DECOMPOSITION OF GAUSSIAN PROCESSES, AND FACTORIZATION OF POSITIVE DEFINITE KERNELS

Palle Jorgensen and Feng Tian

Communicated by P.A. Cojuhari

Abstract. We establish a duality for two factorization questions, one for general positive definite (p.d.) kernels K, and the other for Gaussian processes, say V. The latter notion, for Gaussian processes is stated via Ito-integration. Our approach to factorization for p.d. kernels is intuitively motivated by matrix factorizations, but in infinite dimensions, subtle measure theoretic issues must be addressed. Consider a given p.d. kernel K, presented as a covariance kernel for a Gaussian process V. We then give an explicit duality for these two seemingly different notions of factorization, for p.d. kernel K, vs for Gaussian process V. Our result is in the form of an explicit correspondence. It states that the analytic data which determine the variety of factorizations for K is the exact same as that which yield factorizations for V. Examples and applications are included: point-processes, sampling schemes, constructive discretization, graph-Laplacians, and boundary-value problems.

Keywords: reproducing kernel Hilbert space, frames, generalized Ito-integration, the measurable category, analysis/synthesis, interpolation, Gaussian free fields, non-uniform sampling, optimization, transform, covariance, feature space.

Mathematics Subject Classification: 47L60, 46N30, 46N50, 42C15, 65R10, 05C50, 05C75, 31C20, 60J20.

1. INTRODUCTION

We give an integrated approach to positive definite (p.d.) kernels and Gaussian processes, with an emphasis on factorizations, and their applications. Positive definite kernels serve as powerful tools in such diverse areas as Fourier analysis, probability theory, stochastic processes, boundary theory, potential theory, approximation theory, interpolation, signal/image analysis, operator theory, spectral theory, mathematical physics, representation theory, complex function-theory, moment problems, integral equations, numerical analysis, boundary-value problems for partial differential

equations, machine learning, geometric embedding problems, and information theory. While there is no single book which covers all these applications, the reference [65] goes some of the way. As for the use of RKHS analysis in machine learning, we refer to [72] and [74].

Here, we give a new and explicit duality for positive definite functions (kernels) on the one hand, and Gaussian processes on the other. A covariance kernel for a general stochastic process is positive definite. In general, the stochastic process in question is not determined by its covariance kernel. But in the special case when the process is Gaussian, it is. In fact (Theorem 3.1), every p.d. kernel K is indeed the covariance kernel of a Gaussian process. The construction is natural; starting with the p.d. kernel K, there is a canonical inductive limit construction leading to the Gaussian process for this problem, following a realization of Gaussian processes dating back to Kolmogorov. The interplay between analytic properties of p.d. kernels and their associated Gaussian processes is the focus of our present study.

We formulate two different factorization questions, one for general p.d. kernels K, and the other for Gaussian processes, say V. The latter notion, for Gaussian processes, is a subordination approach. Our approach to factorization for p.d. kernels is directly motivated by matrix factorizations, but in infinite dimensions, there are subtle measure theoretic issues involved. If the given p.d. kernel K is already presented as a covariance kernel for a Gaussian process V, we then give an explicit duality for these two seemingly different notions of factorization. Our main result, Theorem 5.1, states that the analytic data which determine the variety of factorizations for K is the exact same as that which yield factorizations for V.

2. POSITIVE DEFINITE KERNELS
AND THEIR REPRODUCING KERNEL HILBERT SPACES

The notion of a positive definite (p.d.) kernel has come to serve as a versatile tool in a host of problems in pure and applied mathematics. The abstract notion of a p.d. kernel is in fact a generalization of that of a positive definite function, or a positive-definite matrix. Indeed, the matrix-point of view lends itself naturally to the particular factorization question which we shall address in Section 5 below. The general idea of p.d. kernels arose first in various special cases in the first half of 20th century: It occurs in work by J. Mercer in the context of solving integral operator equations; in the work of G. Szegő and S. Bergmann in the study of harmonic analysis and the theory of complex domains; and in the work by N. Aronszajn in boundary value problems for PDEs. It was Aronszajn who introduced the natural notion of reproducing kernel Hilbert space (RKHS) which will play a central role here; see especially (2.5) below. References covering the areas mentioned above include: $[9, 15, 29, 32, 34, 35, 51]$, and [46].

Right up to the present, p.d. kernels have arisen as powerful tools in many and diverse areas of mathematics. A partial list includes the areas listed above in the Introduction. An important new area of application of RKHS theory includes the following $[1–6, 11, 12]$.

Let X be a set and let K be a complex valued function on $X \times X$. We say that K is *positive definite* (p.d.) iff (Def.) for all finite subset F ($\subset X$) and complex numbers $(\xi_x)_{x \in F}$, we have

$$\sum_{x \in F} \sum_{y \in F} \overline{\xi}_x \xi_y K(x, y) \geq 0. \tag{2.1}$$

In other words, the $|F| \times |F|$ matrix $(K(x, y))_{F \times F}$ is positive definite in the usual sense of linear algebra. We refer to the rich literature regarding theory and applications of p.d. functions [7, 19, 24, 27, 44, 68, 71].

We shall also need the Aronszajn [15] reproducing kernel Hilbert spaces (R.K.H.S.), denoted $\mathscr{H}(K)$.

Let F, and $(\xi)_{x \in F}$, are as above, and consider all functions on X of the form

$$\sum_{x \in F} \xi_x K(\cdot, x). \tag{2.2}$$

If F (finite) is fixed, and $(\xi_x)_{x \in F}$, $(\eta_x)_{x \in F}$ are vectors in $\mathbb{C}^{|F|}$, we set

$$\left\langle \sum_{x \in F} \xi_x K(\cdot, x), \sum_{y \in F} \eta_y K(\cdot, y) \right\rangle_{\mathscr{H}(K)} := \sum \sum_{F \times F} \overline{\xi}_x \eta_y K(x, y), \tag{2.3}$$

and

$$\left\| \sum_{x \in F} \xi_x K(\cdot, x) \right\|_{\mathscr{H}(K)}^2 := \sum \sum_{F \times F} \overline{\xi}_x \xi_y K(x, y). \tag{2.4}$$

Then $\mathscr{H}(K)$ is the Hilbert completion of all functions as in (2.2) with respect to the $\|\cdot\|_{\mathscr{H}(K)}$-norm; see (2.3)–(2.4).

With the definition of the R.K.H.S. $\mathscr{H}(K)$, we get directly that the functions $\{K(\cdot, x)\}_{x \in X}$ are automatically in $\mathscr{H}(K)$; and that, for all $h \in \mathscr{H}(K)$, we have

$$\langle K(\cdot, x), h \rangle_{\mathscr{H}(K)} = h(x); \tag{2.5}$$

i.e., the *reproducing* property holds.

Further recall (see e.g. [65]) that, given K, then the R.K.H.S. $\mathscr{H}(K)$ is determined uniquely, up to isometric isomorphism in Hilbert space.

Remark 2.1. In general, when a positive definite kernel K is given, and the reproducing kernel Hilbert space (RKHS) $\mathscr{H}(K)$ is constructed, one only specifies a generating (dense span) system of elements in $\mathscr{H}(K)$. As outlined, this subspace (in $\mathscr{H}(K)$) may be specified as in (2.2), and the $\|\cdot\|_{\mathscr{H}(K)}$-norm is then specified as in (2.4). With this point of view, $\mathscr{H}(K)$ arises as a Hilbert-completion.

However, in many applications (see especially Sections 7–10 below), it is possible to give an explicit formula for *all* the elements ψ in $\mathscr{H}(K)$, and their Hilbert norms $\|\psi\|_{\mathscr{H}(K)}$.

Lemma 2.2. *Let $X \times X \xrightarrow{K} \mathbb{C}$ be a p.d. kernel, and let $\mathscr{H}(K)$ be the corresponding RKHS (see (2.3)–(2.5)). Let h be a function defined on X. Then the following conditions are equivalent:*

(i) $h \in \mathscr{H}(K)$;

(ii) *there is a constant $C = C_h < \infty$ such that, for all finite subset $F \subset X$, and all $(\xi_x)_{x \in F}$, $\xi_x \in \mathbb{C}$, the following a priori estimate holds:*

$$\left| \sum_{x \in F} \xi_x h(x) \right|^2 \leq C_h \sum_{x \in F} \sum_{y \in F} \overline{\xi}_x \xi_y K(x, y). \qquad (2.6)$$

Proof. The implication (i)$\Rightarrow$(ii) is immediate, and in this case, we may take $C_h = \|h\|^2_{\mathscr{H}(K)}$.

Now for the converse, assume (ii) holds for some finite constant. On the $\mathscr{H}(K)$-dense span in (2.2), define a linear functional

$$L_h \left(\sum_{x \in F} \xi_x K(\cdot, x) \right) := \sum_{x \in F} \xi_x h(x). \qquad (2.7)$$

From the assumption (2.6) in (ii), we conclude that L_h (in (2.7)) is a well defined bounded linear functional on $\mathscr{H}(K)$. Initially, L_h is only defined on the span (2.2), but by (2.6), it is bounded, and so extends uniquely by $\mathscr{H}(K)$-norm limits. We may therefore apply Riesz' lemma to the Hilbert space $\mathscr{H}(K)$, and conclude that there is a unique $H \in \mathscr{H}(K)$ such that

$$L_h(\psi) = \langle \psi, H \rangle_{\mathscr{H}(K)} \qquad (2.8)$$

for all $\psi \in \mathscr{H}(K)$. Now, setting $\psi(\cdot) := K(\cdot, x)$, for $x \in X$, we conclude from (2.8) that $h(x) = H(x)$; and so $h \in \mathscr{H}(K)$, proving (i). $\qquad \square$

3. GAUSSIAN PROCESSES

The interest in positive definite (p.d.) functions has at least three roots: (i) Fourier analysis, and harmonic analysis more generally; (ii) Optimization and approximation problems, involving for example spline approximations as envisioned by I. Schöenberg; and (iii) Stochastic processes. See [62, 69].

Below, we sketch a few details regarding (iii). A *stochastic process* is an indexed family of random variables based on a fixed probability space. In some cases, the processes will be indexed by some group G, or by a subset of G. For example, $G = \mathbb{R}$, or $G = \mathbb{Z}$, correspond to processes indexed by real time, respectively discrete time. A main tool in the analysis of stochastic processes is an associated *covariance function.*

A process $\{X_g \mid g \in G\}$ is called *Gaussian* if each random variable X_g is Gaussian, i.e., its distribution is Gaussian. For Gaussian processes, we only need two moments. So if we normalize, setting the mean equal to 0, then the process is determined by its covariance function. In general, the covariance function is a function on $G \times G$,

or on a subset, but if the process is *stationary*, the covariance function will in fact be a p.d. function defined on G, or a subset of G. For a systematic study of positive definite functions on groups G, on subsets of groups, and the variety of the extensions to p.d. functions on G, see e.g. [41].

By a theorem of Kolmogorov [53], every Hilbert space may be realized as a (Gaussian) reproducing kernel Hilbert space (RKHS), see Theorem 3.1 below, and also [32, 64, 73].

Now every positive definite kernel is also the covariance kernel of a Gaussian process; a fact which is a point of departure in our present analysis: Given a positive definite kernel, we shall explore its use in the analysis of the associated Gaussian process; and vice versa.

This point of view is especially fruitful when one is dealing with problems from stochastic analysis. Even restricting to stochastic analysis, we have the exciting area of applications to statistical learning theory [72, 74].

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a *probability space*, i.e., Ω is a fixed set (sample space), $\mathscr{F}$ is a specified sigma-algebra (events) of subsets in Ω, and $\mathbb{P}$ is a probability measure on $\mathscr{F}$.

A Gaussian random variable is a function $V : \Omega \to \mathbb{R}$ (in the real case), or $V : \Omega \to \mathbb{C}$, such that V is measurable with respect to the sigma-algebra $\mathscr{F}$ on Ω, and the corresponding sigma-algebra of Borel subsets in $\mathbb{R}$ (or in $\mathbb{C}$). Let $\mathbb{E}$ denote the expectation defined from $\mathbb{P}$, i.e.,

$$\mathbb{E}(\cdot) = \int_{\Omega} (\cdot) \, d\mathbb{P}. \tag{3.1}$$

The requirement on V is that its distribution is Gaussian. If g denotes a Gaussian on $\mathbb{R}$ (or on $\mathbb{C}$), the requirement is that

$$\mathbb{E}(f \circ V) = \int_{\mathbb{R}(\text{or } \mathbb{C})} f \, dg; \tag{3.2}$$

or equivalently

$$\mathbb{P}(V \in B) = \int_{B} dg = g(B) \tag{3.3}$$

for all Borel sets B; see Figure 1.

If $N \in \mathbb{N}$, and $V_1, \ldots, V_N$ are random variables, the Gaussian requirement is (see Figure 2) that the joint distribution of $(V_1, \ldots, V_N)$ is an N-dimensional Gaussian, say g_N, so if $B \subset \mathbb{R}^N$ then

$$\mathbb{P}((V_1, \ldots, V_N) \in B) = g_N(B). \tag{3.4}$$

For our present purpose we may restrict to the case where the mean (of the respective Gaussians) is assumed zero. In that case, a finite joint distribution is

determined by its covariance matrix. In the $\mathbb{R}^N$ case, it is specified as follows (the extension to $\mathbb{C}^N$ is immediate) $(G_N\,(j_1,j_2))_{j_1,j_2=1}^{N}$,

$$G_N\,(j_1,j_2) = \int\limits_{\mathbb{R}^N} x_{j_1} x_{j_2} g_N\,(x_1,\ldots,x_N)\,dx_1\ldots dx_N \qquad (3.5)$$

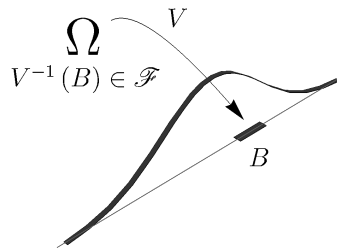where $dx_1\ldots dx_N = \lambda_N$ denotes the standard Lebesgue measure on $\mathbb{R}^N$.

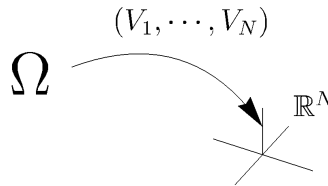**Fig. 1.** A Gaussian random variable and its distribution, see (3.3)

Fig. 2. A Gaussian system and its joint distribution, see (3.4)

The following is known:

Theorem 3.1 (Kolmogorov [54], see also [29,30])**.** *A kernel $K : X \times X \to \mathbb{C}$ is positive definite if and only if there is a (mean zero) Gaussian process $(V_x)_{x \in X}$ indexed by X such that*

$$\mathbb{E}\left(\overline{V}_x V_y\right) = K(x,y), \qquad (3.6)$$

where $\overline{V}_x$ denotes complex conjugation. Moreover (see Hida [28,30]), the process in (3.6) is uniquely determined by the kernel K in question. If $F \subset X$ is finite, then the covariance kernel for $(V_x)_{x \in F}$ is K_F given by

$$K_F(x,y) = G_F(x,y), \qquad (3.7)$$

for all $x,y \in F$, see (3.5) above.

In the subsequent sections, we shall address a number of properties of Gaussian processes important for their stochastic calculus. Our analysis deals with both the general case, and particular examples from applications. We begin in Section 4 with certain Wiener processes which are indexed by sigma-finite measures. For this class, the corresponding p.d. kernel has a special form; see (4.1) in Definition 4.1. (The case of fractal measures is part of Section 6 below.) In Section 5, we address the general case: We prove our duality result for factorization, Theorem 5.1. The remaining sections are devoted to examples and applications.

4. SIGMA-FINITE MEASURE SPACES AND GAUSSIAN PROCESSES

We shall consider functions of σ-finite measure space $(M, \mathscr{F}_M, \mu)$, where M is a set, $\mathscr{F}_M$ a σ-algebra of subsets in M, and μ is a positive measure defined on $\mathscr{F}_M$. It is further assumed that there is a countably indexed $(A_i)_{i \in \mathbb{N}}$ such that $0 < \mu(A_i) < \infty$, $M = \cup_i A_i$; and further that the measure space $(M, \mathscr{F}_M, \mu)$ is complete; so the Radon-Nikodym theorem holds. We shall also restrict to the case when μ is assumed non-atomic. The case when μ is atomic is different, and is addressed in Section 7 below.

Definition 4.1. Set

$$\mathscr{F}_{fin} = \{A \in \mathscr{F}_M \mid 0 < \mu(A) < \infty\}.$$

Note then

$$K^{(\mu)}(A, B) = \mu(A \cap B) \quad A, B \in \mathscr{F}_{fin} \tag{4.1}$$

is positive definite. The corresponding Gaussian process $(W_A^{(\mu)})_{A \in \mathscr{F}_{fin}}$ is called the Wiener process [28, 30]. In particular, we have

$$\mathbb{E}\left(W_A^{(\mu)} W_B^{(\mu)}\right) = \mu(A \cap B), \tag{4.2}$$

and

$$\lim_{(A_i)} \sum_i \left(W_{A_i}^{(\mu)}\right)^2 = \mu(A). \tag{4.3}$$

The precise limit in (4.3), quadratic variation, is as follows: Given μ as above, and $A \in \mathscr{F}_{fin}$, we then take limit over the filter of all partitions of A (see (4.4)) relative to the standard notation of refinement:

$$A = \cup_i A_i, \ A_i \cap A_j = \emptyset \text{ if } i \neq j, \text{ and } \lim \mu(A_i) = 0. \tag{4.4}$$

Details: Let $(\Omega, Cyl, \mathbb{P})$, $\mathbb{P} = \mathbb{P}^{(\mu)}$ be the probability space which realizes $W^{(\mu)}$ as a Gaussian process (or generalized Wiener process), i.e., such that (4.2) holds for all pairs in $\mathscr{F}_{fin}$. In particular, we have that $W_A^{(\mu)} \underset{(\text{dist})}{\sim} N(0, \mu(A))$, i.e., mean zero, Gaussian, and variance $= \mu(A)$. Then the following result holds.

Lemma 4.2 (see e.g., [9]). *With the assumptions as above, we have*

$$\lim_{(A_i)} \mathbb{E}\left(\left|\mu(A) \mathbb{1} - \sum_i (W_{A_i}^{(\mu)})^2\right|^2\right) = 0 \tag{4.5}$$

where (in (4.5)) the limit is taken over the filter of all partitions (A_i) of A, and $\mathbb{1}$ denotes the constant function "one" on Ω.

As a result, we get the following Ito-integral

$$W^{(\mu)}(f) := \int_M f(s) \, dW_s^{(\mu)}, \tag{4.6}$$

defined for all $f \in L^2(M, \mathscr{F}, \mu)$, and

$$\mathbb{E}\left(\left|\int_M f(s)\, dW_s^{(\mu)}\right|^2\right) = \int_M |f(s)|^2 \, d\mu(s). \tag{4.7}$$

We note that the following operator,

$$L^2(M, \mu) \ni f \longmapsto W^{(\mu)}(f) \in L^2(\Omega, \mathbb{P}) \tag{4.8}$$

is isometric.

In our subsequent considerations, we shall need the following precise formula (see Lemma 4.3) for the RKHS associated with the p.d. kernel

$$K^{(\mu)}(A, B) := \mu(A \cap B), \tag{4.9}$$

defined on $\mathscr{F}_{fin} \times \mathscr{F}_{fin}$. We denote the RKHS by $\mathscr{H}(K^{(\mu)})$.

Lemma 4.3. *Let μ be as above, and let $K^{(\mu)}$ be the p.d. kernel on $\mathscr{F}_{fin}$ defined in* (4.9). *Then the corresponding RKHS $\mathscr{H}(K^{(\mu)})$ is as follows: A function Φ on $\mathscr{F}_{fin}$ is in $\mathscr{H}(K^{(\mu)})$ if and only if there is a $\varphi \in L^2(M, \mathscr{F}_M, \mu) \left(=: L^2(\mu)\right)$ such that*

$$\Phi(A) = \int_A \varphi \, d\mu, \tag{4.10}$$

for all $A \in \mathscr{F}_{fin}$. Then

$$\|\Phi\|_{\mathscr{H}(K^{(\mu)})} = \|\varphi\|_{L^2(\mu)}. \tag{4.11}$$

Proof. To show that Φ in (4.10) is in $\mathscr{H}(K^{(\mu)})$, we must choose a finite constant C_Φ such that, for all finite subset $(A_i)_{i=1}^N$, $A_i \in \mathscr{F}_{fin}$, $\{\xi_i\}_{i=1}^N$, $\xi_i \in \mathbb{R}$, we get the following *a priori* estimate:

$$\left|\sum_{i=1}^N \xi_i \Phi(A_i)\right|^2 \leq C_\Phi \sum_i \sum_j \xi_i \xi_j K^{(\mu)}(A_i, A_j). \tag{4.12}$$

But a direct application of Schwarz to $L^2(\mu)$ shows that (4.12) holds, and for a finite C_Φ, we may take $C_\Phi = \|\varphi\|_{L^2(\mu)}^2$, where φ is the $L^2(\mu)$-function in (4.10). The desired conclusion now follows from an application of Lemma 2.2.

We have proved one implication from the statement of the lemma: Functions Φ on $\mathscr{F}_{fin}$ of the formula (4.10) are in the RKHS $\mathscr{H}(K^{(\mu)})$, and the norm $\|\cdot\|_{\mathscr{H}(K^{(\mu)})}$ is as stated in (4.11). In the below, we shall denote these elements in $\mathscr{H}(K^{(\mu)})$ as pairs (Φ, φ). We shall also restrict attention to the case of real valued functions.

For the converse implication, let H be a function on $\mathscr{F}_{fin}$, and assume $H \in \mathscr{H}(K^{(\mu)})$. Then by the Schwarz inequality applied to $\langle \cdot, \cdot \rangle_{\mathscr{H}(K^{(\mu)})}$ we get

$$\left|\langle H, \Phi \rangle_{\mathscr{H}(K^{(\mu)})}\right| \leq \|H\|_{\mathscr{H}(K^{(\mu)})} \|\varphi\|_{L^2(\mu)}, \tag{4.13}$$

where we used (4.11). Hence when Schwarz is applied to $L^2(\mu)$, we get a unique $h \in L^2(\mu)$ such that

$$\langle H, \Phi \rangle_{\mathscr{H}(K^{(\mu)})} = \int_M h\, \varphi\, d\mu \tag{4.14}$$

for all (Φ, φ) as in (4.10). Now specialize to $\varphi = \chi_A$, $A \in \mathscr{F}_{fin}$, in (4.14) and we conclude that

$$H(A) = \int_A h\, d\mu; \tag{4.15}$$

which translates into the assertion that the pair (H, h) has the desired form (4.10). And hence by (4.11) we have $\|H\|_{\mathscr{H}(K^{(\mu)})} = \|h\|_{L^2(\mu)}$ as stated. This concludes the proof of the converse inclusion. □

5. FACTORIZATIONS AND STOCHASTIC INTEGRALS

In Sections 2 and 3, we introduced the related notions of positive definite (p.d.) functions (kernels) on the one hand, and Gaussian processes on the other. One notes the immediate fact that a covariance kernel for a general stochastic process is positive definite. In general, the stochastic process in question is not determined by its covariance kernel. But in the special case when the process is Gaussian, it is.

In Theorem 3.1, we stated that every p.d. kernel K is indeed the covariance kernel of a Gaussian process. The construction is natural; starting with the p.d. kernel K, there is a canonical inductive limit construction leading to the Gaussian process for this problem. The basic idea for this particular construction of Gaussian processes dates back to pioneering work by Kolmogorov [29, 53].

In the present section, we formulate two different factorization questions, one for general p.d. kernels K, and the other for Gaussian processes, say V. For details, see the respective definitions in (5.2) and (5.3) below. If K is indeed the covariance kernel for a Gaussian process V, it is natural to try to relate these two seemingly different notions of factorization. (In the case of Gaussian processes, a better name is perhaps "subordination" (see (5.10) below), but our theorem justifies the use of factorization in both of these contexts.) Our main result, Theorem 5.1, states that the data determining factorization for K is the exact same as that which yields factorization for V.

Let K be a positive definite kernel $X \times X \xrightarrow{K} \mathbb{C}$; and let $V = V_K$ be the corresponding Gaussian (mean zero) process, indexed by X, i.e., $V_x \in L^2(\Omega, \mathbb{P})$, for all $x \in X$, and

$$\mathbb{E}\left(\overline{V}_x V_y\right) = K(x, y), \quad \forall (x, y) \in X \times X. \tag{5.1}$$

We set

$$
\mathscr{F}(K) := \Big\{ (M, \mathscr{F}_M, \mu) \mid \text{s.t. } K(\cdot, x) \longmapsto k_x \in L^2(M, \mu)
$$

extends to an isometry, i.e.,

$$
K(x, y) = \int_M \overline{k_x(s)} k_y(s)\, d\mu(s) = \langle k_x, k_y \rangle_{L^2(\mu)} \text{ for all } x, y \in X \Big\}. \tag{5.2}
$$

Further, if V is the Gaussian process (from (5.1)), we set

$$
\mathscr{M}(V) := \Big\{ (M, \mathscr{F}_M, \mu) \mid V \text{ admits an Ito-integral representation}
$$

$$
V_x = \int_M k_x(s)\, dW_s^{(\mu)},\ \forall x \in X, \text{ where } \{k_x\}_{x \in X} \text{ is}
$$

$$
\text{an indexed system in } L^2(M, \mu) \Big\}. \tag{5.3}
$$

Following parallel terminology from measure theory, we say that a Gaussian process V admits a disintegration, via suitable Ito-integrals, when there is a measure space with measure μ such that the corresponding Wiener process $W^{(\mu)}$ satisfies (5.3). Our theorem below (Theorem 5.1) shows that this disintegration question may be decided instead by the answer to an equivalent spectral decomposition question; the latter of course formulated for the covariance kernel for V. As is shown in the examples/applications below, given a Gaussian process, it is not at all clear what disintegrations hold; see for example Corollary 6.7.

Theorem 5.1. *Let $K : X \times X \to \mathbb{C}$ be given positive definite, and let $\{V_x\}_{x \in X}$ be the corresponding Gaussian (mean zero) process, then*

$$
\mathscr{F}(K) = \mathscr{M}(V). \tag{5.4}
$$

We shall need the following lemma.

Lemma 5.2. *From the definition of $\mathscr{F}(K)$, with K fixed and assumed p.d., we get to every $\big((k_x)_{x \in X}, \mu\big) \in \mathscr{F}(K)$ a natural isometry $T_\mu : \mathscr{H}(K) \longrightarrow L^2(M, \mu)$. It is denoted by*

$$
T_\mu(\underbrace{K(\cdot, x)}_{\in \mathscr{H}(K)}) := k_x \in L^2(\mu); \tag{5.5}
$$

and the adjoint operator $T_\mu^* : L^2(M, \mu) \longrightarrow \mathscr{H}(K)$ is as follows: For all $f \in L^2(M, \mu)$ we have

$$
(T_\mu^* f)(x) = \int_M f(s)\, \overline{k_x(s)}\, d\mu(s). \tag{5.6}
$$

Moreover, we also have

$$
T_\mu^*(k_x) = K(\cdot, x) \text{ for all } x \in X. \tag{5.7}
$$

Proof. Since $(k_x, \mu) \in \mathscr{F}(K)$, we have the factorization property (5.2), and so it follows from (5.5) that this extends by linearity and norm-completion to an isometry $\mathscr{H}(K) \xrightarrow{T_\mu} L^2(\mu)$ as stated.

By the definition of the adjoint operator $L^2(\mu) \xrightarrow{T_\mu^*} \mathscr{H}(K)$, we have for $f \in L^2(\mu)$:

$$\left(T_\mu^* f\right)(x) = \left\langle K(\cdot, x), T_\mu^* f \right\rangle_{\mathscr{H}(K)} = \langle k_x, f \rangle_{L^2(\mu)} = \int_M f(s) \overline{k_x(s)} d\mu(s),$$

which is the assertion in the lemma.

From the properties of $\mathscr{H}(K)$ (see Section 2), it follows that (5.7) holds iff

$$\left\langle K(\cdot, y), T_\mu^*(k_x) \right\rangle_{\mathscr{H}(K)} = \left\langle K(\cdot, y), K(\cdot, x) \right\rangle_{\mathscr{H}(K)} \tag{5.8}$$

for all $y \in X$. But we may compute both sides in eq. (5.8) as follows:

$$\begin{aligned}
\text{LHS}_{(5.8)} \quad &= \quad \langle T_\mu K(\cdot, y), k_x \rangle_{L^2(\mu)} \\
&\underset{\text{by (5.5)}}{=} \langle k_y, k_x \rangle_{L^2(\mu)} \\
&\underset{\text{by (5.2)}}{=} K(y, x) \\
&\underset{\text{by (2.3)}}{=} \text{RHS}_{(5.8)}. \qquad \square
\end{aligned}$$

Proof of Theorem 5.1. The proof is divided into two parts, one for each of the inclusions $\subseteq$ and $\supseteq$ in (5.4).

Part 1 "$\subseteq$". Assume a pair $\left((k_x)_{x \in X}, \mu\right)$ is in $\mathscr{F}(K)$; see (5.2). Then by definition, the factorization (5.3) holds on $X \times X$. Now let $W^{(\mu)}$ denote the Wiener process associated with μ, i.e., $W^{(\mu)}$ is a Gaussian process indexed by $\mathscr{F}_{fin}$, and

$$\mathbb{E}\left(W_A^{(\mu)} W_B^{(\mu)}\right) = \mu(A \cap B), \tag{5.9}$$

for all $A, B \in \mathscr{F}_{fin}$; see (4.1) above. Now form the Ito-integral

$$V_x := \int_M k_x(s) \, dW_s^{(\mu)}, \quad x \in X. \tag{5.10}$$

We stress that then V_x, as defined by (5.10), is a Gaussian process indexed by X. To see this, use the general theory of Ito-integration, see also [28, 29, 45–48, 51]. The approximation in (5.10) is over the filter of all *partitions*

$$\{A_i\}_{i \in \mathbb{N}} \text{ s.t. } A_i \cap A_j = \emptyset, \, i \neq j, \, \cup_{i \in \mathbb{N}} A_i = M, \text{ and } 0 < \mu(A_i) < \infty; \tag{5.11}$$

see (4.4). From the property of $W_{A_i}^{(\mu)}$, $i \in \mathbb{N}$, we conclude that, for all $s_i \in A_i$, we have that

$$\sum_{i \in \mathbb{N}} k_x(s_i) W_{A_i}^{(\mu)} \tag{5.12}$$

is Gaussian (mean zero) with

$$
\mathbb{E}\left|\sum_i k_x\left(s_i\right) W_{A_i}^{(\mu)}\right|^2 = \sum_i \sum_j \overline{k_x\left(s_i\right)} k_x\left(s_j\right) \mu\left(A_i \cap A_j\right)
$$
$$
= \sum_i \left|k_x\left(s_i\right)\right|^2 \mu\left(A_i\right), \qquad (5.13)
$$

where we used (5.11). Passing to the limit over the filter of all partitions of M (as in (5.11)), we then get

$$
\mathbb{E}\left(\int_M \overline{k_x\left(s\right)} dW_s^{(\mu)} \int_M k_y\left(t\right) dW_t^{(\mu)}\right) = \int_M \overline{k_x\left(s\right)} k_y\left(s\right) d\mu\left(s\right);
$$

and with definition (5.10), therefore:

$$
\mathbb{E}\left(\overline{V}_x V_y\right) = \langle k_x, k_y \rangle_{L^2(\mu)} = K(x,y), \quad \forall (x,y) \in X \times X, \qquad (5.14)
$$

where the last step in the derivation (5.14) uses the assumption that $\left((k_x)_{x \in X}, \mu\right) \in \mathscr{F}(K)$; see (5.2).

Part 2 "$\supseteq$". Assume now that some pair $\left((k_x)_{x \in X}, \mu\right)$ is in $\mathscr{M}(V)$ where K is given assumed p.d.; and where $(V_x)_{x \in X}$ is "the" associated (mean zero) Gaussian process; i.e., with K as its covariance kernel; see (5.1).

We claim that $\left((k_x)_{x \in X}, \mu\right)$ must then be in $\mathscr{F}(K)$, i.e., that the factorization (5.3) holds. This in turn follows from the following chain of identities:

$$
\langle k_x, k_y \rangle_{L^2(\mu)} = \mathbb{E}\left(\overline{V}_x V_y\right) \quad \left(\text{since } V_x = \int_M k_x\left(s\right) dW_s^{(\mu)}\right)
$$
$$
= K(x,y) \qquad \left(\text{since } K \text{ is the covariance kernel} \qquad (5.15)\right.
$$
$$
\left. \text{of the Gaussian process } (V_x)_{x \in X}\right)
$$

valid for all $(x,y) \in X \times X$, and the conclusion follows. Note that the first step in the derivation of (5.15) uses the Ito-isometry. Hence, initially K may possibly be the covariance kernel for a mean zero Gaussian process, say (V_x'), different from $V_x := \int_M k_x\left(s\right) dW_s^{(\mu)}$. But we proved that the two Gaussian processes V_x, and V_x', have the same covariance kernel. It follows then the two processes must be equivalent. This is by general theory; see e.g. [9, 33, 34].

The last uniqueness is only valid since we can consider Gaussian processes. Other stochastic processes are typically not determined uniquely from the respective covariance kernels. $\qquad \square$

Remark 5.3. In the statement of Theorem 5.1 there are two isometries: Starting with $\left((k_x)_{x \in X}, \mu\right) \in \mathscr{F}(K)$ we get the canonical isometry $T_\mu : \mathscr{H}(K) \to L^2(\mu)$ given by

$$
T_\mu\left(K\left(\cdot, x\right)\right) = k_x; \qquad (5.16)
$$

see (5.5) of Lemma 5.2. But with μ, we then also get the Wiener process $W^{(\mu)}$ and the Ito-integral

$$L^2\left(M,\mu\right) \ni f \longmapsto \int_M f\, dW^{(\mu)} \in L^2\left(\Omega, Cyl, \mathbb{P}\right) \tag{5.17}$$

as an isometry. Here $(\Omega, Cyl, \mathbb{P})$ denotes the standard probability space, with Cyl abbreviation for the cylinder sigma-algebra of subsets of $\Omega := \mathbb{R}^M$. For finite subsets $(s_1, s_2, \ldots, s_k)$ in M, and Borel subsets B_k in $\mathbb{R}^k$, the corresponding cylinder set

$$Cyl\left(\left(s_i\right)_{i=1}^k\right) := \left\{\omega \in \mathbb{R}^M ; \left(\omega\left(s_1\right), \ldots, \omega\left(s_k\right)\right) \in B_k\right\}.$$

In summary, we get the following diagram of isometries (see Figure 3), corresponding to a fixed $\left(\left(k_x\right)_{x \in X}, \mu\right) \in \mathscr{F}\left(K\right)$, where K is a fixed p.d. function on $X \times X$.
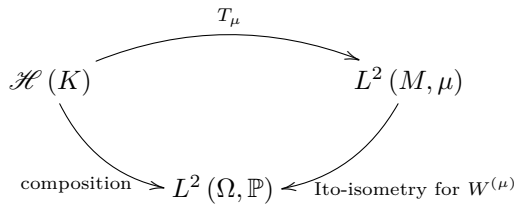
Fig. 3. The two isometries. Factorizations by isometries

6. EXAMPLES AND APPLICATIONS

Below we present four examples in order to illustrate the technical points in Theorem 5.1. In the first example $X = [0, 1]$, the unit interval, and in the next two examples $X = \mathbb{D} = \{z \in \mathbb{C} ; |z| < 1\}$ the open complex disk. In the fourth example, the Drury-Arveson kernel, we have $X = \mathbb{C}^k$.

We begin with a note on identifications. For $t \in [0, 1]$, we set

$$e\left(t\right) := e^{i2\pi t}.$$

We write λ_1 for the Lebesgue measure restricted to $[0, 1]$; and we make the identification:

$$[0, 1] \cong \mathbb{R}/\mathbb{Z} \cong \mathbb{T}^1 = \left\{z \in \mathbb{C} ; |z| = 1\right\}. \tag{6.1}$$

Hence, for $L^2\left([0,1],\lambda_1\right)$ we have the following familiar Fourier expansion. With

$$f \in L^2\left(\lambda_1\right) \;\; \text{and} \;\; c_n := \int_{[0,1]} \overline{e\left(nt\right)}f\left(t\right)d\lambda_1\left(t\right), \; n \in \mathbb{Z}, \tag{6.2}$$

we get

$$f\left(t\right) = \sum_{n \in \mathbb{Z}} c_n e\left(nt\right) \;\; \text{and} \;\; \int_0^1 \left|f\left(t\right)\right|^2 d\lambda_1\left(t\right) = \sum_{n \in \mathbb{Z}} \left|c_n\right|^2. \tag{6.3}$$

On $[0,1]$, we shall also consider the Cantor measure μ_4 with support equal to the Cantor set

$$C_4 = \left\{x \; ; \; x = \sum_{k=1}^{\infty} b^k/4^k, \; b_k \in \{0,2\}\right\} \subseteq [0,1];$$

see Figure 4 and [35, 39].
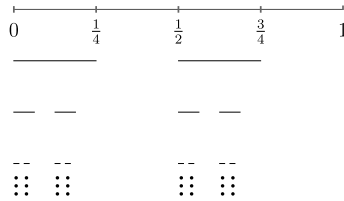
Fig. 4. The 4-Cantor set with double gaps as an iterated function system. This is an iterated-function system construction: Cantor-set and measure; see (6.4) below

It is known that μ_4 is the unique probability measure such that

$$\frac{1}{2}\int_0^1 \left(f\left(\frac{x}{4}\right) + f\left(\frac{x+2}{4}\right)\right) d\mu_4\left(x\right) = \int_0^1 f\,d\mu_4. \tag{6.4}$$

For the Fourier transform $\widehat{\mu}_4$ we have

$$\widehat{\mu}_4\left(t\right) = \prod_{k=1}^{\infty} \frac{1}{2}\left(1 + e^{i\pi t/4^k}\right), \quad t \in \mathbb{R}. \tag{6.5}$$

In Table 1, we summarize the three examples with the data from Theorem 5.1.

**Table 1.**   Three p.d. kernels and their respective Gaussian realizations

	X	K	$k_x, M = [0,1] \cong \mathbb{T}^1,$ $\mathscr{F}(K) = \{(k_x, \mu)\}$	μ
Ex. 6.1	$[0,1]$	$x \wedge y$	$k_x(s) = \chi_{[0,x]}(s)$	λ_1
Ex. 6.2	$\mathbb{D}$	$\dfrac{1}{1 - z\overline{w}}$	$k_z(t) = \dfrac{1}{1 - ze(t)}$	λ_1 on $\mathbb{T}^1$
Ex. 6.3	$\mathbb{D}$	$\prod_{n=0}^{\infty} \left(1 + z^{4^n}\overline{w}^{4^n}\right)$	$k_z(t) = \prod_{n=0}^{\infty}\left(1 + z^{4^n} e\left(4^n t\right)\right)$	μ_4

We now turn to the details of the respective examples.

Example 6.1. If $K(x,y) := x \wedge y$ is considered a kernel on $[0,1] \times [0,1]$, then the corresponding RKHS $\mathscr{H}(K)$ is the Hilbert space of functions f on $[0,1]$ such that the distribution derivative $f' = df/dx$ is in $L^2([0,1], \lambda_1)$, $\lambda_1 = dx$, $f(0) = 0$, and
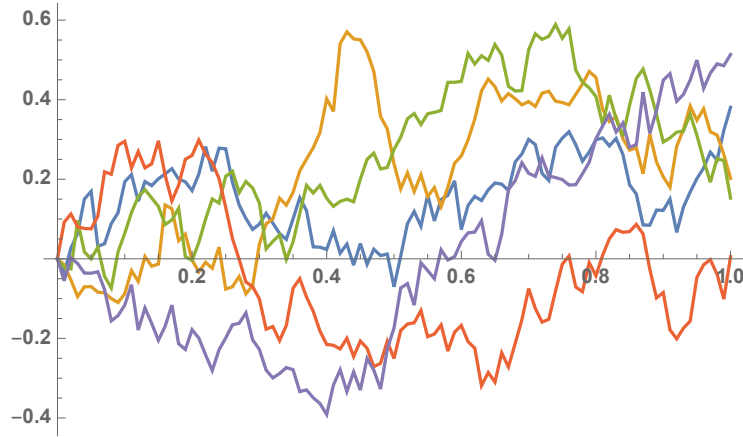
$$\|f\|^2_{\mathscr{H}(K)} := \int_0^1 |f'(x)|^2\,dx; \tag{6.6}$$

and it is immediate that $(k_x, \lambda_1) \in \mathscr{F}(K)$, where $k_x(s) := \chi_{[0,x]}(s)$, the indicator function; see Figure 5.
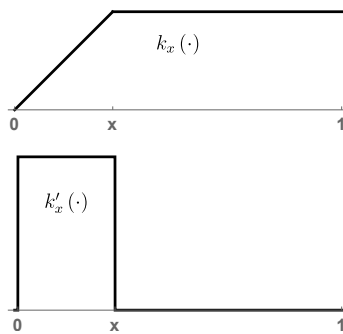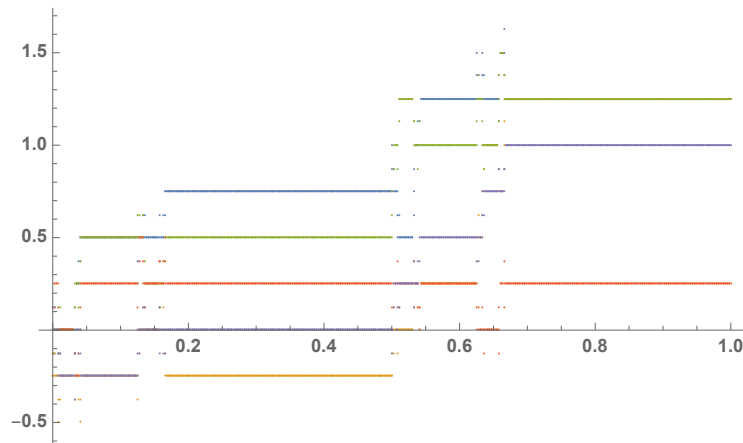

Fig. 5.   The generators of the Cameron-Martin RKHS

The process $W^{(\lambda_1)}$ is of course the standard Brownian motion on $[0,1]$, pinned at $x = 0$; see Figure 6, and compare with the $W^{(\mu_4)}$-process in Figure 7. For Monte Carlo simulation, see e.g. [56, 58].

Fig. 6. Brownian motion on $[0, 1]$. Sample-paths by Monte Carlo. See Example 6.1

Fig. 7. A Wiener process with holding patterns in the gaps of the Cantor set C_4 in Figure 4: The $W^{(\mu_4)}$ process on $[0, 1]$. Sample-paths by Monte Carlo

The Hilbert space characterized by (6.6) is called the Cameron-Martin space, see e.g., [29]. Moreover, to see that (6.6) is indeed the precise characterization of the RKHS for this kernel, one again applies Lemma 2.2.

It immediately follows from Theorem 5.1 that the Gaussian processes corresponding to the data in Table 1 are as follows:

Example 6.2. $z \in \mathbb{D}$:

$$V_z = \int\limits_0^1 \frac{1}{1 - z\overline{e(t)}} dW_t^{(\lambda_1)} \tag{6.7}$$

realized as an Ito-integral.

As an application of Theorem 5.1, we get

$$\mathbb{E}\left(\overline{V}_z V_w\right) = \frac{1}{1 - z\overline{w}}, \quad \forall\, (z, w) \in \mathbb{D} \times \mathbb{D}.$$

Example 6.3. $z \in \mathbb{D}$:

$$V_z = \int_0^1 \prod_{n=0}^{\infty} \left(1 + z^{4^n}\overline{e\left(4^n t\right)}\right) dW_t^{(\mu_4)}, \tag{6.8}$$

where the $W^{(\mu_4)}$-Ito integral is supported on the Cantor set $C_4 \subset [0, 1]$, see Figure 4.

As an application of Theorem 5.1, we get:

$$\mathbb{E}\left(\overline{V}_z V_w\right) = \prod_{n=0}^{\infty} \left(1 + (z\overline{w})^{4^n}\right).$$

The reasoning of Example 6.3 is based on a theorem of the paper [39] (see also [35]). Set

$$\Lambda_4 = \{0, 1, 4, 5, 16, 17, 20, 21, 64, 65, \ldots\}$$
$$= \left\{\sum_{k=0}^{\text{finite}} \alpha_k 4^k \mid \alpha_k \in \{0, 1\}, \text{ fintie summations}\right\}. \tag{6.9}$$

Then the Fourier functions $\{e\left(\lambda t\right); \lambda \in \Lambda_4\}$ forms an orthonormal basis in $L^2\left(C_4, \mu_4\right)$, i.e., every $f \in L^2\left(C_4, \mu_4\right)$ has its Fourier expansion

$$\widehat{f}\left(\lambda\right) = \int_{C_4} \overline{e\left(\lambda t\right)} f\left(t\right) d\mu_4\left(t\right);$$
$$f\left(t\right) = \sum_{\lambda \in \Lambda_4} \widehat{f}\left(\lambda\right) e\left(\lambda t\right);$$

and

$$\int_{C_4} |f|^2 \, d\mu_4 = \sum_{\lambda \in \Lambda_4} \left|\widehat{f}\left(\lambda\right)\right|^2.$$

Lemma 6.4. *Consider the set Λ_4 in (6.9), and, for $s \in \mathbb{D}$, let*

$$F\left(s\right) := \sum_{\lambda \in \Lambda_4} s^{\lambda} \tag{6.10}$$

be the corresponding generating function. Then we have the following infinite-product representation

$$F\left(s\right) = \prod_{n=0}^{\infty} \left(1 + s^{4^n}\right). \tag{6.11}$$

Proof. From (6.9) we have the following self-similarity for Λ_4: It is the following identity of sets

$$\Lambda_4 = \{0, 1\} + 4\Lambda_4. \tag{6.12}$$

Note that (6.12) is an algorithm for generating points in Λ_4. Hence,

$$\begin{aligned}
F(s) &= \sum_{\lambda \in \Lambda_4} s^\lambda = \sum_{\{0,1\}+4\Lambda_4} s^\lambda \\
&= \sum_{4\Lambda_4} s^\lambda + s \sum_{4\Lambda_4} s^\lambda \\
&= (1+s) F(s^4) \\
&= (1+s)(1+s^4)\ldots(1+s^{4^{n-1}})F(s^{4^n})
\end{aligned}$$

and by induction.

Hence, if $s \in \mathbb{D}$, the infinite-product is absolutely convergent, and the desired product formula (6.11) follows. $\qquad\square$

Remark 6.5. Note that, in combination with the theorem from [39] (see also [35]), this property of the generating function $F = F_{\Lambda_4}$ from Lemma 6.4 is used in the derivation of the assertions made about the factorization properties in Example 6.3; this includes the two formulas (Ex 3) as stated in Table 1; as well as of the verification that $(k_z, \mu_4) \in \mathscr{F}(K)$, where k_z, μ_4, and K are as stated.
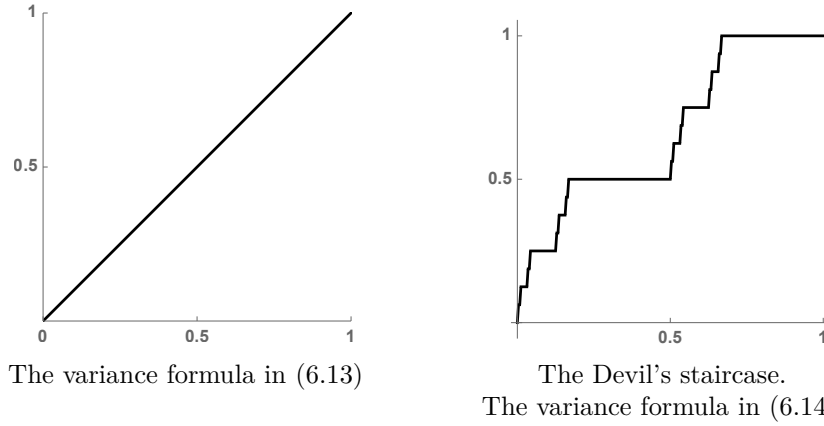
A direct computation of the two cases, Example 6.1 and Example 6.3, is of interest. Our result, Lemma 4.3, is useful in the construction: When computing the two Wiener processes $W^{(\lambda_1)}$ and $W^{(\mu)}$ one notes that the covariance computed on intervals $[0, x]$ as $0 < x < 1$ are as follows:

$$\mathbb{E}\left(\left(W_{[0,x]}^{(\lambda_1)}\right)^2\right) = \lambda_1([0,x]) = x, \tag{6.13}$$

and

$$\mathbb{E}\left(\left(W_{[0,x]}^{(\mu_4)}\right)^2\right) = \mu_4([0,x]). \tag{6.14}$$

So the two functions have the representations as in Figure 8.

The variance formula in (6.13)

The Devil's staircase.
The variance formula in (6.14)

Fig. 8. The two cumulative distributions

Example 6.6. The following example illustrates the need for a distinction between X, and families of choices M in Theorem 5.1. *A priori*, one might expect that if $X \times X \xrightarrow{K} \mathbb{C}$ is given and p.d., it would be natural to try to equip X with a σ-algebra $\mathscr{F}_X$ of subsets, and a measure μ such that the condition in (5.2) holds for $(X, \mathscr{F}_X, \mu)$, i.e.,

$$K(x, y) = \int_X \overline{k}_x k_y \, d\mu, \quad (x, y) \in X \times X \tag{6.15}$$

with $\{k_x\}_{x \in X}$ a system in $L^2(X, \mathscr{F}_X, \mu)$. It turns out that there are interesting examples where this is known to *not* be feasible. The best known such example is perhaps the Drury–Arveson kernel; see [16] and [13, 14].

Specifics. Consider $\mathbb{C}^k$ for $k \geq 2$, and $B_k \subset \mathbb{C}^k$ the complex ball defined for $z = (z_1, \dots, z_k) \in \mathbb{C}^k$,

$$B_k := \Big\{ z \in \mathbb{C}^k \; ; \; \underbrace{\sum\nolimits_{j=1}^{k} |z_j|^2}_{\|z\|_2^2} < 1 \Big\}. \tag{6.16}$$

For $z, w \in \mathbb{C}^k$, set

$$\langle z, w \rangle := \sum_{j=1}^{k} z_j \overline{w}_j$$

and

$$K_{DA}(z, w) := \frac{1}{1 - \langle z, w \rangle}, \ (z, w) \in B_k \times B_k.$$

Corollary 6.7 ([16, Corollary 2])**.** *Let $k \geq 2$, and let $\mathscr{H}(K_{DA})$ be the RKHS of the Drury–Arveson kernel in* (6.6)*. Then there is <u>no</u> Borel measure on $\mathbb{C}^k$ such that $\left(\mathbb{C}^k, \mathscr{B}_k, \mu\right) \in \mathscr{F}(K_{DA})$; i.e., there is <u>no</u> solution to the formula*

$$\|f\|^2_{\mathscr{H}(K_{DA})} = \int\limits_{\mathbb{C}^k} |f(z)|^2 \, d\mu(z),$$

for all $f(z)$ k-polynomials.

Remark 6.8. It is natural to ask about disintegration properties for the Gaussian process V_{DA} corresponding to the Drury–Arveson kernel (6.6). Combining our Theorem 5.1 above with the corollary (Corollary 6.7), we conclude that, in two or more complex dimensions k, the question of finding the admissible disintegrations this Gaussian process V_{DA} is subtle. It must necessarily involve measure spaces going beyond $\mathbb{C}^k$.

7. THE CASE OF $(k_x, \mu) \in \mathscr{F}(K)$ WHEN μ IS ATOMIC

Below we present a case where μ from pairs in $\mathscr{F}(K)$ may be chosen to be atomic. The construction is general, but for the sake of simplicity we shall assume that a given p.d. K is such that the RKHS $\mathscr{H}(K)$ is separable, i.e., when it has an (all) orthonormal basis (ONB) indexed by $\mathbb{N}$.

Definition 7.1. Let $\mathscr{H}$ be a Hilbert space (separable), and let $\{g_n\}_{n \in \mathbb{N}}$ be a system of vectors in $\mathscr{H}$ such that

$$\sum_{n \in \mathbb{N}} |\langle \psi, g_n \rangle_{\mathscr{H}}|^2 = \|\psi\|^2_{\mathscr{H}} \tag{7.1}$$

holds for all $\psi \in \mathscr{H}$. We then say that $\{g_n\}_{n \in \mathbb{N}}$ is a Parseval frame for $\mathscr{H}$. (Also see Definition 10.1.)

An equivalent assumption is that the mapping

$$\mathscr{H} \ni \psi \xmapsto{\quad T \quad} (\langle \psi, g_n \rangle_{\mathscr{H}}) \in l^2(\mathbb{N}) \tag{7.2}$$

is isometric. One checks that then the adjoint $T^* : l^2 \to \mathscr{H}$ is

$$T^*((\xi_n)) = \sum_{n \in \mathbb{N}} \xi_n g_n \in \mathscr{H}.$$

For general background references on frames in Hilbert space, we refer to [20, 22, 25, 26, 47, 52, 57, 66, 70], and also see [17, 50, 57, 63, 75].

Lemma 7.2. *Let K be given p.d. on $X \times X$, and assume that $\{g_n\}_{n \in \mathbb{N}}$ is a Parseval frame in $\mathscr{H}(K)$; then*

$$K(x, y) = \sum_{n \in \mathbb{N}} g_n(x) \overline{g_n(y)} \tag{7.3}$$

with the sum on the RHS in (7.3) *absolutely convergent.*

Proof. By the reproducing property of $\mathscr{H}(K)$, see Section 2, we get, for all $(x, y) \in X \times X$:

$$
\begin{aligned}
K(x, y) &= \langle K(\cdot, x), K(\cdot, y) \rangle_{\mathscr{H}(K)} \\
&\underset{\text{by (7.1)}}{=} \sum_{n \in \mathbb{N}} \langle K(\cdot, x), g_n \rangle_{\mathscr{H}(K)} \langle g_n, K(\cdot, y) \rangle_{\mathscr{H}(K)} \\
&\underset{\text{by (2.5)}}{=} \sum_{n \in \mathbb{N}} g_n(x) \overline{g_n(y)}.
\end{aligned}
$$

$\square$

Now a direct application of the argument in the proof of Theorem 5.1 yields the following:

Corollary 7.3. *Let K be given p.d. on $X \times X$ such that $\mathscr{H}(K)$ is separable, and let $\{g_n\}_{n \in \mathbb{N}}$ be a Parseval frame, for example an ONB in $\mathscr{H}(K)$. Let $\{\zeta_n\}_{n \in \mathbb{N}}$ be a chosen system of i.i.d. (independent identically distributed) system of standard Gaussians, i.e., with $N(0, 1)$-distribution $1/\sqrt{2\pi} e^{-s^2/2}$, $s \in \mathbb{R}$. Then the following sum defines a Gaussian process,*

$$
V_x(\cdot) := \sum_{n \in \mathbb{N}} g_n(x) \zeta_n(\cdot), \tag{7.4}
$$

i.e., $\{V_x\}_{x \in X}$ is well-defined in $L^2(\Omega, Cyl, \mathbb{P})$, as stated, where $\Omega = \mathbb{R}^{\mathbb{N}}$ as a realization in an infinite Cartesian product with the usual cylinder σ-algebra, and $\{V_x\}_{x \in X}$ has K as covariance kernel, i.e.,

$$
\mathbb{E}\left(\overline{V}_x V_y\right) = K(x, y), \quad \forall (x, y) \in X \times X;
$$

see (5.15).

Proof. This is a direct application of Lemma 7.2, and we leave the remaining verifications to the reader. $\square$

8. POINT PROCESSES: THE CASE WHEN $\{\delta_x\} \subset \mathscr{H}(K)$

Let $X \times X \xrightarrow{K} \mathbb{R}$ be a fixed positive definite kernel. We know that the RKHS $\mathscr{H}(K)$ consists of functions h on X subject to the *a priori* estimate in Lemma 2.2. For recent work on point-processes over infinite networks [18, 23, 37, 38, 43, 45, 49, 61, 67], the case when the Dirac measures δ_x are in $\mathscr{H}(K)$ is of special significance. In this case there is an abstract Laplace operator Δ, defined as follows:

$$
(\Delta h)(x) = \langle \delta_x, h \rangle_{\mathscr{H}(K)}, \ \forall h \in \mathscr{H}(K). \tag{8.1}
$$

For the $\|\cdot\|_{\mathscr{H}(K)}$-norm of δ_x, we have

$$
(\Delta \delta_x)(x) = \|\delta_x\|_{\mathscr{H}(K)}^2; \tag{8.2}
$$

immediate from (8.1).

For every finite subset $F \subset X$, we consider the induced $|F| \times |F|$ matrix

$$K_F(x,y) = (K(x,y))_{x,y \in F}. \tag{8.3}$$

Note that K_F is a positive definite square matrix. Its spectrum consists of eigenvalues $\lambda_s(F)$.

If (K, X) is as described, i.e., $X \times X \xrightarrow{K} \mathbb{R}$ (or $\mathbb{C}$) p.d., and if

$$\{\delta_x\}_{x \in X} \subset \mathscr{H}(K), \tag{8.4}$$

we shall see that X must then be discrete. (In interesting cases, also countable.) If (8.4) holds, we shall say that (K, X) is a *point process*. We shall further show that point processes arise by restriction as follows:

Let (K, X) be given with K a p.d. kernel. If a countable subset $S \subset X$ is such that $K^{(S)} := K\big|_{S \times S}$ has

$$\{\delta_x\}_{x \in S} \in \mathscr{H}(K^{(S)}), \tag{8.5}$$

then we shall say that $\left(K^{(S)}, S\right)$ is an *induced point process*.

8.1. NETS OF FINITE SUBMATRICES, AND THEIR LIMITS

Given (K, X) as above with K p.d. and defined on $X \times X$. Then the finite submatrices in the subsection header are indexed by the net of all finite subsets F of X as follows: Given F, then the corresponding $|F| \times |F|$ square matrix K_F is simply the restriction of K to $F \times F$. Of course, each matrix K_F is positive definite, and so it has a finite list of eigenvalues. These eigenvalue lists figure in the discussion below.

Lemma 8.1. *Let K, F, and K_F be as above, with $\lambda_s(F)$ denoting the numbers in the list of eigenvalues for the matrix K_F. Then*

$$1 \le \lambda_s(F) \sum_{x \in F} \|\delta_x\|_{\mathscr{H}(K)}^2. \tag{8.6}$$

Proof. Consider the eigenvalue equation

$$(\xi_x)_{x \in F}, \quad \sum_{x \in F} |\xi_x|^2 = \|\xi\|_2^2 = 1, \quad K_F \xi = \lambda_s(F)\xi. \tag{8.7}$$

From Lemma 2.2 and for $x \in F$, we then get

$$\begin{aligned} |\xi_x|^2 &\le \|\delta_x\|_{\mathscr{H}(K)}^2 \langle \xi, K_F \xi \rangle_{l^2(F)} \\ &= \|\delta_x\|_{\mathscr{H}(K)}^2 \lambda_s(F). \end{aligned} \tag{8.8}$$

Now apply $\sum_{x \in F}$ to both sides in (8.8), and the desired conclusion (8.6) follows. $\quad\square$

Remark 8.2. A consequence of the lemma is that the matrices K_F^{-1} and $K_F^{-1/2}$ automatically are well defined (by the spectral theorem) with associated spectral bounds.

Definition 8.3. Let K, F, and K_F be as above; and with the condition $\delta_x \in \mathscr{H}(K)$ in force. Set

$$\mathscr{H}_K(F) := \text{span}_{x \in F} \{K(\cdot, x)\}. \tag{8.9}$$

It is a finite-dimensional (and therefore closed) subspace in $\mathscr{H}(K)$. The orthogonal projection onto $\mathscr{H}_K(F)$ will be denoted $P_F : \mathscr{H}(K) \to \mathscr{H}_K(F)$.

Lemma 8.4. *Let K, F, K_F, and $\mathscr{H}_K(F)$ be as above. Then the orthogonal projection P_F is as follows: For $h \in \mathscr{H}(K)$, set $h_F = h\big|_F$, restriction:*

$$(P_F h)(\cdot) = \sum_{y \in F} (K_F^{-1} h_F)(y) K(\cdot, y). \tag{8.10}$$

Proof. It is immediate from the definition that $P_F h$ has the form

$$P_F h = \sum_{y \in F} \xi_y K(\cdot, y) \tag{8.11}$$

with $(\xi_y)_{y \in F} \in \mathbb{C}^{|F|}$. Since P_F is the orthogonal projection,

$$(h - P_F h) \perp_{\mathscr{H}(K)} \{K(\cdot, y)\}_{y \in F} \tag{8.12}$$

(orthogonality in the $\mathscr{H}(K)$-inner product) which yields

$$h(x) = (K_F \xi)(x) \left(= \sum_{y \in F} K(x, y)\xi_y \right), \quad \forall x \in F;$$

and therefore, $\xi = K_F^{-1} h_F$, which is the desired formula (8.10). $\square$

Corollary 8.5. *Let X, K, $\mathscr{H}(K)$ be as above, and assume $\delta_x \in \mathscr{H}(K)$ for some $x \in X$. Then a function h on X is in $\mathscr{H}(K)$ if and only if*

$$\sup_F \left\| K_F^{-1/2} h_F \right\|_{l^2(F)} < \infty, \tag{8.13}$$

where the supremum is over all finite subsets F of X. If h is finite energy, then

$$\|h\|_{\mathscr{H}(K)}^2 = \sup_F \left\| K_F^{-1/2} h_F \right\|_{l^2(F)}^2. \tag{8.14}$$

Proof. The proof follows from an application of Hilbert space geometry to the RKHS $\mathscr{H}(K)$, on the family of orthogonal projections P_F indexed by the finite subsets F in X. With the standard lattice operations, applied to projections, we have $\sup_F P_F = I_{\mathscr{H}(K)}$. The conclusions (8.13)–(8.14) follow from this since, by the lemma,

$$\|P_F h\|_{\mathscr{H}(K)}^2 \underset{\text{by (2.3)}}{=} \left\langle K_F^{-1} h_F, K_F K_F^{-1} h_F \right\rangle_{l^2(F)}$$
$$= \left\langle h_F, K_F^{-1} h_F \right\rangle_{l^2(F)} = \left\| K_F^{-1/2} h_F \right\|_{l^2(F)}^2. \tag{8.15}$$

$\square$

Remark 8.6. The advantage with the use of this system of orthogonal projections P_F, indexed by the finite subsets F of X, is that we may then take advantage of the known lattice operations for orthogonal projections in Hilbert space. But it is important that we get approximation with respect to the canonical norm in the RKHS $\mathscr{H}(K)$. This works because by our construction, the orthogonality properties for the projections P_F refers precisely to the inner product in $\mathscr{H}(K)$. Naturally we get the best $\mathscr{H}(K)$-approximation properties when X is further assumed countable. But the formula for the $\mathscr{H}(K)$-norm holds in general.

Corollary 8.7. *Let* $X \times X \xrightarrow{K} \mathbb{C}$ *be fixed, assumed p.d., and let* $\mathscr{H}(K)$ *be the corresponding RKHS. Let* $x \in X$ *be given. Then* $\delta_x \in \mathscr{H}(K)$ *if and only if*

$$\sup_{F \subset X,\, F \text{ is finite},\, x \in F} \left(K_F^{-1}\right)_{x,x} < \infty. \tag{8.16}$$

In this case, we have

$$\|\delta_x\|_{\mathscr{H}(K)}^2 = \text{the supremum in } (8.16).$$

Proof. The result is immediate from Corollary 8.5 applied to $h := \delta_x$, where x is fixed. Here the terms in (8.14) are, for F finite, $x \in F$:

$$\left\langle \delta_x\big|_F, K_F^{-1}\left(\delta_x\big|_F\right)\right\rangle_{l^2(F)} = \left(K_F^{-1}\right)_{x,x}, \tag{8.17}$$

and the stated conclusion is now immediate. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Corollary 8.8. *Let* X, K, *and* $\mathscr{H}(K)$ *be as above, but assume now that* X *is countable, with a monotone net of finite sets:*

$$F_1 \subset F_2 \subset F_3 \ldots, \text{ and } X = \cup_{i \in \mathbb{N}} F_i; \tag{8.18}$$

then a function h *on* X *is in* $\mathscr{H}(K)$ *iff* $\sup_i \left\| K_{F_i}^{-1/2} h\big|_{F_i} \right\|_{l^2(F_i)} < \infty$.

Moreover,

$$\|h\|_{\mathscr{H}_E}^2 = \lim_{i \to \infty} \left\| K_{F_i}^{-1/2} h\big|_{F_i} \right\|_{l^2(F_i)}^2, \tag{8.19}$$

where, the convergence in (8.19) *is monotone.*

Proof. From the definition of the order of orthogonal projections, we have

$$P_{F_1} \leq P_{F_2} \leq P_{F_3} \leq \ldots, \tag{8.20}$$

and therefore,

$$\|P_{F_1} h\|_{\mathscr{H}(K)}^2 \leq \|P_{F_2} h\|_{\mathscr{H}(K)}^2 \leq \|P_{F_3} h\|_{\mathscr{H}(K)}^2 \leq \ldots, \tag{8.21}$$

with $\lim_{i \to \infty} \|P_{F_i} h\|_{\mathscr{H}(K)}^2 = \|h\|_{\mathscr{H}(K)}^2$. But by (8.15) and the proof of Corollary 8.5, we have

$$\left\| K_{F_i}^{-1/2} h\big|_{F_i} \right\|_{l^2(F_i)}^2 = \|P_{F_i} h\|_{\mathscr{H}(K)}^2$$

and, so, by (8.21), we get

$$\left\| K_{F_1}^{-1/2} h \big|_{F_1} \right\|_{l^2(F_1)}^2 \leq \left\| K_{F_2}^{-1/2} h \big|_{F_2} \right\|_{l^2(F_2)}^2 \leq \left\| K_{F_3}^{-1/2} h \big|_{F_3} \right\|_{l^2(F_3)}^2 \leq \ldots$$

The conclusion now follows. $\qquad\square$

8.2. RESTRICTIONS OF P.D. KERNELS

Below we shall be considering pairs (K, X) with K a fixed p.d. kernel defined on $X \times X$, and, as before, we denote by $\mathscr{H}(K)$ the corresponding RKHS with its canonical inner product. In general, X is an arbitrary set, typically of large cardinality, in particular uncountable: It may be a complex domain, a generalized boundary, or it may be a manifold arising from problems in physics, in signal processing, or in machine learning models. Moreover, for such general pairs (K, X), with K a fixed p.d. kernel, the Dirac functions δ_x are typically not in $\mathscr{H}(K)$.

Here we shall turn to induced systems, indexed by suitable countable discrete subsets S of X. Indeed, for a number of sampling or interpolation problems, it is possible to identify countable discrete subsets S of X, such that when K is restricted to $S \times S$, i.e., $K^{(S)} := K\big|_{S \times S}$, then for $x \in S$, the Dirac functions δ_x will be in $\mathscr{H}\left(K^{(S)}\right)$; i.e., we get induced point processes indexed by S. In fact, with Corollary 8.8, we will be able to identify a variety of such subsets S.

Moreover, each such choice of subset S yields point-process, and an induced graph, and graph Laplacian; see (8.1)–(8.2). These issues will be taken up in detail in the two subsequent sections. In the following Example 8.9, for illustration, we identify a particular instance of this, when $X = \mathbb{R}$ (the reals), and $S = \mathbb{Z}$ (the integers), and where K is the covariance kernel of standard Brownian motion on $\mathbb{R}$.

Example 8.9 (Discretizing the covariance function for Brownian motion on $\mathbb{R}$)**.** The present example is a variant of Example 6.1, but with $X = \mathbb{R}$ (instead of the interval $[0, 1]$). We now set

$$K(x, y) := \begin{cases} |x| \wedge |y|, & (x, y) \in \mathbb{R} \times \mathbb{R}, \ xy \geq 0; \\ 0, & xy < 0. \end{cases} \tag{8.22}$$

It is immediate that (6.6) in Example 6.1 carries over, but now with $\mathbb{R}$ in place of $[0, 1]$. The normalization $f(0) = 0$ is carried over. We get that: A function $f(x)$ on $\mathbb{R}$ is in $\mathscr{H}(K)$ iff it has distribution-derivative $f' = df/dx$ in $L^2(\mathbb{R})$, see (8.23). As before, we conclude that the $\mathscr{H}(K)$-norm is

$$\|f\|_{\mathscr{H}(K)}^2 = \int_{\mathbb{R}} |f'|^2 \, dx; \tag{8.23}$$

see also Lemma 4.3.

Set

$$K^{(\mathbb{Z})} = K\big|_{\mathbb{Z} \times \mathbb{Z}}, \tag{8.24}$$

and consider the corresponding RKHS $\mathscr{H}\left(K^{(\mathbb{Z})}\right)$. Using [42, 44], we conclude that functions Φ on $\mathbb{Z}$ are in $\mathscr{H}\left(K^{(\mathbb{Z})}\right)$ iff $\Phi\left(0\right) = 0$, and

$$\sum_{n\in\mathbb{Z}}\left|\Phi\left(n\right) - \Phi\left(n+1\right)\right|^2 < \infty.$$

In that case,

$$\left\|\Phi\right\|^2_{\mathscr{H}\left(K^{(\mathbb{Z})}\right)} = \sum_{n\in\mathbb{Z}}\left|\Phi\left(n\right) - \Phi\left(n+1\right)\right|^2. \tag{8.25}$$

For the $\mathbb{Z}$-kernel, we have: $\{\delta_n\}_{n\in\mathbb{Z}} \subset \mathscr{H}\left(K^{(\mathbb{Z})}\right)$, and

$$\delta_n\left(\cdot\right) = 2K\left(\cdot, n\right) - K\left(\cdot, n+1\right) - K\left(\cdot, n-1\right), \quad \forall n \in \mathbb{Z}. \tag{8.26}$$

Moreover, the corresponding Laplacian Δ from (8.1) is

$$\left(\Delta\Phi\right)\left(n\right) = 2\Phi\left(n\right) - \Phi\left(n+1\right) - \Phi\left(n-1\right), \tag{8.27}$$

i.e., the standard discretized Laplacian.

From the matrices $K_F^{(\mathbb{Z})}$, $F \subset \mathbb{Z}$, we have the following; illustrated with $F = F_N = \{1, 2, \ldots, N\}$.

$$K_{F_N}^{(\mathbb{Z})} = \begin{bmatrix} 1 & 1 & 1 & 1 & \ldots & & \ldots & 1 \\ 1 & 2 & 2 & 2 & \ldots & & \ldots & 2 \\ 1 & 2 & 3 & 3 & \ldots & & \ldots & 3 \\ 1 & 2 & 3 & 4 & \ldots & & \ldots & 4 \\ \vdots & \vdots & \vdots & \vdots & \ddots & & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & & N-1 & N-1 \\ 1 & 2 & 3 & 4 & \ldots & & N-1 & N \end{bmatrix}, \tag{8.28}$$

and

$$\left(K_{F_N}^{(\mathbb{Z})}\right)^{-1} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & \ldots & 0 \\ -1 & 2 & -1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 2 & -1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -1 & 2 & -1 & 0 \\ 0 & 0 & \ldots & 0 & -1 & 2 & -1 \\ 0 & 0 & \ldots & 0 & 0 & -1 & 1 \end{bmatrix}. \tag{8.29}$$

In particular, for $n, m \in \mathbb{Z}$, we have

$$\left\langle \delta_n, \delta_m \right\rangle_{\mathscr{H}\left(K^{(\mathbb{Z})}\right)} = \begin{cases} 2 & \text{if } n = m, \\ -1 & \text{if } |n - m| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 8.10. The determinant of $K_{F_N}^{(\mathbb{Z})}$ is 1 for all N. *Proof.* By eliminating the first column, and then the first row, $\det(K_{F_N}^{(\mathbb{Z})})$ is reduced to $\det(K_{F_{N-1}}^{(\mathbb{Z})})$. So by induction, the determinant is 1.

Note that

$$\sum_{k \in \mathbb{Z}} \chi_{[1,n]}(k) \chi_{[1,m]}(k) = n \wedge m$$

which yields the factorization

$$K_{F_N}^{(\mathbb{Z})} = A_N A_N^*, \tag{8.30}$$

i.e.,

$$K_{F_N}^{(\mathbb{Z})}(n,m) = (A_N A_N^*)_{n,m} = \sum A_N(n,k) A_N^*(k,m),$$

where A_N is the $N \times N$ lower triangular matrix given by

$$A_N = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 1 & 1 & \dots & \dots & 1 \end{bmatrix}.$$

In particular, we get that $\det(K_{F_N}^{(\mathbb{Z})}) = 1$ immediately. This is a special case of Theorem 5.1.

For the general case, let $F_N = \{x_j\}_{j=1}^N$ be a finite subset of $\mathbb{R}$, assuming $x_1 < x_2 < \dots < x_N$. Then the factorization (8.30) holds with

$$A_N = \begin{bmatrix} \sqrt{x_1} & 0 & 0 & \dots & 0 \\ \sqrt{x_1} & \sqrt{x_2 - x_1} & 0 & \dots & \vdots \\ \sqrt{x_1} & \sqrt{x_2 - x_1} & \sqrt{x_3 - x_2} & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \sqrt{x_1} & \sqrt{x_2 - x_1} & \sqrt{x_3 - x_2} & \dots & \sqrt{x_N - x_{N-1}} \end{bmatrix}. \tag{8.31}$$

Thus,

$$\det(K_{F_N}^{(\mathbb{Z})}) = x_1(x_2 - x_1)\dots(x_N - x_{N-1}). \tag{8.32}$$

In the setting of Section 5 (finite sums of standard Gaussians), we have the following: Let $\{x_i\}_{i=1}^N$ be as in (8.31), and let $1 \le n, m \le N$. Let $\{Z_i\}_{i=1}^N$ be a system i.i.d. standard Gaussians $N(0,1)$, i.e., independent identically distributed. Set

$$V_n = Z_1\sqrt{x_1} + Z_2\sqrt{x_2 - x_1} + \dots + Z_n\sqrt{x_n - x_{n-1}}. \tag{8.33}$$

Then one checks that

$$\mathbb{E}(V_n V_m) = x_n \wedge x_m = K(x_n, x_m) \tag{8.34}$$

which is the desired Gaussian realization of K.

Alternatively, $K_{F_N}^{(\mathbb{Z})}$ assumes the following factorization via non-square matrices: Assume $F_N \subset \mathbb{Z}_+$, then

$$K_{F_N}^{(\mathbb{Z})} = AA^*, \tag{8.35}$$

where A is the $N \times x_N$ matrix such that

$$A_{n,k} = \begin{cases} 1 & \text{if } 1 \le k \le x_n, \\ 0 & \text{otherwise.} \end{cases}$$

That is, A takes the form:

$$A = \begin{bmatrix} \underbrace{1 \quad \dots \quad 1}_{x_1} & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \underbrace{1 \quad \dots \quad \dots \quad \dots \quad 1}_{x_2} & 0 & \dots & \dots & \dots & 0 \\ \underbrace{1 \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad 1}_{x_3} & 0 & \dots & 0 \\ \vdots \quad \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \vdots \quad \vdots & & \vdots & \vdots & & \vdots & \vdots & 0 \\ 1 \quad 1 & \dots & \dots & \dots & \dots & \dots & \dots & 1 \quad 1 \end{bmatrix}. \tag{8.36}$$

$$\underbrace{\qquad\qquad\qquad\qquad}_{x_N}$$

Remark 8.11 (Spectrum of the matrices K_F; see also [31])**.** It is known that the factorization as in (8.30) can be used to obtain the spectrum of positive definite matrices. The algorithm is as follows: Let K be a given p.d. matrix.

Initialization: $B := K$;

Iterations: $k = 1, 2, \dots, n - 1$,

(i) $B = AA^*$,

(ii) $B = A^*A$.

Here A in step (i) denotes the lower triangular matrix in the Cholesky decomposition of B (see (8.30)). Then $\lim_{n \to \infty} B$ converges to a diagonal matrix consisting of the eigenvalues of K.

We now resume consideration of the general case of p.d. kernels K on $X \times X$ and their restrictions: a setting for harmonic functions.

Remark 8.12. In the general case of (8.2) and Lemma 8.1, we still have a Laplace operator Δ. It is a densely defined symmetric operator on $\mathscr{H}(K)$. Moreover (general case),

$$\Delta . K(\cdot, x) = \delta_x(\cdot), \quad \forall x \in X \tag{8.37}$$

(assuming that $\delta_x \in \mathscr{H}(K)$). The dot "$\cdot$" in (8.37) refers to the action variable for the operator Δ. In other words, $K(\cdot, \cdot)$ is a generalized Greens kernel.

Definition 8.13. Let $X \times X \xrightarrow{K} \mathbb{C}$ be given p.d., and assume

$$\{\delta_x\}_{x \in X} \subset \mathscr{H}(K). \tag{8.38}$$

Let Δ denote the induced Laplace operator. A function h (in $\mathscr{H}(K)$) is said to be *harmonic* iff $\Delta h = 0$.

Corollary 8.14. *Let $(X, K, \mathscr{H}(K))$ be as above. Assume (8.38), and let Δ be the induced Laplace operator. Then we have the following orthogonal decomposition for $\mathscr{H}(K)$:*

$$\mathscr{H}(K) = \{h \; ; \Delta h = 0\} \oplus \operatorname{clospan}^{\mathscr{H}(K)}\left(\{\delta_x\}_{x \in X}\right) \tag{8.39}$$

where "clospan" in (8.39) refers to the norm in $\mathscr{H}(K)$.

Proof. It is immediate from (8.1) that

$$\{h \in \mathscr{H}(K) \; ; \Delta h = 0\} = \left(\{\delta_x\}_{x \in X}\right)^{\perp} \tag{8.40}$$

where the orthogonality "$\perp$" in (8.40) refers to the inner product $\langle \cdot, \cdot \rangle_{\mathscr{H}(K)}$. Since, by Hilbert space geometry, $\left(\{\delta_x\}_{x \in X}\right)^{\perp\perp} = \operatorname{clospan}^{\mathscr{H}(K)}\left(\{\delta_x\}_{x \in X}\right)$, we only need to observe that $\{h \in \mathscr{H}(K) \; ; \Delta h = 0\}$ is closed in $\mathscr{H}(K)$. But this is immediate from (8.1). $\qquad \square$

Corollary 8.15 (Duality). *Let $X \times X \xrightarrow{K} \mathbb{R}$ be given, assumed p.d., and let $S \subset X$ be a countable subset such that*

$$\mathscr{D}(S) := \{\delta_x\}_{x \in S} \subset \mathscr{H}(K^{(S)}). \tag{8.41}$$

(i) *Then the following duality holds for the two induced kernels:*

$$K^{(S)} := K\big|_{S \times S},$$

   and

$$D^{(S)}(x, y) := \langle \delta_x, \delta_y \rangle_{\mathscr{H}\left(K^{(S)}\right)}, \; \forall (x, y) \in S \times S;$$

   both p.d. kernels on $S \times S$.
   For every pair $x, y \in S$, we have the following matrix-inversion formula:

$$\sum_{z \in S} D^{(S)}(x, z) K^{(S)}(z, y) = \delta_{x,y}, \tag{8.42}$$

   where the summation on the LHS in (8.42) is a limit over a net of finite subsets $\{F_i\}_{i \in \mathbb{N}}$, $F_1 \subset F_2 \subset \ldots$, such that $\cup_i F_i = S$; and the result is independent of choice of net.

(ii) *We get an <u>induced graph</u> with S as the set of vertices, and edge set E as follows: $E \subset (S \times S) \setminus (diagonal)$.*
   An edge is a pair $(x, y) \in (S \times S) \setminus (diagonal)$ such that

$$\langle \delta_x, \delta_y \rangle_{\mathscr{H}\left(K^{(S)}\right)} \neq 0.$$

Proof. The result follows from an application of Corollaries 8.7 and 8.8, and Remark 8.12. □

Let X, K, and S be as stated, S countable infinite, with assumptions as in the previous two results. We showed that then the subset S acquires the structure of a vertex set in an induced infinite graph (Corollary 8.15 (ii)). If Δ denotes the corresponding graph Laplacian, then the following boundary value problem is of great interest: Make precise the boundary conditions at "infinity" for this graph Laplacian Δ. An answer to this will require identification of Hilbert space, and limit at "infinity". The result below is such an answer, and the limit notion will be, limit over the filter of all finite subsets in S; see Corollary 8.7. Another key tool in the arguments below will again be the net of orthogonal projections $\{P_F\}$ from Lemma 8.4, and the convergence results from Corollaries 8.5 and 8.7.

Corollary 8.16. *Let $X \times X \xrightarrow{K} \mathbb{R}$, and $S \subset X$ be as in the statement of Corollary 8.15. Let $\mathscr{F}_{fin}(S)$ denote the filter of finite subsets $F \subset S$. Let $\Delta = \Delta_S$ be the graph Laplacian defined in* (8.2), *i.e.,*

$$(\Delta h)(x) := \langle \delta_x, h \rangle_{\mathscr{H}(K^{(S)})},$$

for all $x \in S$, $h \in \mathscr{H}(K^{(S)})$. Then the following equivalent conditions hold:

(i) *For all $h \in \mathscr{H}(K^{(S)})$,*

$$\begin{aligned}
\|h\|^2_{\mathscr{H}(K^{(S)})} &= \sup_{F \in \mathscr{F}_{fin}(S)} \left\langle h\big|_F, \Delta P_F h \right\rangle_{l^2(F)} \\
&= \sup_{F \in \mathscr{F}_{fin}(S)} \left\langle h\big|_F, K_F^{-1}\left(h\big|_F\right) \right\rangle_{l^2(F)}.
\end{aligned} \tag{8.43}$$

(ii) *For $\forall F \in \mathscr{F}_{fin}(S)$, $x \in F$, $h \in \mathscr{H}(K^{(S)})$,*

$$(\Delta(P_F h))(x) = \left(K_F^{-1}\left(h\big|_F\right)\right)(x). \tag{8.44}$$

(iii) $K_F \Delta P_F h = h\big|_F.$

Proof. On account of Corollary 8.8, we only need to verify (8.44). Let $F \in \mathscr{F}_{fin}(S)$, $h \in \mathscr{H}(K^{(S)})$, then we proved that

$$(P_F h)(\cdot) = \sum_{y \in F} \xi_y K(\cdot, y) \quad \text{with} \tag{8.45}$$

$$\xi_y = \left(K_F^{-1}\left(h\big|_F\right)\right)(y). \tag{8.46}$$

Now apply $\langle \delta_x, \cdot \rangle_{\mathscr{H}(K^{(S)})}$ to both sides in (8.45); and we get

$$(\Delta(P_F h))(x) = \xi_x, \tag{8.47}$$

where we used $\langle \delta_x, K(\cdot, y) \rangle_{\mathscr{H}(K^{(S)})} = \delta_{x,y}$. The desired conclusion (8.44) now follows from (8.47). Also note that $(\Delta(P_F h))(x) = 0$ if $x \in X \backslash F$. □

8.3. CANONICAL ISOMETRIES COMPUTED FROM POINT PROCESSES

Below we consider p.d. kernels K defined initially on $X \times X$. Our present aim is to consider restrictions to $S \times S$ when S is a suitable subset of X. Our first observation is the identification of a canonical isometry T_S between the respective reproducing kernel Hilbert spaces; T_S identifying $\mathscr{H}(K^{(S)})$ as an isometric subspace inside $\mathscr{H}(K)$. This isometry T_S exists in general. However, we shall show that, when the subset S is further restricted, the respective RKHSs, and isometry T_S will admit explicit characterizations. For example, if S is countable, and is the Dirac functions δ_s, $s \in S$, are in $\mathscr{H}(K^{(S)})$ we shall show that this setting leads to a point process. In this case, we further identify an induced (infinite) graph with the set S as vertices, and with associated edges defined by an induced δ_s kernel.

Theorem 8.17. *Let* $X \times X \xrightarrow{K} \mathbb{C}$ *be a p.d. kernel, and let* $S \subset X$ *be a subset. Set* $K^{(S)} := K\big|_{S \times S}$. *Let* $\mathscr{H}(K)$, *and* $\mathscr{H}(K^{(S)})$, *be the respective RKHSs.*

(i) *Then there is a canonical isometric embedding*

$$\mathscr{H}(K^{(S)}) \xrightarrow{T} \mathscr{H}(K),$$

given by the following formula: For $s \in S$, *set*

$$T(K^{(S)}(\cdot, s)) = K(\cdot, s). \tag{8.48}$$

(Note that $K^{(S)}(\cdot, s)$ *on the LHS in (8.48) is a function on* S, *while* $K(\cdot, s)$ *on the RHS is a function on* X.)

(ii) *The adjoint operator* T^*,

$$\mathscr{H}(K) \xrightarrow{T^*} \mathscr{H}(K^{(S)}) \tag{8.49}$$

is given by restriction, i.e., if $f \in \mathscr{H}(K)$, *and* $s \in S$, *then* $(T^* f)(s) = f(s)$; *or equivalently, for all* $f \in \mathscr{H}(K)$,

$$T^* f = f\big|_S. \tag{8.50}$$

Proof. To show that T in (8.48) is isometric, proceed as follows: Let $\{s_i\}_{i=1}^N$ be a finite subset of S, and $\{\xi_i\}_{i=1}^N \in \mathbb{C}^N$, then

$$\left\| T\left(\sum_i \xi_i K^{(S)}(\cdot, s_i) \right) \right\|_{\mathscr{H}(K)}^2 = \left\| \sum_i \xi_i T(K^{(S)}(\cdot, s_i)) \right\|_{\mathscr{H}(K)}^2$$

$$\underset{\text{by (8.48)}}{=} \left\| \sum_i \xi_i K(\cdot, s_i) \right\|_{\mathscr{H}(K)}^2$$

$$\underset{\text{by (2.3)}}{=} \sum_i \sum_j \overline{\xi}_i \xi_j K(s_i, s_j)$$

$$= \sum_i \sum_j \overline{\xi}_i \xi_j K^{(S)}(s_i, s_j)$$

$$= \left\| \sum_i \xi_i K^{(S)}(\cdot, s_i) \right\|_{\mathscr{H}(K^{(S)})}^2$$

which is the desired isometric property.

We now turn to (8.50), the restriction formula: Let $s \in S$, and $f \in \mathscr{H}(K)$, then

$$\left\langle T(K^{(S)}(\cdot, s)), f \right\rangle_{\mathscr{H}(K)} = \left\langle K^{(S)}(\cdot, s), T^* f \right\rangle_{\mathscr{H}(K^{(S)})} \tag{8.51}$$

$$\underset{\text{by } (2.5)}{=} (T^* f)(s).$$

But, for the LHS in (8.51), we have

$$\left\langle T(K^{(S)}(\cdot, s)), f \right\rangle_{\mathscr{H}(K)} \underset{\text{by } (8.48)}{=} \left\langle K(\cdot, s), f \right\rangle_{\mathscr{H}(K)} \underset{\text{by } (2.5)}{=} f(s);$$

and so the desired formula (8.50) follows. $\square$

Remark 8.18. The canonical isometry for Example 8.9 ($\mathbb{Z}$-discretization of the covariance function for Brownian motion on $\mathbb{R}$). From Theorem 8.17, we know that the canonical isometry T maps $\mathscr{H}(K^{(Z)})$ into $\mathscr{H}(K)$; see (8.22). But (8.23) and (8.25) in the Example offer exact characterization of these two Hilbert spaces. So, in the special case of Example 8.9, the canonical isometry T maps from functions Φ on $\mathbb{Z}$ into functions on $\mathbb{R}$. In view of (8.23), this assignment turns out to be a precise spline realization of the point grids realized by these sequences Φ.

Below we present an explicit formula, and graphics, for the spline realizations. By (8.26), the embedding of δ_n from $\mathscr{H}(K^{(\mathbb{Z})})$ into $\mathscr{H}(K)$ is given by

$$(T\delta_n)(x) = 2K(x, n) - K(x, n+1) - K(x, n-1), \quad \forall x \in \mathbb{R}.$$
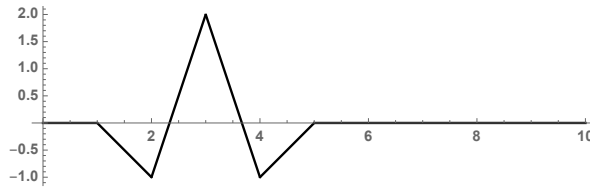
See Figure 9. Therefore, for all $h \in \mathscr{H}(K)$, we get

$$(T^* h)(m) = \sum_{n \in \mathbb{Z}} h(n) \delta_n(m), \quad m \in \mathbb{Z},$$

and

$$(TT^* h)(x) = \sum_{n \in \mathbb{Z}} h(n)(2K(x, n) - K(x, n+1) - K(x, n-1)), \quad x \in \mathbb{R}$$

which is the spline interpolation.

Fig. 9. Isometric extrapolation from functions on $\mathbb{Z}$ to functions on $\mathbb{R}$. An illustration of the isometric embedding of δ_n from $\mathscr{H}(K^{(\mathbb{Z})})$ into $\mathscr{H}(K)$, with $n = 3$

Corollary 8.19. *Let* $X \times X \xrightarrow{K} \mathbb{C}$ *be a p.d. kernel, and let* $S \subset X$ *be a subset. Assume further that* $\{\delta_s\}_{s \in S} \subset \mathscr{H}(K^{(S)})$. *Then every finitely supported function h on S is in* $\mathscr{H}(K^{(S)})$, *and we have the following generalized spline interpolation; i.e., isometrically extending h from S to X:*

$$\widetilde{h}(x) = \sup_{F \supset F_0} \sum_{y \in F} \left(K_F^{-1} h_F \right)(y) K(y,x), \ x \in X, \tag{8.52}$$

where $F_0 = \mathrm{suppt}(h)$, *and the sup is taken over the filter of all finite subsets of X containing F_0.*

Proof. Assume $h \in \mathscr{H}(K^{(S)})$, supported on a finite subset $F_0 \subset S$. Then,

$$\begin{aligned}
\widetilde{h}(x) := Th(x) &= T\left(\sum_{s \in F_0} h(s)\,\delta_s \right)(x) \\
&= \sum_{s \in F_0} h(s)\,(T\delta_s)(x) \\
&= \sum_{s \in F_0} h(s) \sup_{F \supset F_0} (P_F \delta_s)(x) \\
&= \sup_{F \supset F_0} P_F \left(\sum_{s \in F_0} h(s)\,\delta_s \right)(x) \\
&= \sup_{F \supset F_0} (P_F h_{F_0})(x) \\
&= \sup_{F \supset F_0} \sum_{y \in F} \left(K_F^{-1} h_F \right)(y) K(x,y),
\end{aligned}$$

where the last step follows from (8.10), and P_F is the orthogonal projection from $\mathscr{H}(K)$ onto the subspace $\mathscr{H}_K(F)$. $\qquad\square$

Corollary 8.20. *Let* $X \times X \xrightarrow{K} \mathbb{C}$, *p.d. be given, and let* $S \subset X$ *be a subset. Let* $T = T_S$, $\mathscr{H}(K^{(S)}) \xrightarrow{T} \mathscr{H}(K)$, *be the canonical isometry. Then a function f in* $\mathscr{H}(K)$ *satisfies* $\langle f, T(\mathscr{H}(K^{(S)})) \rangle_{\mathscr{H}(K)} = 0$ *if and only if*

$$f(s) = 0 \text{ for all } s \in S. \tag{8.53}$$

Proof. Immediate from part (ii) in Theorem 8.17. $\qquad\square$

Remark 8.21. Let (X, K, S) be as in Corollary 8.20, and let T_S be the canonical isometry. Let $P_S := T_S T_S^*$ be the corresponding projection. Then $I_{\mathscr{H}(K)} - P_S$ is the projection onto the subspace given in (8.53).

Corollary 8.22. *Let* $X \times X \xrightarrow{K} \mathbb{C}$ *be given p.d.; and let* $S \subset X$ *be a subset with induced kernel*

$$K^{(S)} := K\big|_{S \times S}. \tag{8.54}$$

Consider the two sets $\mathscr{F}(S)$ *and* $\mathscr{F}(K^{(S)})$ *from (5.2) and Theorem 5.1. Let* $T_S : \mathscr{H}(K^{(S)}) \to \mathscr{H}(K)$ *be the canonical isometry (8.48) in Theorem 8.17. Then the following implication holds:*

$$\begin{aligned}
\left(\{k_x\}_{x \in X}, \mu \right) &\in& \mathscr{F}(K) \tag{8.55} \\
&\Downarrow& \\
\left(\{k_s\}_{s \in S}, \mu \right) &\in& \mathscr{F}(K^{(S)}) \tag{8.56}
\end{aligned}$$

Proof. Assuming (8.55), we get the representation (5.2):

$$K(x,y) = \int\limits_M \overline{k}_x k_y \, d\mu, \quad \forall (x,y) \in X \times X. \tag{8.57}$$

But then, for all $(s_1, s_2) \in S \times S$, we then have

$$
\begin{aligned}
K^{(S)}(s_1, s_2) \quad &= \quad \left\langle T_S(K^{(S)}(\cdot, s_1)), T_S(K^{(S)}(\cdot, s_2)) \right\rangle_{\mathscr{H}(K)} \\
&\underset{\text{by (8.54)}}{=} \quad K(s_1, s_2) \\
&= \quad \int\limits_M \overline{k}_{s_1} k_{s_2} \, d\mu,
\end{aligned}
$$

which is the desired conclusion. $\square$

9. BOUNDARY VALUE PROBLEMS

Our setting in the present section is the discrete case, i.e., RKHSs of functions defined on a prescribed countable infinite discrete set S. We are concerned with a characterization of those RKHSs $\mathscr{H}$ which contain the Dirac masses δ_x for all points $x \in S$. Of the examples and applications where this question plays an important role, we emphasize two: (i) discrete Brownian motion-Hilbert spaces, i.e., discrete versions of the Cameron-Martin Hilbert space; (ii) energy-Hilbert spaces corresponding to graph-Laplacians.

The problems addressed here are motivated in part by applications to analysis on infinite weighted graphs, to stochastic processes, and to numerical analysis (discrete approximations), and to applications of RKHSs to machine learning. Readers are referred to the following papers, and the references cited there, for details regarding this: [7, 8, 10, 21, 36, 37, 40, 55, 59, 60, 76].

The discrete case can be understood as restrictions of analogous PDE-models. In traditional numerical analysis, one builds discrete and algorithmic models (finite element methods), each aiming at finding approximate solutions to PDE-boundary value problems. They typically use multiresolution-subdivision schemes, applied to the continuous domain, subdividing into simpler discretized parts, called finite elements. And with variational methods, one then minimize various error-functions. In this paper, we turn the tables: our object of study are the discrete models, and analysis of suitable continuous PDE boundary problems serve as a tool for solutions in the discrete world.

Definition 9.1. Let $X \times X \xrightarrow{K} \mathbb{C}$ be a given p.d. kernel on X. The RKHS $\mathscr{H} = \mathscr{H}(K)$ is said to have the *discrete mass* property ($\mathscr{H}$ is called a *discrete RKHS*) if $\delta_x \in \mathscr{H}$, for all $x \in X$.

In fact, it is known ([44]) that every fundamental solution for a Dirichlet boundary value problem on a bounded open domain Ω in $\mathbb{R}^\nu$, allows for discrete restrictions (i.e., vertices sampled in Ω), which have the desired "discrete mass" property.

We recall the following result to stress the distinction of the discrete models vs their continuous counterparts.

Let Ω be a bounded, open, and connected domain in $\mathbb{R}^\nu$ with smooth boundary $\partial\Omega$. Let $K : \Omega \times \Omega \to \mathbb{R}$ continuous, p.d., given as the Green's function of Δ_0, where

$$\Delta_0 := -\sum_{j=1}^{\nu} \left(\frac{\partial}{\partial x_j} \right)^2,$$

$$\mathrm{dom}\,(\Delta_0) = \left\{ f \in L^2(\Omega) \mid \Delta f \in L^2(\Omega), \text{ and } f\big|_{\partial\Omega} \equiv 0 \right\}. \tag{9.1}$$

for the Dirichlet boundary condition. Thus, Δ_0 is positive selfadjoint, and

$$\Delta_0 K = \delta(x - y) \text{ on } \Omega \times \Omega, \tag{9.2}$$

$$K(x, \cdot)\big|_{\partial\Omega} \equiv 0. \tag{9.3}$$

Let $\mathscr{H}_{CM}(\Omega)$ be the corresponding Cameron-Martin RKHS.
For $\nu = 1$, $\Omega = (0,1)$, take

$$\mathscr{H}_{CM}(0,1) = \Big\{ f \mid f' \in L^2(0,1),\ f(0) = f(1) = 0,$$

$$\|f\|_{CM}^2 := \int_0^1 |f'|^2\, dx < \infty \Big\} \tag{9.4}$$

For $\nu > 1$, let

$$\mathscr{H}_{CM}(\Omega) = \left\{ f \mid \nabla f \in L^2(\Omega),\ f\big|_{\partial\Omega} \equiv 0,\ \|f\|_{CM}^2 := \int_\Omega |\nabla f|^2\, dx < \infty \right\},$$

$$\text{where } \nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_\nu} \right). \tag{9.5}$$

Theorem 9.2. *Let Ω, and $S \subset \Omega$, be given. Then*

(i) *Discrete case: Fix $S \subset \Omega$, $\#S = \aleph_0$, where $S = \{x_j\}_{j=1}^\infty$, $x_j \in \Omega$. Assume that there exists $\varepsilon > 0$ such that $\|x_i - x_j\| \geq \varepsilon$ for all i, j, where $i \neq j$. Let*

$$\mathscr{H}(S) = RKHS \text{ of } K^{(S)} := K\big|_{S \times S};$$

then $\delta_{x_j} \in \mathscr{H}(S)$.

(ii) *Continuous case: By contrast, $K_x^{(S)} \in \mathscr{H}_{CM}(S)$, but $\delta_x \notin \mathscr{H}_{CM}(\Omega)$, $x \in \Omega$.*

Proof. The result follows from an application of Corollaries 8.7 and 8.8. It extends earlier results [42, 44] by the co-authors. $\qquad\square$

10. SAMPLING IN $\mathscr{H}(K)$

In the present section, we study classes of reproducing kernels K on general domains with the property that there are non-trivial restrictions to countable discrete sample subsets S such that every function in $\mathscr{H}(K)$ has an S-sample representation. In this general framework, we study properties of positive definite kernels K with respect to sampling from "small" subsets, and applying to all functions in the associated Hilbert space $\mathscr{H}(K)$.

We are motivated by concrete kernels which are used in a number of applications, for example, on one extreme, the Shannon kernel for band-limited functions, which admits many sampling realizations; and on the other, the covariance kernel of Brownian motion which has no non-trivial countable discrete sample subsets.

Definition 10.1. Let $X \times X \xrightarrow{K} \mathbb{C}$ be a p.d. kernel, and $\mathscr{H}(K)$ be the associated RKHS. We say that K has non-trivial sampling property, if there exists a countable subset $S \subset X$, and $a, b \in \mathbb{R}_+$, such that

$$a \sum_{s \in S} |f(s)|^2 \le \|f\|^2_{\mathscr{H}(K)} \le b \sum_{s \in S} |f(s)|^2, \quad \forall f \in \mathscr{H}(K). \tag{10.1}$$

If equality holds in (10.1) with $a = b = 1$, then we say that $\{K(\cdot, s)\}_{s \in S}$ is a Parseval frame. (Also see Definition 7.1.)

It follows that sampling holds in the form

$$f(x) = \sum_{s \in S} f(s) K(x, s), \quad \forall f \in \mathscr{H}(K), \forall x \in X$$

if and only if $\{K(\cdot, s)\}_{s \in S}$ is a Parseval frame.

Lemma 10.2. *Suppose K, X, a, b, and S satisfy the condition in* (10.1)*, then the linear span of $\{K(\cdot, s)\}_{s \in S}$ is dense in $\mathscr{H}(K)$. Moreover, there is a positive operator B in $\mathscr{H}(K)$ with bounded inverse such that*

$$f(\cdot) = \sum_{s \in S} (Bf)(s) K(\cdot, s)$$

is a convergent interpolation formula valid for all $f \in \mathscr{H}(K)$. Equivalently,

$$f(x) = \sum_{s \in S} f(s) B(K(\cdot, s))(x), \quad \forall x \in X.$$

Proof. Define $A : \mathscr{H}(K) \to l^2(S)$ by $(Af)(s) = f(s)$, $s \in S$. Then the adjoint operator $A^* : l^2(S) \to \mathscr{H}(K)$ is given by $A^*\xi = \sum_{s \in S} \xi_s K(\cdot, s)$, for all $\xi \in l^2(S)$, and

$$A^*Af = \sum_{s \in S} f(s) K(\cdot, s)$$

holds in $\mathscr{H}(K)$, with $\mathscr{H}(K)$-norm convergence. Now set $B = (A^*A)^{-1}$, and note that $\|B\|_{\mathscr{H}(K) \to \mathscr{H}(K)} \le a^{-1}$, where a is in the lower bound in (10.1). $\qquad\square$

Theorem 10.3. *Let $K : X \times X \to \mathbb{R}$ be a p.d. kernel, and let $S \subset X$ be a countable discrete subset. For all $s \in S$, set $K_s(\cdot) = K(\cdot, s)$. Then the following conditions are equivalent:*

(i) *The family $\{K_s\}_{s \in S}$ is a Parseval frame in $\mathscr{H}(K)$;*

(ii)

$$\|f\|^2_{\mathscr{H}(K)} = \sum_{s \in S} |f(s)|^2, \quad \forall f \in \mathscr{H}(K);$$

(iii)

$$K(x, x) = \sum_{s \in S} |K(x, s)|^2, \quad \forall x \in X;$$

(iv)

$$f(x) = \sum_{s \in S} f(s) K(x, s), \quad \forall f \in \mathscr{H}(K), \ \forall x \in X,$$

where the sum converges in the norm of $\mathscr{H}(K)$.

Proof. The proof is simple, and follows the steps in the proof of Lemma 7.2. Details are left to the reader. □

We now turn to dichotomy: Existence of countably discrete sampling sets vs non-existence.

Example 10.4. Let $X = \mathbb{R}$, and let $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be the Shannon kernel, where

$$
\begin{aligned}
K(x, y) &:= \operatorname{sinc} \pi(x - y) \\
&= \frac{\sin \pi(x - y)}{\pi(x - y)}, \quad \forall x, y \in \mathbb{R}.
\end{aligned}
\tag{10.2}
$$

We may choose $S = \mathbb{Z}$, and then $\{K(\cdot, n)\}_{n \in \mathbb{Z}}$ is even an orthonormal basis (ONB) in $\mathscr{H}(K)$, but there are many other examples of countable discrete subsets $S \subset \mathbb{R}$ such that (10.1) holds for finite $a, b \in \mathbb{R}_+$.

The RKHS $\mathscr{H}(K)$ in (10.2) is the Hilbert space $\subset L^2(\mathbb{R})$ consisting of all $f \in L^2(\mathbb{R})$ such that $\operatorname{suppt}(\hat{f}) \subset [-\pi, \pi]$, where "suppt" stands for support of the Fourier transform $\hat{f}$. Note $\mathscr{H}(K)$ consists of functions on $\mathbb{R}$ which have entire analytic extensions to $\mathbb{C}$. Using the above observations, we get

$$
\begin{aligned}
f(x) &= \sum_{n \in \mathbb{Z}} f(n) K(x, n) \\
&= \sum_{n \in \mathbb{Z}} f(n) \operatorname{sinc} \pi(x - n), \quad \forall x \in \mathbb{R}, \ \forall f \in \mathscr{H}(K).
\end{aligned}
$$

Example 10.5. Let K be the covariant kernel of standard Brownian motion, with $X := [0, \infty)$ or $[0, 1)$, and
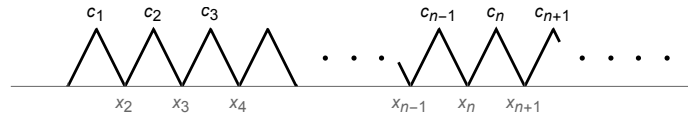
$$K(x, y) := x \wedge y = \min(x, y), \quad \forall (x, y) \in X \times X. \tag{10.3}$$

Theorem 10.6. *Let K, X be as in* (10.3); *then there is no countable discrete subset $S \subset X$ such that $\{K(\cdot, s)\}_{s \in S}$ is dense in $\mathscr{H}(K)$.*

Proof. Suppose $S = \{x_n\}$, where

$$0 < x_1 < x_2 < \ldots < x_n < x_{n+1} < \ldots \qquad (10.4)$$

Then consider the function given in Figure 10.

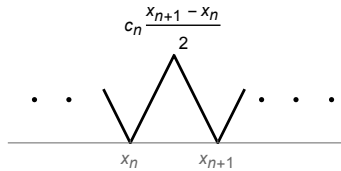**Fig. 10.** Example of a saw-tooth function vanishing on S

On the respective intervals $[x_n, x_{n+1}]$, the function f is as follows:

$$f(x) = \begin{cases} c_n(x - x_n) & \text{if } x_n \leq x \leq \frac{x_n + x_{n+1}}{2}, \\ c_n(x_{n+1} - x) & \text{if } \frac{x_n + x_{n+1}}{2} < x \leq x_{n+1}. \end{cases}$$

In particular, $f(x_n) = f(x_{n+1}) = 0$, and on the midpoints:

$$f\left(\frac{x_n + x_{n+1}}{2}\right) = c_n \frac{x_{n+1} - x_n}{2},$$

see Figure 11.

Fig. 11. The saw-tooth function

Choose $\{c_n\}_{n \in \mathbb{N}}$ such that

$$\sum_{n \in \mathbb{N}} |c_n|^2 (x_{n+1} - x_n) < \infty. \qquad (10.5)$$

Admissible choices for the slope-values c_n include

$$c_n = \frac{1}{n\sqrt{x_{n+1} - x_n}}, \quad n \in \mathbb{N}.$$

We will now show that $f \in \mathscr{H}(K)$. For the distribution derivative computed from the function given in Figure 10, we get
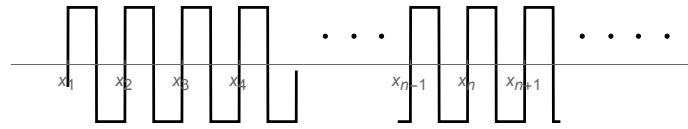
Fig. 12. The distribution derivative of the saw-tooth function

$$\int\limits_0^\infty |f'(x)|^2 \, dx = \sum_{n \in \mathbb{N}} |c_n|^2 (x_{n+1} - x_n) < \infty$$

which is the desired conclusion; see Figure 12. $\qquad\square$

Corollary 10.7. *For the kernel $K(x,y) = x \wedge y$ in (10.3), $X = [0, \infty)$, the following holds: Given $\{x_j\}_{j \in \mathbb{N}} \subset \mathbb{R}_+$, $\{y_j\}_{j \in \mathbb{N}} \subset \mathbb{R}$, then the interpolation problem*

$$f(x_j) = y_j, \quad f \in \mathscr{H}(K) \tag{10.6}$$

is solvable if

$$\sum_{j \in \mathbb{N}} (y_{j+1} - y_j)^2 / (x_{j+1} - x_j) < \infty. \tag{10.7}$$

Proof. Let f be the piecewise linear spline (see Figure 13) for the problem (10.6). Then the $\mathscr{H}(K)$-norm is as follows:

$$\int\limits_0^\infty |f'(x)|^2 \, dx = \sum_{j \in \mathbb{N}} \left(\frac{y_{j+1} - y_j}{x_{j+1} - x_j} \right)^2 (x_{j+1} - x_j) < \infty$$
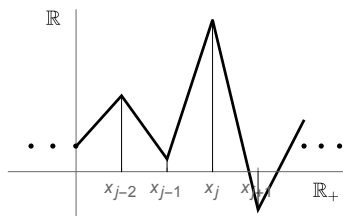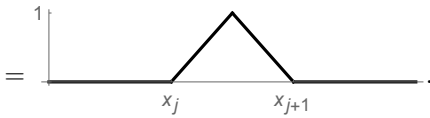
when (10.7) holds. $\qquad\square$

Fig. 13. Piecewise linear spline

Remark 10.8. Let K be as in (10.3), $X = [0, \infty)$. For all $0 \leq x_j < x_{j+1} < \infty$, let

$$f_j(x) := \frac{2}{x_{j+1} - x_j} \left(K\left(x - x_j, \frac{x_{j+1} - x_j}{2}\right) - K\left(x - \frac{x_j + x_{j+1}}{2}, \frac{x_{j+1} - x_j}{2}\right) \right)$$

Assuming (10.5) holds, then

$$f(x) = \sum_j c_j f_j(x) \in \mathscr{H}(K).$$

Theorem 10.9. *Let X be a set of cardinality c of the continuum, and let $K : X \times X \to \mathbb{R}$ be a positive definite kernel. Let $S = \{x_j\}_{j \in \mathbb{N}}$ be a discrete subset of X. Suppose there are weights $\{w_j\}_{j \in \mathbb{N}}$, $w_j \in \mathbb{R}_+$, such that*

$$(f(x_j)) \in l^2(\mathbb{N}, w) \tag{10.8}$$

for all $f \in \mathscr{H}(K)$. Suppose further that there is a point $t_0 \in X \backslash S$, a $y_0 \in \mathbb{R} \backslash \{0\}$, and $\alpha \in \mathbb{R}_+$ such that the infimum

$$\inf_{f \in \mathscr{H}(K)} \left\{ \sum_j w_j |f(x_j)|^2 + |f(t_0) - y_0|^2 + \alpha \|f\|^2_{\mathscr{H}(K)} \right\} \tag{10.9}$$

is strictly positive. Then S is not a interpolation set for (K, X).

Proof. This results follows from Lemma 10.2 and Theorem 10.3 above. We also refer readers to [45]. □

REFERENCES

[1] D. Alpay, V. Bolotnikov, *On tangential interpolation in reproducing kernel Hilbert modules and applications*, [in:] *Topics in interpolation theory*, vol. 95, Oper. Theory Adv. Appl., Birkhäuser, Basel, 1997, 37–68.

[2] D. Alpay, V. Bolotnikov, H. Turgay Kaptanoğlu, *The Schur algorithm and reproducing kernel Hilbert spaces in the ball*, Linear Algebra Appl. **342** (2002), 163–186.

[3] D. Alpay, P. Dewilde, H. Dym, *Lossless inverse scattering and reproducing kernels for upper triangular operators*, [in:] *Extension and interpolation of linear operators and matrix functions*, vol. 47, Oper. Theory Adv. Appl., Birkhäuser, Basel, 1990, 61–135.

[4] D. Alpay, A. Dijksma, J. Rovnyak, H.S.V. de Snoo, *Realization and factorization in reproducing kernel Pontryagin spaces*, [in:] *Operator theory, system theory and related topics*, vol. 123, Oper. Theory Adv. Appl., Birkhäuser, Basel, 2001, 43–65.

[5] D. Alpay, C. Dubi, *Some remarks on the smoothing problem in a reproducing kernel Hilbert space*, J. Anal. Appl. **4** (2006) 2, 119–132.

[6] D. Alpay, H. Dym, *On a new class of structured reproducing kernel spaces*, J. Funct. Anal. **111** (1993) 1, 1–28.

[7] D. Alpay, P.E.T. Jorgensen, *Stochastic processes induced by singular operators*, Numer. Funct. Anal. Optim. **33** (2012) 7–9, 708–735.

[8] D. Alpay, P. Jorgensen, D. Levanony, *A class of Gaussian processes with fractional spectral measures*, J. Funct. Anal. **261** (2011) 2, 507–541.

[9] D. Alpay, P. Jorgensen, D. Levanony, *On the equivalence of probability spaces*, J. Theoret. Probab. **30** (2017) 3, 813–841.

[10] D. Alpay, P. Jorgensen, G. Salomon, *On free stochastic processes and their derivatives*, Stochastic Process. Appl. **124** (2014) 10, 3392–3411.

[11] D. Alpay, D. Levanony, *On the reproducing kernel Hilbert spaces associated with the fractional and bi-fractional Brownian motions*, Potential Anal. **28** (2008) 2, 163–184.

[12] D. Alpay, T.M. Mills, *A family of Hilbert spaces which are not reproducing kernel Hilbert spaces*, J. Anal. Appl. **1** (2003) 2, 107–111.

[13] N. Arcozzi, R. Rochberg, E. Sawyer, *Carleson measures for the Drury-Arveson Hardy space and other Besov-Sobolev spaces on complex balls*, Adv. Math. **218** (2008) 4, 1107–1180.

[14] N. Arcozzi, R. Rochberg, E. Sawyer, *Two variations on the Drury-Arveson space*, [in:] *Hilbert spaces of analytic functions*, vol. 51, CRM Proc. Lecture Notes, Amer. Math. Soc., Providence, RI, 2010, 41–58.

[15] N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. **68** (1950), 337–404.

[16] W. Arveson, *Subalgebras of C^*-algebras. III. Multivariable operator theory*, Acta Math. **181** (1998) 2, 159–228.

[17] R. Balan, M. Begué, C. Clark, K. Okoudjou, *Optimization methods for frame conditioning and application to graph Laplacian scaling*, [in:] *Frames and other bases in abstract and function spaces*, Appl. Numer. Harmon. Anal., Birkhäuser/Springer, Cham, 2017, 27–45.

[18] X. Chen, Y. Hou, *A sharp lower bound on the least signless Laplacian eigenvalue of a graph*, Bull. Malays. Math. Sci. Soc. **41** (2018) 4, 2011–2018.

[19] X. Chang, H. Xu, Shing-Tung Yau, *Spanning trees and random walks on weighted graphs*, Pacific J. Math. **273** (2015) 1, 241–255.

[20] B. Currey, A. Mayeli, *The orthonormal dilation property for abstract Parseval wavelet frames*, Canad. Math. Bull. **56** (2013) 4, 729–736.

[21] M. D'Elia, M. Gunzburger, *The fractional Laplacian operator on bounded domains as a special case of the nonlocal diffusion operator*, Comput. Math. Appl. **66** (2013) 7, 1245–1260.

[22] D. Freeman, D. Poore, A.R. Wei, M. Wyse, *Moving Parseval frames for vector bundles*, Houston J. Math. **40** (2014) 3, 817–832.

[23] A. Ghosal, P. Deb, *Quantum walks over a square lattice*, Phys. Rev. A **98** (2018), 032104.

[24] S. Haeseler, M. Keller, D. Lenz, J. Masamune, M. Schmidt, *Global properties of Dirichlet forms in terms of Green's formula*, arXiv e-prints (2014).

[25] D. Han, W. Jing, D. Larson, P. Li, R.N. Mohapatra, *Dilation of dual frame pairs in Hilbert C^*-modules*, Results Math. **63** (2013) 1–2, 241–250.

[26] D. Han, K. Kornelson, D. Larson, E. Weber, *Frames for undergraduates*, vol. 40, Student Mathematical Library, American Mathematical Society, Providence, RI, 2007.

[27] S. Hersonsky, *Boundary value problems on planar graphs and flat surfaces with integer cone singularities, I: The Dirichlet problem*, J. Reine Angew. Math. **670** (2012), 65–92.

[28] T. Hida, *Quadratic functionals of Brownian motion*, J. Multivariate Anal. **1** (1971) 1, 58–69.

[29] T. Hida, *Brownian Motion*, vol. 11, Applications of Mathematics, Springer-Verlag, New York–Berlin, 1980.

[30] T. Hida, *Stochastic variational calculus*, [in:] *Stochastic partial differential equations and their applications*, vol. 176, Lect. Notes Control Inf. Sci., Springer, Berlin, 1992, 123–134.

[31] R.R. Holmes, H. Huang, Tin-Yau Tam, *Asymptotic behavior of Iwasawa and Cholesky iterations*, Linear Algebra Appl. **438** (2013) 10, 3755–3768.

[32] K. Itô, H.P. McKean, Jr., *Diffusion Processes and Their Sample Paths*, Die Grundlehren der Mathematischen Wissenschaften, Band 125. Academic Press Inc., Publishers, New York, 1965.

[33] K. Itô, *Stochastic Processes: Lectures Diven at Aarhus University*, Springer-Verlag, Berlin, 2004 (reprint of the 1969 original, edited and with a foreword by O.E. Barndorff-Nielsen and K. Sato).

[34] O.G. Jørsboe, *Equivalence or Singularity of Gaussian Measures on Function Spaces*, Various Publications Series, no. 4, Matematisk Institut, Aarhus Universitet, Aarhus, 1968.

[35] P.E.T. Jorgensen, *Harmonic Analysis: Smooth and Non-smooth*, CBMS Regional Conference Series in Mathematics, Conference Board of the Mathematical Sciences, 2018.

[36] P.E.T. Jorgensen, E.P.J. Pearse, *Gel′fand triples and boundaries of infinite networks*, New York J. Math. **17** (2011), 745–781.

[37] P.E.T. Jorgensen, E.P.J. Pearse, *Spectral comparisons between networks with different conductance functions*, J. Operator Theory **72** (2014) 1, 71–86.

[38] P.E.T. Jorgensen, E.P.J. Pearse, *Continuum versus discrete networks, graph Laplacians, and reproducing kernel Hilbert spaces*, J. Math. Anal. Appl. **469** (2019) 2, 765–807.

[39] P.E.T. Jorgensen, S. Pedersen, *Dense analytic subspaces in fractal L^2-spaces*, J. Anal. Math. **75** (1998), 185–228.

[40] P. Jorgensen, S. Pedersen, F. Tian, *Spectral theory of multiple intervals*, Trans. Amer. Math. Soc. **367** (2015) 3, 1671–1735.

[41] P. Jorgensen, S. Pedersen, F. Tian, *Extensions of Positive Definite Functions: Applications and Their Harmonic Analysis*, vol. 2160, Lecture Notes in Mathematics, Springer, Cham, 2016.

[42] P. Jorgensen, F. Tian, *Discrete reproducing kernel Hilbert spaces: sampling and distribution of Dirac-masses*, J. Mach. Learn. Res. **16** (2015), 3079–3114.

[43] P. Jorgensen, F. Tian, *Infinite networks and variation of conductance functions in discrete Laplacians*, J. Math. Phys. **56** (2015) 4, 043506.

[44] P. Jorgensen, F. Tian, *Graph Laplacians and discrete reproducing kernel Hilbert spaces from restrictions*, Stoch. Anal. Appl. **34** (2016) 4, 722–747.

[45] P. Jorgensen, F. Tian, *Nonuniform sampling, reproducing kernels, and the associated Hilbert spaces*, Sampl. Theory Signal Image Process. **15** (2016), 37–72.

[46] P. Jorgensen, F. Tian, *Positive definite kernels and boundary spaces*, Adv. Oper. Theory **1** (2016) 1, 123–133.

[47] P. Jorgensen, F. Tian, *Generalized Gramians: creating frame vectors in maximal subspaces*, Anal. Appl. (Singap.) **15** (2017) 1, 123–135.

[48] P. Jorgensen, F. Tian, *Transfer operators, induced probability spaces, and random walk models*, Markov Process. Related Fields **23** (2017) 2, 187–210.

[49] P. Jorgensen, F. Tian, *Infinite weighted graphs with bounded resistance metric*, Math. Scand. **123** (2018) 1, 5–38.

[50] P.E.T. Jorgensen, Myung-Sin Song, *Infinite-dimensional measure spaces and frame analysis*, Acta Appl. Math. **155** (2018), 41–56.

[51] P.E.T. Jorgensen, Myung-Sin Song, *Markov chains and generalized wavelet multiresolutions*, J. Anal. **26** (2018) 2, 259–283.

[52] V. Kaftal, D.R. Larson, S. Zhang, *Operator-valued frames*, Trans. Amer. Math. Soc. **361** (2009) 12, 6349–6385.

[53] A.N. Kolmogorov, *On logical foundations of probability theory*, [in:] *Probability theory and mathematical statistics*, vol. 1021, Lecture Notes in Math., Springer, Berlin, 1983, 1–5.

[54] A.N. Kolmogorov, Ju.A. Rozanov, *On a strong mixing condition for stationary Gaussian processes*, Teor. Verojatnost. i Primenen **5** (1960), 222–227.

[55] C. Kreuzer, *Reliable and efficient a posteriori error estimates for finite element approximations of the parabolic p-Laplacian*, Calcolo **50** (2013) 2, 79–110.

[56] D.P. Kroese, T. Brereton, T. Taimre, Z.I. Botev, *Why the monte carlo method is so important today*, Wiley Interdisciplinary Reviews: Computational Statistics **6** (2014) 6, 386–392.

[57] G. Kutyniok, K.A. Okoudjou, F. Philipp, E.K. Tuley, *Scalable frames*, Linear Algebra Appl. **438** (2013) 5, 2225–2238.

[58] H.A. Lay, Z. Colgin, V. Reshniak, A.Q.M. Khaliq, *On the implementation of multilevel Monte Carlo simulation of the stochastic volatility and interest rate model using multi-GPU clusters*, Monte Carlo Methods Appl. **24** (2018) 4, 309–321.

[59] M.Z. Nashed, Q. Sun, *Function spaces for sampling expansions*, [in:] *Multiscale signal analysis and modeling*, Springer, New York, 2013, 81–104.

[60] Z. Nashed, *Operator parts and generalized inverses of linear manifolds with applications*, [in:] *Trends in theory and practice of nonlinear differential equations*, vol. 90, Lecture Notes in Pure and Appl. Math., Dekker, New York, 1984, 395–412.

[61] B. Ning, X. Peng, The Randić index and signless Laplacian spectral radius of graphs. Discrete Math. **342** (2019) 3, 643–653.

[62] J. von Neumann, I.J. Schoenberg, *Fourier integrals and metric geometry*, Trans. Amer. Math. Soc. **50** (1941), 226–251.

[63] K.A. Okoudjou, *Preconditioning techniques in frame theory and probabilistic frames*, [in:] *Finite frame theory*, vol. 73, Proc. Sympos. Appl. Math., Amer. Math. Soc., Providence, RI, 2016, 105–142.

[64] K.R. Parthasarathy, K. Schmidt, *Stable positive definite functions*, Trans. Amer. Math. Soc. **203** (1975), 161–174.

[65] V.I. Paulsen, M. Raghupathi, *An introduction to the theory of reproducing kernel Hilbert spaces*, vol. 152, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2016.

[66] I.Z. Pesenson, *Paley-Wiener-Schwartz nearly Parseval frames on noncompact symmetric spaces*, [in:] *Commutative and noncommutative harmonic analysis and applications*, vol. 603, Contemp. Math. Amer. Math. Soc., Providence, RI, 2013, 55–71.

[67] Z. Qiu, M.Y. Li, Z. Shen, *Global dynamics of an infinite dimensional epidemic model with nonlocal state structures*, J. Differential Equations **265** (2018) 10, 5262–5296.

[68] G.J. Rodgers, K. Austin, B. Kahng, D. Kim, *Eigenvalue spectra of complex networks*, J. Phys. A **38** (2005) 43, 9431–9437.

[69] I.J. Schoenberg, *Interpolating splines as limits of polynomials*, Linear Algebra Appl. **52/53** (1983), 617–628.

[70] F.A. Shah, L. Debnath, *Tight wavelet frames on local fields*, Analysis (Berlin) **33** (2013) 3, 293–307.

[71] M. Skopenkov, *The boundary value problem for discrete analytic functions*, Adv. Math. **240** (2013), 61–87.

[72] S. Smale, Ding-Xuan Zhou, *Learning theory estimates via integral operators and their approximations*, Constr. Approx. **26** (2007) 2, 153–172.

[73] B. Sz.-Nagy, C. Foias, H. Bercovici, L. Kérchy, *Harmonic analysis of operators on Hilbert space*, Universitext, Springer, New York, enlarged edition, 2010.

[74] J. Weston, *Statistical learning theory in practice*, [in:] *Empirical Inference*, Springer, Heidelberg, 2013, 81–93.

[75] C. Wickman, K. Okoudjou, *Duality and geodesics for probabilistic frames*, Linear Algebra Appl. **532** (2017), 198–221.

[76] H. Zhang, Y. Xu, J. Zhang, *Reproducing kernel Banach spaces for machine learning*, J. Mach. Learn. Res. **10** (2009), 2741–2775.

Palle E.T. Jorgensen
palle-jorgensen@uiowa.edu

The University of Iowa
Department of Mathematics
Iowa City, IA 52242-1419, USA

Feng Tian
feng.tian@hamptonu.edu

Hampton University
Department of Mathematics
Hampton, VA 23668, USA