in M,\ \delta_x : f \in \mathbb{H}^*_K \mapsto f(x) \text{ is continuous.}$$

Moreover,

$$y \mapsto K(x,y) = K_x(y) \in \mathbb{H}_K, \quad \forall f \in \mathbb{H}_K,\ \delta_x(f) = \langle f, K_x\rangle_{\mathbb{H}_K},$$

and $(K_x)_{x\in M}$ is a total set in $\mathbb{H}_K$. The space $\mathbb{H}_K$ is the completion of span $\{K(x,\cdot) : x \in M\}$, more precisely

$$\mathbb{H}_K^\circ := \Big\{h(y) = \sum_{i\in F} \alpha_i K(x_i, y) : \ \|h\|_{\mathbb{H}}^2 = \sum_{i,j\in F} \alpha_i\alpha_j K(x_i, x_j) = \sum_{j\in F} \alpha_j h(x_j)\Big\}.$$

It is also well known (see e.g. [15]) that

$$\|h\|_{\mathbb{H}}^2 = 0 \ \text{ for } \ h \in \mathbb{H}_K^\circ \iff h(y) = 0, \ \forall y \in M.$$

Furthermore (see [33]),

$$K(x,y) = \sum_{i\in I} g_i(x)g_i(y) \iff g_i \in \mathbb{H}_K, \ \forall i \quad \text{and} \quad (g_i)_{i\in I} \text{ is a tight frame for } \mathbb{H}_K.$$

In our geometric framework, (4.1) entails the following representation of K

$$K(x,y) = \sum_k \nu_k u_k(x)u_k(y) \quad \text{and} \quad \nu_k \geq 0. \tag{4.4}$$

Therefore, $(\sqrt{\nu_k}u_k)_{k\in\mathbb{N},\nu_k\neq 0}$ is a tight frame of $\mathbb{H}$, and moreover $(\delta_x)_{x\in M} \subset \mathbb{H}_K^*$ is dense in $\mathbb{H}_K^*$ in the weak $\sigma(\mathbb{H}_K^*, \mathbb{H}_K)$ topology. Actually, by Mercer's theorem we have (see [40], [24]): Let $\mathbb{N}(\nu) := \{k \in \mathbb{N}, \nu_k \neq 0\}$ and define

$$\mathcal{H} := \Big\{f : M \mapsto \mathbb{R} : f(x) = \sum_{k\in\mathbb{N}(\nu)} \alpha_k \sqrt{\nu_k}\, u_k(x), \ (\alpha_k) \in \ell^2\Big\}$$

with inner product

$$\langle f, g\rangle_{\mathcal{H}} = \Big\langle \sum_{k\in\mathbb{N}(\nu)} \alpha_k\sqrt{\nu_k}\, u_k(\cdot), \sum_{k\in\mathbb{N}(\nu)} \beta_k\sqrt{\nu_k}u_k(\cdot)\Big\rangle_{\mathcal{H}} := \langle(\alpha_k), (\beta_k)\rangle_{\ell^2(\mathbb{N}(\nu))}.$$

Then $\mathcal{H}$ is a Hilbert space of continuous functions and $(\sqrt{\nu_k}u_k)_{k\in\mathbb{N}(\nu)}$ is an orthonormal basis for $\mathcal{H}$ and hence $\mathbb{H}_K = \mathcal{H}$.

In fact, the following theorem holds.

Theorem 4.8. *We have for $s > 0$*

$$\mathbb{H}_K \subseteq B^{\frac{s}{2}}_{\infty,\infty} \iff \sup_{x\in M} \|K(x,\bullet)\|_{B^s_{\infty,\infty}} < \infty.$$

4.3. **Seminal example: the Neumann operator on $[0,1]$ and the Brownian motion.** Here we show that the classical Brownian motion on $[0,1]$ is a particular case of our general theory.

4.3.1. *The Neumann operator on* $[0,1]$. Let $H^2([0,1])$ be the space of the functions $f \in L^2([0,1])$ twice weakly differentiable and such that $f', f'' \in L^2([0,1])$. Consider the operator

$$Af := -f'', \ \ D(A) := \{f \in H^2([0,1]) : f'(0) = f'(1) = 0\}.$$

Clearly,

$$\int_0^1 (Af)g dx = \int_0^1 f'g' dx = \int_0^1 fAg dx$$

and hence A is positive and symmetric. In fact, A generates a Dirichlet space, and also

$$\cos k\pi x \in D(A) \quad \text{and} \quad A(\cos k\pi\bullet)(x) = (\pi k)^2 \cos k\pi x, \ \ k \geq 1.$$

Therefore, $\{\mathbb{1}, (\sqrt{2}\cos k\pi x)_{k\in\mathbb{N}}\}$ is an orthonormal basis of $L^2([0,1])$ consisting of eigenvectors of A. Write $H^1([0,1]) := \big\{f \in L^2([0,1]) : \int_0^1 |f'(u)|^2 du < \infty\big\}$. This allows to define a Dirichlet form:

$$A, D(A) = \Big\{f \in H^1([0,1]) : \Big|\int_0^1 f'(x)\phi'(x)dx\Big| \leq c\|\phi\|_2,\ \forall \phi \in H^1([0,1])\Big\}.$$

Thus

$$\int_0^1 f'(x)\phi'(x)dx = \int_0^1 Af(x)\phi(x)dx$$

and the distance is defined by

$$\rho(x,y) = \sup_{\phi\in H^1([0,1]):|\phi'|\leq 1} \phi(x) - \phi(y) = |x-y|.$$

The Poincaré inequality is well known to be valid in this case. Hence we are now in the setting described above.

4.3.2. *Brownian motion.* Clearly, $\psi(x,y) = |x-y|$ is a negative definite function on $[0,1]$ (see the appendix) as

$$|x-y| = \int_{[0,1]} |1_{[0,x]}(u) - 1_{[0,y]}(u)|^2 du$$

Therefore, there is a natural positive definite function $\tilde{K}(x,y)$ associated to ψ (see again the appendix):

$$\begin{aligned}\tilde{K}(x,y) &= \frac{1}{2}\Big(\int_0^1 |x-u|du + \int_0^1 |y-u|du - |x-y|\Big)\\ &= \frac{1}{4}[x^2 + (1-x)^2 + y^2 + (1-y)^2 - 2|x-y|]\\ &= x\wedge y + \frac{(1-x)^2 + (1-y)^2 - 1}{2}.\end{aligned}$$

It is easy to verify that $\tilde{K}$ and A commute, as

$$\tilde{K}(\cos k\pi\bullet)(x) = \frac{\cos k\pi x}{(\pi k)^2},\quad \forall k\in\mathbb{N},\quad \text{and}\quad \tilde{K}\mathbb{1} = (1/6)\mathbb{1}.$$

(It is easy to see that $\int_0^1 |x-y|\cos k\pi y\ dy = -\frac{2\cos k\pi x}{(\pi k)^2} + \frac{1+(-1)^k}{(\pi k)^2}$.)

$$\text{So: } \tilde{K}(x,y) = \frac{1}{6} + 2\sum_{k\geq 1} \frac{\cos k\pi x.\cos k\pi y}{(\pi k)^2}.$$

Also, $\tilde{K}(x,\bullet)$ is uniformly Lip 1. Therefore, Z_x the centered Gaussian process associated to $\tilde{K}$ is almost surely Lip α, $\alpha < \frac{1}{2}$. The process $Y_x(\omega) = Z_x(\omega) - Z_0(\omega)$ has the same regularity and

$$\mathbb{E}(Y_xY_y) = \frac{1}{2}(|x| + |y| - |x-y|) = x\wedge y$$

is the well known associated kernel. So, $\{Y_x : x\in[0,1]\}$ is the classical Brownian motion.

5. Proof of the main results

The purpose of this section is to prove Theorems 4.5, 4.7, 4.8. For this we need some preparation.

5.1. **Uniform Besov property of $K(x,y)$ and discretization.** Recalling (4.4) we next represent the Besov norm of $K(x,\bullet)$ in terms of the eigenvalues and eigenfunctions of K and A.

Theorem 5.1. *Let $s>0$. Then*

$$\sup_{x\in M}\|K(x,\bullet)\|_{B^s_{\infty,\infty}} \sim \max\Big\{\sup_{x\in M}\sum_{k:\sqrt{\lambda_k}\le 1}\nu_k u_k^2(x),\ \sup_{j\ge1}2^{js}\sup_{x\in M}\sum_{k:2^{j-1}<\sqrt{\lambda_k}\le 2^j}\nu_k u_k^2(x)\Big\}. \tag{5.1}$$

Proof. Note first that from (3.19) it follows that (with Ψ_j from (3.17))

$$\sup_x\|K(x,\bullet)\|_{B^s_{\infty,\infty}}\sim\sup_{j\ge0}2^{js}\sup_x\|\Psi_j(\sqrt{A})K(x,\bullet)\|_\infty.$$

But, using (4.4) we have $\big(\Psi_j(\sqrt{A})K(x,\bullet)\big)(y)=\sum_k\Psi_j(\sqrt{\lambda_k})\nu_k u_k(x)u_k(y)$ and hence, applying the Cauchy-Schwarz inequality it follows that

$$\sup_{x,y}\big|\big(\Psi_j(\sqrt{A})K(x,\bullet)\big)(y)\big|=\sup_x\sum_k\Psi_j(\sqrt{\lambda_k})\nu_k u_k^2(x).$$

Consequently,

$$\sup_x\|K(x,\bullet)\|_{B^s_{\infty,\infty}}\sim\sup_j 2^{js}\sup_x\sum_k\Psi_j(\sqrt{\lambda_k})\nu_k u_k^2(x). \tag{5.2}$$

Clearly, from (3.17) we have $0\le\Psi_j\le1$, $\operatorname{supp}\Psi_0\cap\mathbb{R}_+\subset[0,2]$, and $\operatorname{supp}\Psi_j\cap\mathbb{R}_+\subset[2^{j-1},2^{j+1}]$ for $j\ge1$. Therefore,

$$\sup_x\sum_k\Psi_0(\sqrt{\lambda_k})\nu_k u_k^2(x)\le\sup_x\sum_{\sqrt{\lambda_k}<2}\nu_k u_k^2(x)\quad\text{and}$$
$$\sup_x\sum_k\Psi_j(\sqrt{\lambda_k})\nu_k u_k^2(x)\le\sup_x\sum_{2^{j-1}<\sqrt{\lambda_k}<2^{j+1}}\nu_k u_k^2(x),\quad j\ge1.$$

These estimates and (5.2) readily imply that the left-hand side quantity in (5.1) is dominated by a constant multiple of the right-hand side.

In the other direction, observe that by construction $\Psi_0(\lambda)=1$ for $\lambda\in[0,1]$ and $\Psi_{j-1}(\lambda)+\Psi_j(\lambda)=1$ for $\lambda\in[2^{j-1},2^j]$, $j\ge1$. Hence

$$\sup_x\sum_{\sqrt{\lambda_k}\le1}\nu_k u_k^2(x)\le\sup_x\sum_k\Psi_0(\sqrt{\lambda_k})\nu_k u_k^2(x)\quad\text{and}$$
$$\sup_x\sum_{2^{j-1}<\sqrt{\lambda_k}\le2^j}\nu_k u_k^2(x)\le\sup_x\sum_k\Psi_{j-1}(\sqrt{\lambda_k})\nu_k u_k^2(x)$$
$$+\sup_x\sum_k\Psi_j(\sqrt{\lambda_k})\nu_k u_k^2(x),\quad j\ge1.$$

These inequalities and (5.2) imply that the right-hand side in (5.1) is dominated by a constant multiple of the left-hand side. This completes the proof. $\square$

The following corollary is an indication of how the Besov regularity relates with the "dimension" d of the set M, which appears here through the doubling condition (3.3).

Corollary 5.2. *Let $\gamma > d$ and $s = \gamma - d$. Then*

$$\nu_k = O\big(\sqrt{\lambda_k}\big)^{-\gamma} \implies \sup_x \|K(x,\bullet)\|_{B^s_{\infty,\infty}} \le c.$$

Proof. If $\nu_k \le c\big(\sqrt{\lambda_k}\big)^{-\gamma}$, then using (3.10) and (3.4) we get for any $j \ge 1$ and $x \in M$

$$\sum_{k:2^{j-1}\le\sqrt{\lambda_k}\le 2^j} \nu_k u_k^2(x) \le c2^{-\gamma(j+1)} \sum_{k:2^{j-1}\le\sqrt{\lambda_k}\le 2^j} u_k^2(x) \le c2^{-\gamma j} \sum_{k:\sqrt{\lambda_k}\le 2^j} u_k^2(x)$$
$$= c2^{-\gamma j}\Pi_{2^j}(x,x) \le c2^{-\gamma j}|B(x,2^{-j})|^{-1} \le c2^{-j(\gamma-d)}.$$

A similar estimate with $j = 0$ holds for all k such that $\sqrt{\lambda_k} \le 1$. Then the corollary follows by Theorem 5.1. $\square$

Remark 5.3. *Observe that*

$$\sup_x \sum_{k:2^{j-1}\le\sqrt{\lambda_k}\le 2^j} \nu_k u_k^2(x) \le c2^{-js}$$

implies

$$\sum_{k:2^{j-1}\le\sqrt{\lambda_k}\le 2^j} \nu_k = \sum_{k:2^{j-1}\le\sqrt{\lambda_k}\le 2^j} \int_M \nu_k u_k^2(x) d\mu(x) \le c2^{-js}|M|.$$

We will utilize maximal δ-nets on M along with Proposition 3.2 for discretization. For any $j \ge 0$ we denote by $\mathcal{X}_j$ the maximal δ-net from Proposition 3.2 with $\delta := \gamma 2^{-j-1}$ such that

$$2^{-1}\|g\|_\infty \le \max_{\xi\in\mathcal{X}_j} |g(\xi)| \le \|g\|_\infty, \quad \forall g \in \Sigma_{2^{j+1}}. \tag{5.3}$$

The following claim will be instrumental in the proof of Theorem 4.5.

Proposition 5.4. *We have*

$$\sup_{x\in M} \sum_{k:\sqrt{\lambda_k}\le 1} \nu_k u_k^2(x) \sim \max_{\xi\in\mathcal{X}_0} \sum_{k:\sqrt{\lambda_k}\le 1} \nu_k u_k^2(\xi)$$

and for any $j \ge 1$

$$\sup_{x\in M} \sum_{k:2^{j-1}<\sqrt{\lambda_k}\le 2^j} \nu_k u_k^2(x) \sim \max_{\xi\in\mathcal{X}_j} \sum_{k:2^{j-1}<\sqrt{\lambda_k}\le 2^j} \nu_k u_k^2(\xi)$$

with absolute constants of equivalence.

This proposition follows readily from the following

Lemma 5.5. *Let $\mathcal{X}_j$ be the maximal δ-net from above with $\delta := \gamma 2^{-j}$, $j \ge 0$, and let*

$$H(x,y) := \sum_{\sqrt{\lambda_k}\le 2^j} \alpha_k u_k(x) u_k(y), \quad \textit{where} \quad \alpha_k \ge 0.$$

Then

$$\max_{\xi\in\mathcal{X}_j} H(\xi,\xi) \le \sup_{x,y\in M} |H(x,y)| \le 4 \max_{\xi\in\mathcal{X}_j} H(\xi,\xi).$$

Proof. Clearly $H(x,y)$ is a positive definite function and hence $|H(x,y)| \le \sqrt{H(x,x)H(y,y)}$, implying

$$\max_{\xi,\eta\in\mathcal{X}_j} |H(\xi,\eta)| = \max_{\xi\in\mathcal{X}_j} H(\xi,\xi). \tag{5.4}$$

Evidently, for any fixed $x \in M$ the function $H(x,y) \in \Sigma_{2^j}$ as a function of y and by (5.3)

$$\sup_{y\in M} |H(x,y)| \le 2\max_{\eta\in\mathcal{X}_j} |H(x,\eta)|.$$

Now, using that $H(x,\eta) \in \Sigma_{2^j}$ as a function of x, we again apply (5.3) to obtain

$$\begin{aligned}\sup_{x,y\in M} |H(x,y)| &\le 2\sup_{x\in M}\max_{\eta\in\mathcal{X}_j} |H(x,\xi)| = 2\max_{\eta\in\mathcal{X}_j}\sup_{x\in M} |H(x,\eta)|\\ &\le 4\max_{\eta\in\mathcal{X}_j}\max_{\xi\in\mathcal{X}_j} |H(\xi,\eta)| = 4\max_{\xi\in\mathcal{X}_j} H(\xi,\xi).\end{aligned}$$

Here for the last equality we used (5.4). This completes the proof. □

5.2. **Proof of Theorem 4.5.** (a) Assume $\sup_{x\in M}\|K(x,\bullet)\|_{B^s_{\infty,\infty}} < \infty$. Let $(B_k(\omega))_{k\ge1}$ be a sequence of independent $N(0,1)$ variables. Then as alluded in §2.2

$$\tilde{Z}_x(\omega) := \sum_k \sqrt{\nu_k}u_k(x)B_k(\omega)$$

is also a version of $Z_x(\omega)$. Let Ψ_j, $j \ge 0$, be the functions from (3.17) and observe that $f \in B^s_{\infty,1}$ if and only if $\|f\|_{B^s_{\infty,1}} \sim \sum_{j\ge0} 2^{js}\|\Psi_j(\sqrt{A})f\|_\infty < \infty$. Clearly,

$$\big(\Psi_j(\sqrt{A})\tilde{Z}_\bullet(\omega)\big)(x) = \sum_k \Psi_j(\sqrt{\lambda_k})\sqrt{\nu_k}u_k(x)B_k(\omega). \tag{5.5}$$

For each $x \in M$ this is a Gaussian variable of variance

$$\sigma_j^2(x) = \sum_k \Psi_j^2(\sqrt{\lambda_k})\nu_k u_k(x)^2 \le c2^{-js}.$$

Here we used that $\Psi_j^2(\sqrt{\lambda_k}) \le 1$, the assumption $\sup_{x\in M}\|K(x,\bullet)\|_{B^s_{\infty,\infty}} < \infty$, and Theorem 5.1.

For any $\alpha > 0$ we have

$$\begin{aligned}\mathbb{E}\Big(\sum_j 2^{j\alpha}\|\Psi_j(\sqrt{A})\tilde{Z}_\bullet(\omega)\|_\infty\Big) &= \sum_j 2^{j\alpha}\mathbb{E}\big(\|\Psi_j(\sqrt{A})\tilde{Z}_\bullet(\omega)\|_\infty\big)\\ &\sim \sum_j 2^{j\alpha}\mathbb{E}\big(\sup_{\xi\in\mathcal{X}_j} |\big(\Psi_j(\sqrt{A})\tilde{Z}_\bullet(\omega)\big)(\xi)|\big)\\ &\le c\sum_j 2^{j\alpha}2^{-js/2}(1+\log(\mathrm{card}(\mathcal{X}_j))^{1/2}.\end{aligned}$$

Above for the equivalence we used (5.3) and for the last inequality the following well known inequality (see e.g. [21, Lemma 2.3.4] or [32, lemma 10.1]): If $Z_1,\dots,Z_N$ are centered Gaussian variables (with arbitrary variances), then

$$\mathbb{E}\big(\max_{1\le k\le N} |Z_k|\big) \le c(1+\log N)^{1/2}\max_k \big(\mathbb{E}|Z_k|^2\big)^{1/2}.$$

By (3.14), we have $\mathrm{card}(\mathcal{X}_j) \le c2^{jd}$. Therefore, if $\alpha < \frac{s}{2}$, then

$$\sum_j 2^{j\alpha}2^{-js/2}(1+\log(\mathrm{card}(\mathcal{X}_j))^{1/2} \le c\sum_j 2^{-j(s/2-\alpha)}\big(\log(c2^{jd})\big)^{1/2} < \infty.$$

Consequently, $\mathbb{E}\Big(\sum_j 2^{j\alpha}\|\Psi_j(\sqrt{A})Z_\bullet(\omega)\|_\infty\Big) < \infty$ and hence $x \mapsto \tilde{Z}_x(\omega) \in B^\alpha_{\infty,1}$, $0 < \alpha < s/2$, ω-a.s.

(b) Suppose now that $\omega - a.e., x \mapsto Z_x(\omega) \in B^\alpha_{\infty,\infty}$, $\alpha > 0$. Then by (5.5) and (3.19)

$$\sup_j 2^{j\alpha}\Big\|\sum_k \Psi_j(\sqrt{\lambda_k})\sqrt{\nu_k}u_k(x)B_k(\omega)\Big\|_\infty < \infty, \quad \omega - \text{a.s.}$$

By (5.3) this is equivalent to

$$\sup_j 2^{j\alpha}\max_{\xi\in\mathcal{X}_j}\Big|\sum_k \Psi_j(\sqrt{\lambda_k})\sqrt{\nu_k}u_k(\xi)B_k(\omega)\Big| < \infty, \quad \omega - \text{a.s.} \tag{5.6}$$

However, $\{2^{j\alpha}\sum_k \Psi_j(\sqrt{\lambda_k})\sqrt{\nu_k}u_k(\xi)B_k(\omega)\}_{j\in\mathbb{N},\xi\in\mathcal{X}_j}$ is a countable set of Gaussian centered variables. The Borell-Ibragimov-Sudakov-Tsirelson theorem (see e.g. [29], §7), in particular, asserts that if $(G_t)_{t\in T}$ is a centered Gaussian process indexed by a countable parameter set T and $\sup_{t\in T} G_t < \infty$ almost surely, then $\sup_{t\in T}\mathbb{E}(G_t^2) < \infty$. Consequently, (5.6) implies

$$\sup_{j\in\mathbb{N},\xi\in\mathcal{X}_j}\mathbb{E}\Big(2^{j\alpha}\sum_k \Psi_j(\sqrt{\lambda_k})\sqrt{\nu_k}u_k(\xi)B_k\Big)^2 < \infty.$$

Therefore, there exists a constant $C > 0$ such that

$$\max_{\xi\in\mathcal{X}_j}\sum_k \Psi_j^2(\sqrt{\lambda_k})\nu_k u_k^2(\xi) \le C2^{-2j\alpha}.$$

But as before, this yields

$$\max_{\xi\in\mathcal{X}_0}\sum_{k:\sqrt{\lambda_k}\le 1}\nu_k u_k^2(\xi) \le \max_{\xi\in\mathcal{X}_0}\sum_k \Psi_0^2(\sqrt{\lambda_k})\nu_k u_k^2(\xi)$$

and, for $j \ge 1$,

$$\begin{aligned}\max_{\xi\in\mathcal{X}_j}\sum_{k,2^{j-1}\le\sqrt{\lambda_k}\le 2^j}\nu_k u_k^2(\xi) &\le 2\max_{\xi\in\mathcal{X}_j}\sum_k \Psi_{j-1}^2(\sqrt{\lambda_k})\nu_k u_k^2(\xi)\\ &\quad + 2\max_{\xi\in\mathcal{X}_j}\sum_k \Psi_j^2(\sqrt{\lambda_k})\nu_k u_k^2(\xi) \le c2^{-2j\alpha}.\end{aligned}$$

Here we used that $\Psi_{j-1}(\lambda) + \Psi_j(\lambda) = 1$ for $\lambda \in [2^{j-1}, 2^j]$, implying $\Psi_{j-1}^2(\lambda) + \Psi_j^2(\lambda) \ge 1/2$.

Finally, applying Proposition 5.4 we conclude from above that

$$\sup_{x\in M}\|K(x,\bullet)\|_{B^{2\alpha}_{\infty,\infty}} < \infty. \quad \square$$

5.3. **Proof of Theorem 4.7.** We begin with the following

Lemma 5.6. *Assume $s > 0$ and $1 \le p \le \infty$, and let Ψ_j, $j \ge 0$, be the functions from* (3.17). *Then*

$$f \in B^s_{p,1} \Longleftrightarrow \sum_{j\ge 0}\|\Psi_j(\sqrt{A})f\|_{B^s_{p,1}} < \infty \quad \text{and} \quad \|f\|_{B^s_{p,1}} \sim \sum_{j\ge 0}\|\Psi_j(\sqrt{A})f\|_{B^s_{p,1}}.$$

Proof. From (3.18) we have for any $f \in L^p$

$$f = \sum_{j\ge 0}\Psi_j(\sqrt{A})f, \quad \forall f \in L^p, \tag{5.7}$$

implying $\|f\|_{B^s_{p,1}} \le \sum_{j\ge 0} \|\Psi_j(\sqrt{A})f\|_{B^s_{p,1}}$.

For the estimate in the other direction, note that by (3.19)

$$\|\Psi_j(\sqrt{A})f\|_{B^s_{p,1}} \sim \sum_{\ell\ge 0} 2^{\ell s}\|\Psi_\ell(\sqrt{A})\Psi_j(\sqrt{A})f\|_p.$$

However, $\operatorname{supp}\Psi_j \cap \mathbb{R}_+ \subset [2^{j-1}, 2^{j+1}]$, $j \ge 1$, and hence $\Psi_\ell(\sqrt{A})\Psi_j(\sqrt{A}) = 0$ if $|\ell - j| > 1$. Therefore,

$$\|\Psi_j(\sqrt{A})f\|_{B^s_{p,1}} \sim \sum_{j-1\le \ell\le j+1} 2^{\ell s}\|\Psi_\ell(\sqrt{A})\Psi_j(\sqrt{A})f\|_p.$$

On the other hand, by estimate (3.12) it follows that $\|\Psi_j(\sqrt{A})g\|_p \le c\|g\|_p$, $\forall g \in L^p$, and hence $\|\Psi_\ell(\sqrt{A})\Psi_j(\sqrt{A})f\|_p \le c\|\Psi_j(\sqrt{A})f\|_p$, implying

$$\|\Psi_j(\sqrt{A})f\|_{B^s_{p,1}} \le c2^{js}\|\Psi_j(\sqrt{A})f\|_p.$$

This in turn leads to

$$\sum_{j\ge 0}\|\Psi_j(\sqrt{A})f\|_{B^s_{p,1}} \le c\sum_{j\ge 0} 2^{js}\|\Psi_j(\sqrt{A})f\|_p \le c\|f\|_{B^s_{p,1}}.$$

The proof is complete. □

We now precise Theorem 4.5 with the following

Proposition 5.7. *Under the hypotheses of Theorem 4.5 and with the functions Ψ_j, $j \ge 0$, from (3.17), if $\sup_{x\in M} \|K(x,\bullet)\|_{B^s_{\infty,\infty}} < \infty$, then*

$$\mathbb{E}\Big(\sum_{j\ge 0}\|\Psi_j(\sqrt{A})Z_\bullet(\omega))\|_{B^\alpha_{\infty,1}}\Big) \sim \mathbb{E}\Big(\sum_{j\ge 0} 2^{j\alpha}\|\Psi_j(\sqrt{A})Z_\bullet(\omega)\|_\infty\Big) < \infty, \tag{5.8}$$

the map

$$I : \omega \in \Omega \mapsto \sum_j \psi_j(\sqrt{A})Z_\bullet(\omega)(\cdot) \in B^\alpha_{\infty,1}$$

is measurable, the series converges in the norm of $B^\alpha_{\infty,1}$, and the image probability Q on $B^\alpha_{\infty,1}$ satisfies:

$$\omega \in B^\alpha_{\infty,1} \xrightarrow{\delta_x} \omega(x)$$

is a centered Gaussian process with covariance $K(x,y)$.

Proof. The equivalence (5.8) follows from the proof of Theorem 4.5, (a) and Lemma 5.6.

As is well known, for any Banach space B with a measure space $(\Omega, \mathcal{B})$, if G is a finite set of indices $b_i \in B$ and $X_i(\omega)$ are real-valued measurable functions, then $\omega \mapsto \sum_{i\in G} X_i(\omega)b_i$ is measurable from Ω to B. Hence,

$$\omega \in \Omega \mapsto \Psi_j(\sqrt{A})Z_\bullet(\omega) = \sum_k \Psi_j(\sqrt{\lambda_k})\sqrt{\nu}_k u_k(\bullet)B_k(\omega) \in B^\alpha_{\infty,1}$$

is measurable. Consequently, by almost everywhere convergence

$$I : \omega \in \Omega \mapsto \sum_j \Psi_j(\sqrt{A})Z_\bullet(\omega)(\cdot) \in B^\alpha_{\infty,1}$$

is also measurable, and $I^*(P) = Q$ is a probability measure on the Borelian sigma-algebra such that under Q the family of random variables δ_x

$$\omega \in B^\alpha_{\infty,1} \xrightarrow{\delta_x} \omega(x)$$

is a centered Gaussian process with covariance $K(x,y) = \int_{B^\alpha_{\infty,1}} \omega(x)\omega(y)dQ(\omega)$. $\square$

Finally, Theorem 4.7 holds due the fact that $B^\alpha_{\infty,1}$ is separable. (see Appendix II). It also proves Part (b) of Theorem 4.5.

5.4. **Proof of theorem 4.8.** Suppose that $\sup_{x\in M}\|K(x,\bullet)\|_{B^s_{\infty,\infty}} < \infty$ and let $f(x) = \sum_{k\in\mathbb{N}(\nu)} \alpha_k\sqrt{\nu_k}\, u_k(x)$, where $(\alpha_k)\in\ell^2$. Then

$$\Psi_j(\sqrt{A})f(x) = \sum_{k\in\mathbb{N}(\nu)} \Psi_j(\sqrt{\lambda_k})\alpha_k\sqrt{\nu_k}\; u_k(x),$$

implying, for $j\ge 1$,

$$\begin{aligned}|\Psi_j(\sqrt{A})f(x)| &\le \Big(\sum_{k\in\mathbb{N}(\nu)}|\alpha_k|^2\Big)^{\frac12}\Big(\sum_{k\in\mathbb{N}(\nu)}|\Psi_j(\sqrt{\lambda_k})|^2\nu_k|u_k(x)|^2\Big)^{\frac12}\\ &\le \|f\|_{\mathbb{H}_K}\Big(\sum_{k:2^{j-1}\le\lambda_k\le 2^{j+1}}\nu_k|u_k(x)|^2\Big)^{\frac12} \le c\|f\|_{\mathbb{H}_K}2^{-js/2},\end{aligned}$$

where for the last inequality we used the assumption and Theorem 5.1. Similarly $|\Psi_0(\sqrt{A})f(x)|\le c\|f\|_{\mathbb{H}_K}$. Therefore, in light of (3.19),

$$\|f\|_{B^{\frac{s}{2}}_{\infty,\infty}} \le c\|f\|_{\mathbb{H}_K}. \tag{5.9}$$

Assume that (5.9) holds. Then for every sequence $(\alpha_k)\in\ell^2$ with $\|(\alpha_k)\|_{\ell^2}\le 1$ we have

$$\Big|\sum_{k\in\mathbb{N}(\nu)}\Psi_j(\sqrt{\lambda_k})\alpha_k\sqrt{\nu_k}\; u_k(x)\Big| \le c2^{-js/2},\quad \forall x\in M,$$

which by duality implies

$$\Big(\sum_{k\in\mathbb{N}(\nu)}|\psi_j(\sqrt{\lambda_k})|^2\nu_k\;|u_k(x)|^2\Big)^{\frac12}\le c2^{-js/2},\quad j\ge 0.$$

Just as in the proof of Theorem 5.1 we get for $j\ge1$

$$\begin{aligned}\sum_{k:2^{j-1}\le\sqrt{\lambda_k}\le 2^j}\nu_k u_k^2(x) &\le \sum_{k\in\mathbb{N}(\nu)}|\Psi_{j-1}(\sqrt{\lambda_k})|^2\nu_k\;|u_k(x)|^2\\ &\quad+\sum_{k\in\mathbb{N}(\nu)}|\Psi_j(\sqrt{\lambda_k})|^2\nu_k\;|u_k(x)|^2\le c2^{-js}\end{aligned}$$

and similarly $\sum_{k:\sqrt{\lambda_k}\le1}\nu_k u_k^2(x)\le c$. Consequently, $\sup_{x\in M}\|K(x,\bullet)\|_{B^s_{\infty,\infty}}<\infty$. $\square$

Remark 5.8. *Let $f\in L^2(M,\mu)$. Clearly*

$$\tilde f(\omega):\omega\in W = B^\alpha_{\infty,1}\mapsto\int_M f(x)\omega(x)d\mu(x)$$

belongs to W^*. Hence, under Q_α, $\tilde{f}$ is a Gaussian variable and

$$\begin{aligned}\mathbb{E}(\tilde{f})^2 &= \int_W \Big(\int_M f(x)\omega(x)d\mu(x)\Big)^2 dQ_\alpha(\omega)\\ &= \int_W \int_M f(x)\omega(x)d\mu(x)\int_M f(y)\omega(y)d\mu(y)dQ_\alpha(\omega)\\ &= \int_M\int_M f(x)f(y)\Big(\int_W \omega(x)\omega(y)dQ_\alpha(\omega)\Big)d\mu(x)d\mu(y) = \langle Kf, f\rangle_{L^2(M,\mu)}.\end{aligned}$$

Consequently,

$$\int_W e^{i\tilde{f}(\omega)}dQ_\alpha(\omega) = e^{-\frac{1}{2}\langle Kf,f\rangle_{L^2(M,\mu)}}.$$

6. Positive and negative definite functions on compact homogeneous spaces

For reader's convenience we recall the basics of the general theory of positive definite (P.D.) and negative definite (N.D.) functions in Appendix I. Here we present some basic facts about positive and negative definite kernels in the general setting of compact two point homogeneous spaces. In the next section, we use these results and Theorem 4.5 to establish the Besov regularity of Gaussian processes indexed by the sphere.

6.1. **Group acting on a space.** Let (M,μ) be a compact space equipped with a positive Radon measure μ. Assume that there exists a group G acting transitively on (M,μ), that is, there exists a map $(g,x)\in G\times M \mapsto g\cdot x \in M$ such that

1. $h\cdot(g\cdot x) = (hg)\cdot x$, $\forall g,h\in G$,
2. $\exists e\in G$ s.t. $e\cdot x = x$, $\forall x\in M$ (e is the neutral element in G),
3. $\forall x,y\in M$, $\exists g\in G$ s.t. $g\cdot x = y$ (transitivity),
4. $\int_M (\gamma(g)f)(x)d\mu(x) = \int_M f(g^{-1}\cdot x)d\mu(x) = \int_M f(x)d\mu(x)$ $\forall g\in G$, $\forall f\in L^1$, where $(\gamma(g)f)(x) := f(g^{-1}\cdot x)$. Hence, $(\gamma(g))_{g\in G}$ is a group of isometry of L^1.

Definition 6.1. *A continuous real-valued kernel $K(x,y)$ on $M\times M$ is said to be G-invariant if*

$$K(g\cdot x, g\cdot y) = K(x,y), \quad \forall g\in G,\ \forall x,y\in M.$$

If K is the operator on L^2 with kernel $K(x,y)$, then K is called $G-$invariant if $\gamma(g)K = K\gamma(g)$, $\forall g\in G$, that is,

$$\int_M K(g^{-1}\cdot x, y)f(y)d\mu(y) = \int_M K(x,y)f(g^{-1}\cdot y)d\mu(y), \quad \forall f\in L^2.$$

Remark 6.2. *(a) If $K(x,y)$ is a continuous G-invariant kernel, then*

(i) $K(x,x) = K(g\cdot x, g\cdot x)$ and hence $K(x,x)\equiv |M|^{-1}\,\mathrm{Tr}(K)$, and

(ii)

$$\int_M K(x,y)d\mu(y) = \int_M K(x,g\cdot y)d\mu(y) = \int_M K(g^{-1}\cdot x,y)d\mu(y), \quad \forall g\in G,$$

and hence $\mathbb{1} := \mathbb{1}_M$ is an eigenfunction of K, i.e.,

$$\int_M K(x,y)\mathbb{1}(y)d\mu(y) = \lambda\mathbb{1}(x), \quad \int_M\int_M K(x,y)d\mu(x)d\mu(y) = \lambda|M|.$$

(b) If $K(x,y)$ is a continuous positive G-invariant kernel, then

$$\begin{aligned}\psi_K(x,y) &:= K(x,x) + K(y,y) - 2K(x,y)\\ &= 2(C - K(x,y)) = 2(|M|^{-1}\operatorname{Tr}(K) - K(x,y))\end{aligned}$$

is G-invariant and by (8.3),

$$\tilde{K}(x,y) = K(x,y) + |M|^{-1}(\operatorname{Tr}(K) - 2C').$$

(c) Suppose $\psi(x,y)$ is a $G-$invariant N.D. kernel and consider the associated P.D. kernel $\tilde{K}$, defined as in (8.2). *Then $\tilde{K}(x,y)$ is G-invariant, and*

$$x \mapsto \frac{1}{|M|}\int_M \psi(x,u)d\mu(u) \equiv C_0 \text{ and } \tilde{K}(x,y) = C_0 - \frac{1}{2}\psi(x,y).$$

Thus, in this framework there is one-to-one correspondence up to a constant between invariant P.D. and N.D. kernels.

6.2. **Composition of operators.** Let $K(x,y)$ and $H(x,y)$ be two continuous kernels on $M \times M$ as above, and let K and H be the associate operators. The operator $K \circ H$ is also a kernel operator with kernel $K \circ H(x,y)$:

$$K \circ H(x,y) = \int_M K(x,u)H(u,y)d\mu(u).$$

Observe that:

(1) If $K(x,y) = K(y,x), H(x,y) = H(y,x)$ then

$$K \circ H(x,y) = \int_M K(x,u)H(u,y)d\mu(u) = \int_M H(y,u)K(u,x)d\mu(u) = H \circ K(y,x).$$

(2) If $K(x,y)$ and $H(x,y)$ are $G-$invariant, then so is $K \circ H$. Indeed,

$$\begin{aligned}K \circ H(g\cdot x, g\cdot y) &= \int_M K(g\cdot x,u)H(u,g\cdot y)d\mu(u)\\ &= \int_M K(g\cdot x, g\cdot u)H(g\cdot u, g\cdot y)d\mu(u)\\ &= \int_M K(x,u)H(u,y)d\mu(u) = K\circ H(x,y).\end{aligned}$$

6.3. **Group action and metric.** Assume that we are in the setting of a Dirichlet space defined through a non-negative self-adjoint operator on $L^2(M,\mu)$ just as in §3.1. Suppose now that,

$$\gamma(g)A = A\gamma(g), \quad \forall g \in G$$

or equivalently

$$\gamma(g)P_t = P_t\gamma(g), \quad \forall t > 0,\ \forall g \in G,$$

i.e. $\forall t > 0,\ p_t(x,y)$ is $G-$invariant. Clearly $\Gamma(f_1,f_2)$ is also G-invariant: $\Gamma(f_1,f_2) = \Gamma(\gamma(g)f_1, \gamma(g)f_2)$ and the associate metric $\rho(x,y)$ is G-invariant:

$$\rho(g\cdot x, g\cdot y) = \rho(x,y), \quad \forall g \in G.$$

Definition 6.3. *In the current framework, (M,μ,A,ρ,G) is said to be a two point homogeneous space if*

$$\forall x,y,x',y' \in M \text{ s.t. } \rho(x,y) = \rho(x',y'),\ \exists g \in G \text{ s.t. } g\cdot x = x',\ g\cdot y = y'.$$

In particular, $\forall (x,y) \in M\times M,\ \exists g \in G$ s.t. $g\cdot x = y,\ g\cdot y = x$.

Theorem 6.4. *Let (M, μ, A, ρ, G) be a compact two point homogeneous space. Then*

(1) *Any G-invariant continuous kernel $K(x,y)$ is symmetric.*

(2) *If $K(x,y)$ and $H(x,y)$ are two G-invariant continuous kernels, then $K \circ H = H \circ K$. In particular, if $K(x,y)$ is a G-invariant continuous kernel, then $KA = AK$.*

(3) *Any G-invariant real-valued continuous kernel $K(x,y)$ depends only on the distance $\rho(x,y)$, that is, there exist a continuous function $k : \mathbb{R} \mapsto \mathbb{R}$, such that*

$$K(x,y) = k(\rho(x,y)), \quad \forall x, y \in M.$$

This theorem is a straightforward consequence of the observations from §6.2 and the definition of two point homogeneous spaces.

Let now M be a compact Riemannian manifold and assume that $A := -\Delta_M$ is the Laplacian on M, ρ is the Riemannian metric, and μ is the Riemannian measure. Also, assume that there exists a compact Lie group G of isometries on M such that $(M, \mu, -\Delta_M, \rho, G)$ is a compact two point homogeneous space. For the connection of the above setting with Gaussian processes, see [6], [20].

Let $0 \le \lambda_1 < \lambda_2 < \cdots$ be the spectrum of $-\Delta_M$. Then the eigenspaces $\mathcal{H}_{\lambda_k} := \mathrm{Ker}(\Delta_M + \lambda_k \,\mathrm{Id})$ are finite dimensional and

$$L^2(M, \mu) = \bigoplus_{k \ge 1} \mathcal{H}_{\lambda_k}.$$

Let $P_{\mathcal{H}_{\lambda_k}}(x,y)$ be the kernel of the orthogonal projector onto $\mathcal{H}_{\lambda_k}$. Then if $K(x,y)$ is a G-invariant positive definite kernel we have the following decomposition of $K(x,y)$, which follows from Bochner-Godement theorem ([17], [23]):

$$K(x,y) = \sum_{k \ge 0} \nu_k P_{\mathcal{H}_{\lambda_k}}(x,y), \quad \nu_k \ge 0.$$

7. Gaussian process on the sphere

In this section we apply our main result (Theorem 4.5) to a Gaussian process parametrized by the unit sphere $\mathbb{S}^d$ in $\mathbb{R}^{d+1}$. This is a Riemannian manifold and a compact two point homogeneous space. More explicitly,

$$G = SO(d+1), \ H = SO(d), \ G/H = \mathbb{S}^d.$$

The geodesic distance ρ on $\mathbb{S}^d$ is given by

$$\rho(\xi, \eta) = \arccos\langle \xi, \eta \rangle,$$

where $\langle \xi, \eta \rangle$ is the inner product of $\xi, \eta \in \mathbb{R}^{d+1}$. Clearly,

$$\forall \xi, \eta \in \mathbb{S}^d, \ \forall g \in G, \ \ \rho(g \cdot \xi, g \cdot \eta) = \rho(\xi, \eta), \quad \text{and} \quad \forall \xi, \eta \in \mathbb{S}^d, \ \exists g \in G \ \text{ s.t. } \ g \cdot \xi = \eta.$$

Thus G acts isometrically and transitively on $\mathbb{S}^d$. Furthermore, $\forall \xi, \eta, \xi', \eta' \in \mathbb{S}^d$ s.t. $\rho(\xi, \eta) = \rho(\xi', \eta')$ there exists $g \in G$ s.t. $g \cdot \xi = \xi'$ and $g \cdot \eta = \eta'$. Therefore, $\mathbb{S}^d$ is a compact two point homogeneous space.

Let $-\Delta_{\mathbb{S}^d}$ be the (positive) Laplace-Beltrami operator on $\mathbb{S}^d$. As is well known the eigenspaces of $-\Delta_{\mathbb{S}^d}$ are the spaces of spherical harmonics, defined by

$$\mathcal{H}_{\lambda_k} := \mathrm{Ker}(\Delta_{\mathbb{S}^d} + \lambda_k I_d), \quad \lambda_k := k(k+d-1) = k(k+2\nu), \ \ k \ge 0 \quad \nu := \frac{d-1}{2}.$$

One has $L^2(\mathbb{S}^d) = \bigoplus_{k\geq 0} \mathcal{H}_{\lambda_k}$ and the kernel of the orthogonal projector $P_{\mathcal{H}_{\lambda_k}}$ onto $\mathcal{H}_{\lambda_k}$ is given by

$$P_{\mathcal{H}_{\lambda_k}}(\xi,\eta) = L_k^d(\langle \xi,\eta\rangle), \quad L_k^d(x) := |\mathbb{S}^d|^{-1}\Big(1+\frac{k}{\nu}\Big)C_k^\nu(x).$$

Here $C_k^\nu(x)$, $k\geq 0$, are the Gegenbauer polynomials defined on $[-1,1]$ by the generating function

$$\frac{1}{(1-2xr+r^2)^\nu} = \sum_{k\geq 0} r^k C_k^\nu(x).$$

Therefore,

$$-\Delta_{\mathbb{S}^d} f = \sum_{k\geq 0} k(k+2\nu) P_{\mathcal{H}_{\lambda_k}} f$$

and the invariant continuous positive definite functions on $\mathbb{S}^d$ are of the form

$$K(\xi,\eta) = \sum_k \nu_k L_k^d(\langle\xi,\eta\rangle) = \sum_k \nu_k L_k^d(\cos\rho(\xi,\eta)),$$

where

$$\sum_k \nu_k L_k^d(1) = \sum_k \nu_k L_k^d(\langle \xi,\xi\rangle) < \infty.$$

Note that

$$L_k^\nu(1)|\mathbb{S}^d| = \int_{\mathbb{S}^d} L_k^\nu(\langle\xi,\xi\rangle)d\mu(\xi) = \dim(\mathcal{H}_{\lambda_k}(\mathbb{S}^d)) = \binom{k+d}{d} - \binom{k-2+d}{d} \sim k^{d-1}.$$

Let

$$W_k^\nu(x) := \frac{L_k^\nu(x)}{L_k^\nu(1)} = \frac{C_k^\nu(x)}{C_k^\nu(1)}. \quad \text{Clearly, } W_k^\nu(1) = \sup_{x\in[-1,1]} |W_k^\nu(x)| = 1.$$

Then (see [8])

$$\lim_{\nu\mapsto 0} \frac{C_k^\nu(x)}{C_k^\nu(1)} = T_k(x) \ \ (= W_k^0(x) \text{ by convention}),$$

$$\lim_{\nu\mapsto\infty} \frac{C_k^\nu(x)}{C_k^\nu(1)} = x^k \ \ (= W_k^\infty(x) \text{ by convention}).$$

Here T_k is the Chebyshev polynomial of first kind ($T_k(\cos\theta) = \cos k\theta$). The invariant continuous positive definite functions on $\mathbb{S}^d$ are of the form

$$K^\nu(\xi,\eta) = \sum_{k\geq 0} a_k^\nu W_k^\nu(\langle\xi,\eta\rangle) = \sum_{k\geq 0} a_k^\nu W_k^\nu(\cos\rho(\xi,\eta)), \quad a_k^\nu \geq 0, \quad \sum_k a_k^\nu < \infty.$$

Clearly,

$$\sum_k a_k^\nu W_k^\nu(\cos\rho(\xi,\eta)) = \sum_k \frac{a_k^\nu}{L_k^\nu(1)} L_k^\nu(\cos\rho(\xi,\eta)), \quad L_k^\nu(1) \sim k^{d-1}. \tag{7.1}$$

Therefore,

$$\nu_k = |\mathbb{S}^d| \frac{a_k^\nu}{\dim(\mathcal{H}_{\lambda_k})} = O\Big(\frac{a_k^\nu}{k^{d-1}}\Big).$$

The following **Schoenberg-Bingham result** (see e.g. [8]) plays a key role here: *If f is a continuous function defined on $[-1,1]$, then $f(\langle\xi,\eta\rangle)$ is a positive definite function on $\mathbb{S}^d$ and invariant with respect to $SO(d+1)$ for all $d\in\mathbb{N}$ if and only if*

$$f(x) = \sum_{n\geq 0} a_n x^n, \quad \textit{where} \quad a_n \geq 0 \quad \textit{and} \quad \sum_{n\geq 0} a_n = f(1) < \infty.$$

Therefore, for such a function f

$$f(x)=\sum_{k\ge 0}a_k^\nu W_k^\nu(x),\quad a_k^\nu\ge 0,\quad\text{and}\quad \sum_{k\ge 0}a_k^\nu=\sum_{k\ge 0}a_k=f(1),$$

and hence

$$f(\langle\xi,\eta\rangle)=\sum_{k\ge 0}a_k^\nu W_k^\nu(\langle\xi,\eta\rangle)=\sum_{k\ge 0}\frac{a_k^\nu}{L_k^\nu(1)}L_k^\nu(\langle\xi,\eta\rangle)=f(\cos\rho(\xi,\eta)).$$

7.1. **Fractional Brownian process on the sphere.**

Theorem 7.1. *For any* $0<\alpha\le 1$ *the function*

$$\psi(\xi,\eta)=\rho(\xi,\eta)^\alpha,\quad \xi,\eta\in\mathbb{S}^d,$$

is negative definite, and the associated Gaussian process has almost everywhere regularity $B^\gamma_{\infty,1}$, $\gamma<\frac{\alpha}{2}$.

Proof. Consider first the case when $\alpha=1$ (*Brownian process*). We will show that for some constant $C>0$ the function $C-\rho(\xi,\eta)$ is an invariant positive definite function. To this end, by Schoenberg-Bingham result we have to prove that there exists a function

$$f(x)=\sum a_nx^n,\quad\text{with } a_n\ge 0,\ \sum_{n\ge 0}a_n<\infty,$$

such that $f(\langle\xi,\eta\rangle)=f(\cos\rho(\xi,\eta))=C-\rho(\xi,\eta)$. Luckily the function $\frac{\pi}{2}-\arccos x$ does the job. Indeed, it is easy to see that

$$f(x):=\frac{\pi}{2}-\arccos x=\arcsin x=\sum_{j\ge 0}\frac{(\frac12)_j(\frac12)_j}{j!(\frac32)_j}x^{2j+1}$$

and

$$\sum_{j\ge 0}\frac{(\frac12)_j(\frac12)_j}{j!(\frac32)_j}=\frac{\pi}{2}\quad\text{(Gauss)}.$$

Here we use the standard notation $(a)_j:=a(a+1)\cdots(a+j-1)=\Gamma(a+j)/\Gamma(a)$. Therefore,

$$f(\langle\xi,\eta\rangle)=\frac{\pi}{2}-\arccos\langle\xi,\eta\rangle=\frac{\pi}{2}-\rho(\xi,\eta).$$

Clearly, $|f(\langle\xi,\eta\rangle)-f(\langle\xi,\eta'\rangle)|\le\rho(\eta,\eta')$ and by Theorem 4.5 the associated Gaussian process $(Z^d_\xi(\omega))_{\xi\in\mathbb{S}^d}$ is almost surely in $B^s_{\infty,1}(\mathbb{S}^d)$ (hence in $\operatorname{Lip} s$) for $0<s<\frac12$. Furthermore,

$$\mathbb{E}(Z^d_\xi-Z^d_\eta)^2=2f(1)-2f(\langle\xi,\eta\rangle)=2\rho(\xi,\eta).$$

Consider now the general case: $0<\alpha\le 1$ (*Fractional Brownian process*). From above it follows that $\psi(\xi,\eta):=\rho(\xi,\eta)$ is an invariant negative definite kernel. Then the general theory of negative definite kernels yields that for any $0<\alpha\le 1$ the kernel $\psi_\alpha(\xi,\eta)=\rho(\xi,\eta)^\alpha$ is invariant and negative definite. Therefore, for a sufficiently large constant $C>0$,

$$K(\xi,\eta)=C-\frac12\rho(\xi,\eta)^\alpha$$

is an invariant positive definite kernel. On the other hand,

$$|K(\xi,\eta)-K(\xi,\eta')|=\frac12|\rho(\xi,\eta)^\alpha-\rho(\xi,\eta')^\alpha|\le\frac12\rho(\eta',\eta))^\alpha.$$

By Theorem 4.5 it follows that the associated Gaussian process $(Z^d_\xi(\omega))_{\xi\in\mathbb{S}^d}$ is almost surely in $B^\gamma_{\infty,1}$, $\gamma<\frac{\alpha}{2}$, and hence in $\operatorname{Lip} s$, $s<\frac{\alpha}{2}$, and the proof is complete. □

Remark 7.2. *From the definition of the process, we have*

$$\mathbb{E}(Z^\alpha_\xi - Z^\alpha_\eta)^2 = \rho(\xi,\eta)^\alpha.$$

This directly connects to the regularity proof of such a process using generalization of Kolmogorov-Csensov inequalities. See for instance [3] *and* [27].

Remark 7.3. *If $\alpha>1$, then $\rho(\xi,\eta)^\alpha$ is no more a negative definite function on the sphere $\mathbb{S}^d$. In fact, to prove such a result, it is suffices to prove it for $\mathbb{S}^1$, as the closed geodesic of $\mathbb{S}^d$ are isometric to $\mathbb{S}^1$. As $\mathbb{S}^1$ is a commutative group, one can apply the Bochner theorem: $K(x-y)$ is a positive definite function if and only if the Fourier coefficient of K are nonnegative.*

Let $\alpha>0$ and let ϕ be the 2π-periodical function, such that for $x\in[-\pi,\pi]$, $\phi(x)=|x|^\alpha$, so that on $\mathbb{S}^1=\mathbb{R}/2\pi\mathbb{Z}$, $\phi(x-y)=d_{\mathbb{S}^1}(x,y)^\alpha$. Clearly, for any $k\in\mathbb{Z}$

$$\hat{\phi}(k)=\frac{1}{2\pi}\int_{-\pi}^{\pi}|x|^\alpha e^{-ikx}dx=\frac{1}{\pi}\int_0^\pi x^\alpha\cos kx dx.$$

Integrating by parts we obtain, for $k\geq 1$,

$$\int_0^\pi x^\alpha \cos kx dx = -\frac{\alpha}{k}\int_0^\pi x^{\alpha-1}\sin kx dx = -\frac{\alpha}{k^{\alpha+1}}\int_0^{k\pi} u^{\alpha-1}\sin u du$$

and in going further

$$\int_0^{k\pi} u^{\alpha-1}\sin u du=\sum_{j=0}^{k-1}\int_{j\pi}^{(j+1\pi} u^{\alpha-1}\sin u du = \int_0^\pi \sum_{j=0}^{k-1}(-1)^j(u+j\pi)^{\alpha-1}\sin u du.$$

Now, if $\alpha>1$ it is easy to see that for $0<u<\pi$ and $k\geq 1$

$$\sum_{j=0}^{k-1}(-1)^j(u+j\pi)^{\alpha-1}>0 \quad \textit{if} \quad k\equiv 1 \,(\mathrm{mod}\ 2)$$

and

$$\sum_{j=0}^{k-1}(-1)^j(u+j\pi)^{\alpha-1}<0 \quad \textit{if} \quad k\equiv 0 \,(\mathrm{mod}\ 2).$$

Therefore, if $\alpha>1$, then $K(x-y)=C-d_{\mathbb{S}^1}(x,y)^\alpha$ is never a positive definite function.

7.2. **Regularity of Gaussian processes on the sphere: General result.**

Theorem 7.4. *Let*

$$f(x)=\sum_{n\geq 0}\frac{A_n}{n!}x^n, \quad \textit{where} \quad A_n\geq 0, \ \ \textit{and} \ \ \frac{A_n}{n!}=O\big(\frac{1}{n^{1+\alpha}}\big), \quad \alpha>0.$$

Then

$$K(\xi,\eta):=f(\cos\langle\xi,\eta\rangle), \ \ \xi,\eta\in\mathbb{S}^d, \ \ d\geq 1,$$

is an invariant positive definite function and the associated Gaussian process $(Z^d_\xi(\omega))_{\xi\in\mathbb{S}^d}$ is almost surely in $B^\gamma_{\infty,1}$ for $\gamma<\alpha$.

Proof. By Corollary 5.2 it suffices to show that $f(x)$ can be represented in the following form (see (7.1)):

$$f(x) = \sum_j B_j W_j^\nu(x), \quad 0 \le B_j = O\Big(\frac{1}{j^{1+2\alpha}}\Big),$$

implying $\nu_j = O\Big(\frac{1}{j^{d+2\alpha}}\Big) = O(\sqrt{\lambda_j})^{2\alpha+d}$. By [8, Lemma 1] and the obvious identity $\Gamma(x+n) = (x)_n\Gamma(x)$ we obtain the representation

$$x^n = \frac{n!}{2^n} \sum_{0\le 2k\le n} \frac{n-2k+\nu}{k!(\nu)_{n-k+1}} \frac{(2\nu)_{n-2k}}{(n-2k)!} W_{n-2k}^\nu(x).$$

Substituting this in the definition of $f(x)$ we obtain

$$\begin{aligned}
f(x) = \sum_{n\ge 0} \frac{A_n}{n!} x^n &= \sum_{n\ge 0} \frac{A_n}{2^n} \sum_{0\le 2k\le n} \frac{n-2k+\nu}{k!(\nu)_{n-k+1}} \frac{(2\nu)_{n-2k}}{(n-2k)!} W_{n-2k}^\nu(x) \\
&= \sum_{j\ge 0} \frac{(j+\nu)(2\nu)_j}{j!} W_j^\nu(x) \sum_{n-2k=j} \frac{A_n}{2^n k!(\nu)_{n-k+1}} \\
&= \sum_{j\ge 0} \frac{(j+\nu)(2\nu)_j}{j!} W_j^\nu(x) \frac{1}{2^j} \sum_{k\ge 0} \frac{A_{j+2k}}{2^{2k} k!(\nu)_{j+k+1}} =: \sum_{j\ge 0} B_j W_j^\nu(x),
\end{aligned}$$

where for the third equality we applied the substitution $j = n-2k$ and shifted the order of summation. We also have

$$\begin{aligned}
B_j &:= \frac{(j+\nu)(2\nu)_j}{j!2^j} \sum_{k\ge 0} \frac{A_{j+2k}}{2^{2k}k!(\nu)_{j+k+1}} \\
&= \frac{(j+\nu)(2\nu)_j}{j!2^j(\nu)_{j+1}} \sum_{k\ge 0} \frac{A_{j+2k}}{2^{2k}k!(\nu+j+1)_k} \\
&= \frac{(2\nu)_j}{2^j j!(\nu)_j} \sum_{k\ge 0} \frac{A_{j+2k}}{2^{2k}k!(\nu+j+1)_k}.
\end{aligned}$$

However, for $n > \alpha$ we have $\frac{c_1(\alpha)}{n^{1+\alpha}} \le \frac{\Gamma(n-\alpha)}{n!} \le \frac{c_2(\alpha)}{n^{1+\alpha}}$ and hence

$$\frac{A_n}{n!} = O\Big(\frac{1}{n^{1+\alpha}}\Big) \iff A_n = O(\Gamma(n-\alpha)).$$

We use this to obtain for $j > \alpha$ (with $c = c(\alpha)$)

$$\begin{aligned}
\sum_{k\ge 0} \frac{A_{j+2k}}{2^{2k}k!(\nu+j+1)_k} &\le c \sum_{k\ge 0} \frac{\Gamma(j+2k-\alpha)}{2^{2k}k!(\nu+j+1)_k} \\
&= c\Gamma(j-\alpha) \sum_{k\ge 0} \frac{\Gamma(j+2k-\alpha)}{\Gamma(j-\alpha)} \frac{1}{2^{2k}k!(\nu+j+1)_k} \\
&= c\Gamma(j-\alpha) \sum_{k\ge 0} \frac{(j-\alpha)_{2k}}{2^{2k}} \frac{1}{k!(\nu+j+1)_k} \\
&= c\Gamma(j-\alpha) \sum_{k\ge 0} \Big(\frac{j-\alpha}{2}\Big)_k \Big(\frac{j-\alpha+1}{2}\Big)_k \frac{1}{k!(\nu+j+1)_k},
\end{aligned}$$

where we used the Legendre duplication formula (see e.g. [4]):

$$\frac{(b)_{2k}}{2^{2k}} = \frac{\Gamma(b+2k)}{2^{2k}\Gamma(b)} = \Big(\frac{b}{2}\Big)_k \Big(\frac{b+1}{2}\Big)_k.$$

By the Gaussian identity (see e.g. [4, Theorem 2.2.2])

$$\begin{aligned}\sum_{k\geq 0} \Big(\frac{j-\alpha}{2}\Big)_k \Big(\frac{j-\alpha+1}{2}\Big)_k & \frac{1}{k!(\nu+j+1)_k} \\ &= \frac{\Gamma(\nu+j+1)\Gamma(\nu+j+1-\frac{j-\alpha}{2}-\frac{j-\alpha+1}{2})}{\Gamma(\nu+j+1-\frac{j-\alpha}{2})\Gamma(\nu+j+1-\frac{j-\alpha+1}{2})} \\ &= \frac{\Gamma(\nu+j+1)\Gamma(\nu+\frac{1}{2}+\alpha)}{\Gamma(\nu+\frac{j}{2}+1+\frac{\alpha}{2})\Gamma(\nu+\frac{j}{2}+\frac{1}{2}+\frac{\alpha}{2})}\end{aligned}$$

and hence

$$B_j \leq c\frac{(2\nu)_j}{j!2^j(\nu)_j}\frac{\Gamma(j-\alpha)\Gamma(\nu+j+1)\Gamma(\nu+\frac{1}{2}+\alpha)}{\Gamma(\nu+\frac{j}{2}+1+\frac{\alpha}{2})\Gamma(\nu+\frac{j}{2}+\frac{1}{2}+\frac{\alpha}{2})}.$$

Applying again the Legendre duplication formula, we get

$$\Gamma\Big(\frac{1}{2}\Big)\Gamma(2\nu+j+1+\alpha) = \Gamma\Big(\nu+\frac{j}{2}+1+\frac{\alpha}{2}\Big)\Gamma\Big(\nu+\frac{j}{2}+\frac{1}{2}+\frac{\alpha}{2}\Big)2^{2\nu+j+\alpha}.$$

We use this above to obtain for $j \geq 2\alpha$

$$\begin{aligned}B_j &\leq c\frac{(2\nu)_j}{j!(\nu)_j}\frac{\Gamma(j-\alpha)\Gamma(\nu+j+1)\Gamma(\nu+\frac{1}{2}+\alpha)}{\Gamma(\frac{1}{2})\Gamma(2\nu+j+1+\alpha)2^{-2\nu-\alpha}} \\ &= c\frac{\Gamma(2\nu+j)\Gamma(\nu)}{\Gamma(j+1)\Gamma(2\nu)\Gamma(\nu+j)}\frac{\Gamma(j-\alpha)\Gamma(\nu+j+1)\Gamma(\nu+\frac{1}{2}+\alpha)}{\Gamma(\frac{1}{2})\Gamma(2\nu+j+1+\alpha)2^{-2\nu-\alpha}} \\ &= c2^{\alpha+1}(j+\nu)\frac{\Gamma(\nu+\frac{1}{2}+\alpha)}{\Gamma(\nu+\frac{1}{2})}\frac{\Gamma(j-\alpha)}{\Gamma(j-\alpha+1+\alpha)}\frac{\Gamma(2\nu+j)}{\Gamma(2\nu+j+1+\alpha)} \\ &\leq c(j+\nu)\frac{1}{(j-\alpha)^{1+\alpha}}\frac{1}{(2\nu+j)^{1+\alpha}} \leq \frac{c}{j^{1+2\alpha}}.\end{aligned}$$

Here we used once again the the Legendre duplication formula. It is easy to show that $B_j \leq c(\alpha)$, if $j < 2\alpha$. Therefore, $B_j = O\big(\frac{1}{j^{1+2\alpha}}\big)$ and this completes the proof. ∎

Corollary 7.5. *Let $a > 0$, $b > 0$, $c > a+b$, $\alpha = c-a-b$, and let*

$$F_{a,b;c}(x) := \sum_n \frac{(a)_n(b)_n}{(c)_n}\frac{x^n}{n!}.$$

Then $F_{a,b;c}(\langle\xi,\eta\rangle)$ is an invariant positive definite function on the sphere $\mathbb{S}^d$ and the associated Gaussian process has regularity $B^{\gamma}_{\infty,1}$, $\gamma < \alpha$, almost everywhere.

8. Appendix

8.1. **Appendix I: Positive and negative definite functions.**

We recall in this appendix some well known (or not so well known) facts about positive definite and negative definite functions. For details we refer the reader to [5], [7] [38], [9], [17].

Recall first the definitions of positive and negative definite functions:

Definition 8.1. *Given a set M, a real-valued function $K(x,y)$ defined on $M\times M$ is said to be* positive definite *(P.D.), if $K(x,y)=K(y,x)$, $\forall x,y\in M$, and*

$$\forall \alpha_1,\dots,\alpha_n\in\mathbb{R},\ \forall x_1,\dots,x_n\in M,\quad \sum_{i,j=1}^{n}\alpha_i\alpha_j K(x_i,x_j)\ge 0.$$

Clearly, if $K(x,y)$ is P.D. then $|K(x,y)|\le\sqrt{K(x,x)}\sqrt{K(y,y)}$. It is well known that the following characterization is valid:

$$K(x,y) \text{ is P.D.} \quad\Longleftrightarrow\quad K(x,y)=\mathbb{E}(Z_xZ_y),$$

where $(Z_x)_{x\in M}$ is some (centered) Gaussian process.

Definition 8.2. *For any $u\in M$ we associate to $K(x,y)$ the following P.D. function*

$$K_u(x,y):=K(x,y)+K(u,u)-K(x,u)-K(y,u)=\mathbb{E}[(Z_x-Z_u)(Z_y-Z_u)],$$

where (Z_x-Z_u) is the process "killed" at the point $u\in M$.

Clearly,

$$K_u\equiv K\Longleftrightarrow K(u,u)=0.$$

Definition 8.3. *Given a set M, a real-valued function $\psi(x,y)$ defined on $M\times M$ is said to be* negative definite *(N.D.), if*

$$\psi(x,y)=\psi(y,x),\ \ \forall x,y\in M,\quad \psi(x,x)\equiv 0,\quad \textit{and}$$

$$\forall \alpha_1,\dots,\alpha_n\in\mathbb{R}\ \ s.t.\ \sum_i\alpha_i=0,\ \ \forall x_1,\dots,x_n\in M,\quad \sum_{i,j=1}^{n}\alpha_i\alpha_j\psi(x_i,x_j)\le 0.$$

The following characterization is valid (see e.g. [7, Proposition 3.2]):

$$\psi(x,y)\ \text{ is N.D.}\quad\Longleftrightarrow\quad \psi(x,y)=\mathbb{E}(Z_x-Z_y)^2,$$

where $(Z_x)_{x\in M}$ is some Gaussian process.

Consequently, if $\psi(x,y)$ is N.D. then $\psi(x,y)\ge 0$, $\forall x,y\in M$, and $\sqrt{\psi(x,y)}$ verifies the triangular inequality.

The following proposition is easy to verified.

Proposition 8.4. (*a*) *Let $K(x,y)$ be a P.D. kernel on a set M, and set*

$$\psi_K(x,y):=K(x,x)+K(y,y)-2K(x,y). \tag{8.1}$$

Then ψ_K is negative definite. The function ψ_K will be termed the N.D. function associated to K. In fact, if $K(x,y)=\mathbb{E}(Z_xZ_y)$, then $\psi_K(x,y)=\mathbb{E}(Z_x-Z_y)^2$. Furthermore, $\psi_K\equiv\psi_{K_u}$, $\forall u\in M$.

(*b*) *Let ψ be a N.D. function, and for any $u\in M$ define*

$$N(u,\psi)(x,y):=\frac{1}{2}[\psi(x,u)+\psi(y,u)-\psi(x,y)].$$

Thus, if $\psi(x,y)=\mathbb{E}(Z_x-Z_y)^2$, then $N(u,\psi)(x,y):=\mathbb{E}\big[(Z_x-Z_u)(Z_y-Z_u)\big]$. Then $N(u,\psi)$ is P.D. Moreover,

$$N(u,\psi_K)=K_u.$$

(*c*) *If K is P.D., then $K(x,y)\equiv$ constant $\Longleftrightarrow\psi_K\equiv 0$.*

Proposition 8.5. *Let $\psi(x,y)$ be a real-valued continuous N.D. function on the compact space M, μ a positive Radon measure, with support M, and set*

$$\tilde{K}(x,y) := \frac{1}{2|M|}\int_M [\psi(x,u)+\psi(y,u)-\psi(x,y)]d\mu(u).$$

Then

(*a*) *$\tilde{K}$ is positive definite, and $\psi_{\tilde{K}} = \psi$.*

(*b*) *$\mathbb{1}$ is an eigenfunction of the operator $\tilde{K}$ with kernel $\tilde{K}(x,y)$, that is,*

$$\int_M \tilde{K}(x,y)d\mu(y) \equiv \tilde{\lambda}, \quad \tilde{\lambda} = \frac{1}{2|M|}\int_M\int_M \psi(u,y)d\mu(u)d\mu(y)(\geq 0).$$

(*c*)

$$\exists z\in M \ \ s.t.\ \ \tilde{K}(z,z)=0 \iff \tilde{K}(x,y)\equiv 0 \iff \psi(x,y)\equiv 0.$$

Proof. Parts (a) and (b) are straightforward. For the proof of (c) we first observe the obvious implications:

$$\psi(x,y)\equiv 0 \implies \tilde{K}(x,y)\equiv 0 \implies \tilde{K}(z,z)=0,\ \forall z\in M.$$

Now, let $\tilde{K}(z,z)=0$ for some $z\in M$. Then

$$\frac{1}{2|M|}\int_M [\psi(z,u)+\psi(z,u)-\psi(z,z)]d\mu(u) = 0.$$

By definition $\psi(z,z)=0$ and hence $\int_M \psi(z,u)d\mu(u)=0$. However, $\psi(z,u)$ is continuous, $\psi(z,u)\geq 0$ and $\operatorname{supp}(\mu)=M$. Therefore, $\psi(z,u)=0$, $\forall u\in M$. Now, by the triangular inequality, we obtain for $x,y\in M$

$$0\leq \sqrt{\psi(x,y)} \leq \sqrt{\psi(x,z)}+\sqrt{\psi(z,y)} = 0,$$

and hence $\psi(x,y)\equiv 0$. This completes the proof. $\square$

Remark 8.6. *One can verify easily that if $K(x,y)$ is P.D. on M, then*

$$\begin{aligned} K_u(x,y) &:= K(x,y)+K(u,u)-K(x,u)-K(y,u) \\ &= \frac{1}{2}[\psi_K(x,u)+\psi_K(y,u)-\psi_K(x,y)]. \end{aligned}$$

The proof of the following proposition is straightforward.

Proposition 8.7. *Let M be a compact space, equipped with a Radon measure μ. Assume that $K(x,y)$ is a continuous P.D. kernel and as previously let:*

$\psi(x,y) := \psi_K(x,y) = K(x,x)+K(y,y)-2K(x,y)$ *be the associated N.D. kernel,*

$$\begin{aligned} K_u(x,y) &:= K(x,y)+K(u,u)-K(x,u)-K(y,u) \\ &= \frac{1}{2}[\psi(x,u)+\psi(y,u)-\psi(x,y)], \end{aligned}$$

$$\tilde{K}(x,y) := \frac{1}{2|M|}\int_M [\psi(x,u)+\psi(y,u)-\psi(x,y)]d\mu(u) = \frac{1}{|M|}\int_M K_u(x,y)d\mu(u).$$

Denote by K and $\tilde{K}$ the operators with kernels $K(x,y)$ and $\tilde{K}(x,y)$. Then

$$\tilde{K}(x,y) = K(x,y)+|M|^{-1}\operatorname{Tr}(K)-|M|^{-1}K\mathbb{1}(x)-|M|^{-1}K\mathbb{1}(y). \tag{8.2}$$

Moreover, $\psi_{\tilde{K}} = \psi$, $\tilde{K}_u = K_u$, *and* $\tilde{K}\mathbb{1} = \tilde{\lambda}\mathbb{1}$, *where*

$$\begin{aligned}\tilde{\lambda} &= Tr(K) - \frac{1}{|M|}\int_M\int_M K(x,y)d\mu(x)d\mu(y)\\ &= \frac{1}{2|M|}\int_M\int_M \psi(u,y)d\mu(u)d\mu(y) \geq 0.\end{aligned}$$

In addition,

$$K = \tilde{K} + C \iff K\mathbb{1} = \lambda\mathbb{1} \tag{8.3}$$

and, if so, $\tilde{\lambda} = (\mathrm{Tr}(K) - \lambda)$, $C = \frac{1}{|M|}(Tr(K) - 2\lambda)$.

Remark 8.8. *The following useful assertions can be found in e.g.* [7], [38], [9], [37]. *For N.D. functions there exists a functional calculus that has no equivalent for P.D. functions:*

(1) *Let* F *a bounded completely continuous function, i.e.*

$$\forall z > 0,\ \forall n \in \mathbb{N},\ D^n F(z) \geq 0$$

or equivalently

$$F(z) = \int_0^\infty e^{-tz}d\mu(t), \quad \mu \geq 0, \quad \mu([0,\infty)) < \infty.$$

Then

$$\psi \text{ is N.D.} \Longrightarrow F(\psi) \text{ is P.D.}$$

(2) *If* G *is a Bernstein function, i.e.*

$$G(z) = az + \int_0^\infty (1 - e^{-tz})d\mu(t), \quad a \geq 0,\ ;\mu \geq 0, \quad \int_0^\infty \frac{t}{1+t}d\mu(t) < \infty,$$

then

$$\psi \text{ N.D.} \Longrightarrow G(\psi) \text{ is N.D.}$$

For instance we have:

$$\begin{aligned}&\psi \text{ is N.D.} \iff \forall t > 0,\ e^{-t\psi} \text{ is P.D.}\\ &\psi \text{ is N.D.} \Longrightarrow \forall\, 0 < \alpha \leq 1,\ \psi^\alpha \text{ is N.D.}\\ &\psi \text{ is N.D.} \Longrightarrow \log(1+\psi) \text{ is N.D.}\end{aligned}$$

8.2. **Appendix II: Gaussian probability on separable Banach spaces.**

For detailed account of the material in this section we refer the reader to [10].

Let E be a Banach space and let $\mathcal{B}(E)$ be the sigma-algebra of Borel sets on E. Let E^* be its topological dual, and assume $\mathcal{F}$ is a vector space of real-valued functions defined on E, and $\gamma(\mathcal{F}, E)$ is the sigma-algebra generated by $\mathcal{F}$.

If $\mathcal{F} = \mathcal{C}_b(E,\mathbb{R})$ is the vector space of continuous bounded functions on E, then $\gamma(\mathcal{C}_b(E,\mathbb{R}), E) = \mathcal{B}(E)$ is the Borel sigma-algebra.

If E is separable, it is well known that the sigma-algebra $\gamma(E^*, E)$ generated by E^* is $\mathcal{B}(E)$.

Proposition 8.9. *Let* E *be a separable Banach space. Let* H *be a subspace of* E^*, *endowed with the* $\sigma(E^*, E)$ *topology. Then*

$$H \text{ is closed} \iff H \text{ is stable by simple limit.}$$

Proof. The implication $\Rightarrow$ is obvious. We now prove $\Leftarrow$. By Banach-Krein-Smulian theorem, H is $\sigma(E^*,E)$-closed if and only if $\forall R>0$, $B(0,R)\cap H$ is $\sigma(E^*,E)$-closed. As E is a separable Banach space, we have: For all $R>0$

$$B(0,R)=\{f\in E^*:\|f\|_{E^*}\leq R\} \text{ is metrizable (and compact) for } \sigma(E^*,E).$$

Hence we only have to verify that for every sequence $(f_n)\subset B(0,R)\cap H$ such that $\lim_{n\mapsto\infty} f_n=f$ in the $\sigma(E^*,E)-$topology, we have $f\in B(0,R)\cap H$. But clearly this implies $\forall x\in E,\ \lim_{n\mapsto\infty} f_n(x)=f(x)$, so we have $f\in B(0,R)\cap H$. □

Corollary 8.10. *Let E is a separable Banach space and H is a subspace of E^*. Then*

(1) *$\overline{H}^{\sigma(E^*,E)}$ coincides with the smallest vector space of functions on E, stable by simple limits containing H.*

(2)
$$\gamma(H,E)=\gamma(\overline{H}^{\sigma(E^*,E)},E).$$

(3) *If H is a subspace of E^* separating E, then*
$$\gamma(H,E)=\gamma(E^*,E)=\mathcal{B}(E).$$

Proof. (1) Clearly, as E^* is stable by simple limits (by Banach-Steinhaus theorem), the smallest vector space of functions on E, stable by simple limits containing H is contained in E^* and by the previous proposition it is $\overline{H}^{\sigma(E^*,E)}$.

(2) Let $\gamma(H,E)$ is the sigma-algebra generated by H. The vector subspace $V=\{u\in E^*:u,\ \gamma(H,E)-\text{measurable}\}$ is stable by simple limits. Hence $\overline{H}^{\sigma(E^*,E)}\subset V$.

(3) By the Hahn-Banach theorem, if H is separating, $\overline{H}^{\sigma(E^*,E)}=E^*$ and hence

$$\gamma(H,E)=\gamma(E^*,E)=\mathcal{B}(E).\quad\square$$

Lemma 8.11. *Let E be a separable Banach space, and H be a subspace of E^* separating E. There is at most one probability measure P on the Borel sets of E such that, under P, $\gamma\in H$ is a centered Gaussian variable with a given covariance $K(\gamma,\gamma')$:*

$$K(\gamma,\gamma')=\int_E\gamma(\omega)\gamma'(\omega)DP(\omega)$$

on H. Moreover, if such a probability exists, then

(1) *E^* is a Gaussian space, and $\overline{E^*}^{L^2(E,P)}$ is the Gaussian space generated by H.*

(2) *There exists $\alpha>0$ such that*

$$\int_E e^{\alpha\|x\|_E^2}dP(x)<\infty. \tag{8.4}$$

Proof. If $K(\gamma,\gamma')$ is a positive definite function on H, it determines an additive function on the algebra of cylindrical sets related to H:

$$\big\{x\in E:(\gamma_1(x),\dots,\gamma_n(x))\in C\big\},\ \gamma_i\in H,\ C \text{ Borelian set of } \mathbb{R}^n.$$

Now, the sigma-algebra generated by this algebra is the Borelian of E.

Assume that such a probability P exists . Let $\mathcal{H}=E^*\cap\overline{H}^{L^2(E,P)}$. Clearly $\overline{H}^{L^2(E,P)}$ is the Gaussian space generated by H, and if $(\gamma_n)_{n\geq1}\in\mathcal{H}$ is such that $\forall x \in E$, $\lim_{n\mapsto\infty} \gamma_n(x) = \gamma(x)$ exists, then clearly $\gamma \in E^*$ by the Banach-Stheinhauss theorem, and $\gamma \in \overline{H}^{L^2(E,P)}$ since a simple limit of random variables in a closed Gaussian space belongs to this Gaussian space. Therefore, $\gamma \in \mathcal{H}$, which by Proposition 8.9 implies that $\mathcal{H}$ is closed. But $H \subset \mathcal{H}$ and $\overline{H}^{\sigma(E^*,E)} = E^*$ leads to $\mathcal{H} = E^*$.

Finally, (8.4) is just the Fernique theorem. $\square$

8.2.1. *Cameron-Martin space.* Let us recall that, due to Fernique theorem, and Bochner integration we have the following map from E^* to E :

$$I : \gamma \in E^* \mapsto \int_E^E \omega\gamma(\omega)dP(\omega) \in E$$

as

$$\Big\| \int_E^E \omega\gamma(\omega)dP(\omega)\Big\|_E \leq \int^E \|\omega\||\gamma(\omega)|dP(\omega) \leq \Big(\int^E \|\omega\|^2|dP(\omega)\Big)^{\frac{1}{2}}\|\gamma\|_{L^2(P,E)}$$

$$\text{and} \quad \forall\gamma,\gamma' \in E^*,\ \gamma'(I(\gamma)) = \int_E \gamma'(\omega)\gamma(\omega)dP(\omega).$$

Therefore, I can be extended to $\bar{I} : \overline{E^*}^{L^2(E,P)} \mapsto E$. The subspace

$$\mathbb{H} \subset E, \quad \mathbb{H} = \bar{I}(\overline{E^*}^{L^2(E,P)})$$

with the induced Hilbert structure is the Cameron-Martin space associated to the Gaussian probability space $(E, \mathcal{B}(E), P)$, (see [10]).

Important special case. Let M be a set and let E be a separable Banach space of real-valued functions on M. Let

$$\forall x \in M,\ f \in E \xrightarrow{\delta_x} f(x) \in \mathbb{R}.$$

Suppose $\delta_x \in E^*$. So, $\mathcal{H} = \{\sum_{\text{finite}} \alpha_i\delta_{x_i}\}$ is dense in E^* in the $\sigma(E^*,E)-$ topology.

Let $K(x,y)$ be a positive definite function on $M \times M$. There is at most one probability measure P on the Borelian sets of E such that, under P, E^* is a Gaussian space and $(\delta_x)_{x\in M}$ is a centered Gaussian process with covariance

$$K(x,y) = \int_E \delta_x(\omega)\delta_y(\omega)dP(\omega), \ \text{ i.e. } \forall t \in \mathbb{R},\ \int_E e^{-it\delta_x(\omega)}dP(\omega) = e^{-\frac{1}{2}t^2K(x,x)}.$$

The Cameron-Martin space is identified with the Reproducing Kernel Hilbert Space $\mathbb{H}_K$ associated to K, i.e. the closure of

$$\Big\{y \in M \mapsto f(y) = \sum_i \lambda_i K(x_i,y)\}; \quad \|f\|^2_{\mathbb{H}_K} = \sum_{i,j} \lambda_i\lambda_j K(x_i,x_j)\Big\}.$$

$\mathbb{H}_K$ is characterized as a Hilbert space of functions on M such that

$$\forall x \in M, f \in \mathbb{H}_K \mapsto f(x) = \langle K(x,.), f\rangle_{\mathbb{H}_K} \text{ (is continuous)}.$$

Therefore, if such a P exists on E, then $\mathbb{H}_K \subseteq M$.

References

[1] Adler, R.J.: An introduction to continuity, extrema, and related topics for general Gaussian processes. Institute of Mathematical Statistics Lecture NotesMonograph Series, 12. Institute of Mathematical Statistics, Hayward, CA (1990)

[2] Adler, R.J., Taylor J.E.: Random fields and geometry. Springer Monographs in Mathematics. Springer, New York (2007)

[3] Andreev, R., Lang, A.: Kolmogorov-Chentsov theorem and differentiability of random field on manifolds. Potential Anal. **41**(3), 761-769 (2014)

[4] Andrew, G., Askey, R., Roy, R.: Special functions. Encyclopedia of Mathematics and its Applications, 71. Cambridge University Press, Cambridge (1999)

[5] Aronszajn, N.: Theory of reproducing kernels. Trans. Amer. Math. Soc. **68**(3), 337-404 (1950)

[6] Askey, R., Bingham, N. H.: Gaussian processes on compact symmetric spaces. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **37**(2), 127-143 (1976/77)

[7] Berg, C., Christensen, J.P.R., Ressel, P.: Harmonic analysis on semigroups. Theory of positive definite and related functions. Graduate Texts in Mathematics, 100. Springer-Verlag, New York (1984)

[8] Bingham, N.H.: Positive definite functions on spheres. Proc. Cambridge Philos. Soc. **73**, 145-156 (1973)

[9] Bochner, S.: Harmonic analysis and the theory of probability. University of California Press, Berkeley and Los Angeles (1955)

[10] Bogachev, V.I.: Gaussian Measures. Mathematical Surveys and Monographs, 62. American Mathematical Society, Providence, RI (1998)

[11] Bouleau, N., Hirsch, F.: Dirichlet forms and analysis on Wiener space. de Gruyter Studies in Mathematics, 14. Walter de Gruyter & Co., Berlin (1991)

[12] Ciesielski, Z., Kerkyacharian, G., Roynette, B.: Quelques espaces fonctionnels associes a des processus Gaussiens. Studia Math. **107**(2), 171–204 (1993)

[13] Coifman, R., Weiss, G.: Analyse harmonique non-commutative sur certains espaces homogènes. Lecture Notes in Mathematics, Vol. 242. Springer-Verlag, Berlin-New York (1971)

[14] Coulhon, T., Kerkyacharian, G., Petrushev, P.: Heat kernel generated frames in the setting of Dirichlet spaces, J. Fourier Anal. Appl. **18**(5), 995–1066 (2012)

[15] Cucker, F., Smale, S.: On the mathematical foundations of learning. Bull. Amer. Math. Soc. (N.S.) **39**(1), 1-49 (2002)

[16] Davies, E. B.: Linear operators and their spectra. Cambridge Studies in Advanced Mathematics, 106. Cambridge University Press, Cambridge (2007)

[17] Faraut, J., Harzallah, K.: Distances hilbertiennes invariantes sur un espace homogène. Ann. Inst. Fourier (Grenoble) **24**(3), 171-217 (1974)

[18] Fernique, X.: Regularité des trajectoires des fonctions aléatoires gaussiennes. École d'Été de Probabilités de Saint-Flour, IV-1974, pp. 1-96. Lecture Notes in Math., Vol. 480, Springer, Berlin (1975)

[19] Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet forms and symmetric Markov processes. Second revised and extended edition. de Gruyter Studies in Mathematics, 19. Walter de Gruyter & Co., Berlin (2011)

[20] Gangolli, R.: Positive definite kernels on homogeneous spaces and certain stochastic processes related to Lévy's Brownian motion of several parameters. Ann. Inst. H. Poincaré Sect. B (N.S.) **3**(2), 121-226 (1967)

[21] Gine, E., Nickl, R. : Mathematical foundations of infinite-dimensional statistical models. 690p., Cambridge University Press (2016).

[22] Grigor'yan, A.: Heat kernel and analysis on manifolds. AMS/IP Studies in Advanced Mathematics, 47. American Mathematical Society, Providence, RI; International Press, Boston, MA (2009)

[23] Helgason, S.: Differential geometry and symmetric spaces. Pure and Applied Mathematics, Vol. XII. Academic Press, New York-London (1962)

[24] Steinwart, I., Scovel, C.: Mercer's Theorem on General Domains; On the Interaction between Measures, Kernels, and RKHSs. Constructive Approximation, 35,365-417 (2012)

[25] Kerkyacharian, G., Petrushev, P.: Heat kernel based decomposition of spaces of distributions in the framework of Dirichlet spaces. Trans. Amer. Math. Soc. **367**(1), 121-189 (2015)

[26] Knapik, B. T., van der Vaart, A. W., van Zanten, J. H.: Bayesian inverse problems with Gaussian priors. Ann. Statist. **39**(5), 2626-2657 (2011)
[27] Lang, A., Schwab, C.: Isotropic Gaussian random field on the sphere: regularity, fast simulation, and stochastic partial differential equation. arxiv:1305.1170v2. 16 May 2014.
[28] Ledoux, M.: Isoperimetry and Gaussian analysis. Lectures on probability theory and statistics (Saint-Flour, 1994), 165-294, Lecture Notes in Math., 1648, Springer, Berlin (1996)
[29] Ledoux, M.: The concentration of measure phenomenon. Mathematical Surveys and Monographs, 89. American Mathematical Society, Providence, RI (2001)
[30] Ledoux, M., Talagrand, M.: Probability in Banach spaces. Isoperimetry and processes. Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 23. Springer-Verlag, Berlin (1991)
[31] Li, W.V., Shao, Q.M.: Gaussian processes: inequalities, small ball probabilities and applications. Stochastic processes: theory and methods, 533–597, Handbook of Statistcs 19, North-Holland, Amsterdam (2001)
[32] Lifshits, M.: Lectures on Gaussian processes. Springer Briefs in Mathematics. Springer, Heidelberg (2012)
[33] Luschgy, H., Pages, G.: Expansions for Gaussian processes and Parseval frames. Electron. J. Probab. **14**(42), 1198-1221 (2009)
[34] Marcus, M.B., Rosen, J.: Markov processes, Gaussian processes, and local times. Cambridge Studies in Advanced Mathematics, 100. Cambridge University Press, Cambridge (2006)
[35] Peetre, J.: New thoughs on Besov spaces, Duke University (1976)
[36] Rasmussen, C. E., Williams, Ch.: Gaussian processes for machine learning. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA (2006)
[37] Schilling R.L., Song, R., Vondracek, Z.: Bernstein functions, theory and applications Studies in mathematics 37, De Gruyter, 2010.
[38] Schoenberg, I.J.: Metric spaces and positive definite functions. Trans. Amer. Math. Soc. **44**(3), 522–536 (1938)
[39] Seeger, M.: Gaussian processes for machine learning. International Journal of Neural Systems, **14**(02), 69–106 (2004)
[40] Steinwart, I., Christmann, A.: Support vector machines. Information Science and Statistics. Springer, New York (2008)
[41] Sturm, K.T.: Analysis on local Dirichlet spaces I. Recurrence, conservatisness and L_p-Liouville properties, J. Reine Angew. Math. **456**, 173–196 (1994)
[42] Sturm, K.T.: Analysis on local Dirichlet spaces II. Upper Gaussian estimates for the fundamental solutions of parabolic equations, Osaka J. Math. **32**(2), 275–312 (1995)
[43] Sturm, K.T.: Analysis on local Dirichlet spaces III. The parabolic Harnack inequality, J. Math. Pures Appl. **75**(3), 273–297 (1998)
[44] Sturm, K.T.: The geometric aspect of Dirichlet forms. New directions in Dirichlet forms, 233–277, AMS/IP Stud. Adv. Math., **8**, Amer. Math. Soc., Providence, RI(1998)
[45] Talagrand, M.: Regularity of Gaussian processes. Acta Math. **159**(1-2), 99–149 (1987)
[46] Talagrand, M.: Mean field models for spin glasses, Springer-Verlag, Berlin (2011)
[47] Talagrand, M.: Upper and lower bounds for stochastic processes. Springer, Heidelberg (2014)
[48] van der Vaart, A., van Zanten, H.: Bayesian inference with rescaled Gaussian process priors. Electron. J. Stat. **1**, 433–448 (2007)

LPMA, CNRS-UMR 7599, and Crest
E-mail address: `kerk@math.univ-paris-diderot.fr`

Ritsumeikan University, Kyoto
E-mail address: `ogawa-s@se.ritsumei.ac.jp`

Department of Mathematics, University of South Carolina, Columbia, SC 29208
E-mail address: `pencho@math.sc.edu`

Université Paris Diderot - Paris 7, LPMA
E-mail address: `picard@math.univ-paris-diderot.fr`